Relevant impact of central pathology review on nodal classification in individual breast cancer patients


1Department of Internal Medicine, Division of Medical Oncology, Maastricht University Medical Centre, Maastricht; 2Department of Epidemiology, Biostatistics and HTA, Radboud University Medical Centre Nijmegen, Nijmegen; 3Department of Pathology, University Medical Centre Utrecht, Utrecht; 4Department of Internal Medicine, Division of Medical Oncology, Erasmus Medical Centre Dand Hoed Cancer Centre, Rotterdam; 5Department of Surgery, Erasmus Medical Centre-Daniel den Hoed Cancer Centre, Rotterdam; 6School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht; 7Department of Internal Medicine, Division of Medical Oncology, GROW—School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht; 8Department of Pathology, University Medical Centre Utrecht, Utrecht; 9Department of Internal Medicine, Division of Medical Oncology, Leiden University Medical Centre, Leiden; 10Department of Surgery, The Netherlands Cancer Institute, Amsterdam; 11Department of Internal Medicine, Division of Medical Oncology, Department of Surgery, Erasmus Medical Centre-Daniel den Hoed Cancer Centre, Rotterdam; 12Department of Pathology, Radboud University Medical Centre Nijmegen, Nijmegen, The Netherlands

Received 15 October 2011; revised 2 February 2012; accepted 8 February 2012

Background: In the MIRROR study, pN0(i +) and pN1mi were associated with reduced 5-year disease-free survival (DFS) compared with pN0. Nodal status (N-status) was assessed after central pathology review and restaging according to the sixth AJCC classification. We addressed the impact of pathology review.

Patients and methods: Early favorable primary breast cancer patients, classified pN0, pN0(i +), or pN1(mi) by local pathologists after sentinel node procedure, were included. We assessed the impact of pathology review on N-status (n = 2842) and 5-year DFS for those without adjuvant therapy (n = 1712).

Results: In all, 22% of the 1082 original pN0 patients was upstaged. Of the 623 original pN0(i +) patients, 1% was downstaged, 26% was upstaged. Of 1137 patients staged pN1mi, 15% was downstaged, 11% was upstaged. Originally, 5-year DFS was 85% for pN0, 74% for pN0(i +), and 73% for pN1mi; HR 1.70 [95% confidence interval (CI) 1.27–2.27] and HR 1.57 [95% CI 1.16–2.13], respectively, compared with pN0. By review staging, 5-year DFS was 86% for pN0, 77% for pN0(i +), 77% for pN1mi, and 74% for pN1 +.

Correspondence to: Prof. V. C. G. Tjan-Heijnen, Department of Internal Medicine, Division of Medical Oncology, Maastricht University Medical Centre, P. Debyeelaan 25, 6229 HX Maastricht, The Netherlands. Tel: +31-43-3877025; Fax: +31-43-3876281. E-mail: vcg.tjan.heijnen@mumc.nl

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com.
Conclusion: Pathology review changed the N-classification in 24%, mainly upstaging, with potentially clinical relevance for individual patients. The association of isolated tumor cells and micrometastases with outcome remained unchanged. Quality control should include nodal breast cancer staging.

Key words: breast cancer, pathology review, sentinel node

introduction

The axillary lymph node status remains one of the most important prognostic factors in breast cancer. With the introduction of the sentinel node (SN) procedure, the pathologic work-up is intensified, resulting in an increased detection frequency of small nodal metastases [1]. Until recently, the significance of isolated tumor cells (classified as pN0(i +), with deposits < 0.2 mm) or micrometastases (classified as pN1mi, with deposits > 0.2 to < 2.0 mm) in relation to clinical outcome was unclear. The MIRROR (Micrometastases and isolated tumor cells: relevant and robust) study was the first large breast cancer cohort study on the relevance of pN0(i +) and pN1mi in the SN era [2]. It was found that patients with favorable early-stage breast cancer and isolated tumor cells or micrometastases in regional lymph nodes, not having received adjuvant systemic therapy, had a reduced disease-free survival (DFS) at 5 years compared with patients with node-negative disease. pN0(i +) as final nodal status (N-status) was shown to be equally independently associated with DFS as pN1mi. Patients with pN0(i +) or pN1(mi) who received adjuvant systemic therapy had a better DFS. In the MIRROR study, all available original slides of the SNs and of positive nodes obtained from patients who underwent an axillary lymph node dissection (ALND) were centrally reviewed by three pathologists specialized in breast cancer. Restaging according to the sixth AJCC classification took then place [2]. In daily practice, however, treatment decisions are based on the pathology reports assessed by local pathologists. In the present analysis, we investigated the difference in the final lymph node status according to the review pathologists versus the local pathologists and its association with DFS per subgroup in the patients without adjuvant systemic therapy.

methods

patients

Women with invasive breast cancer, who had undergone an SN biopsy between 1998 and 2005, were identified from the Netherlands Cancer Registry. Only women with favorable primary tumor characteristics (i.e. tumors < 1 cm in diameter, irrespective of grade, or tumors > 1 to < 3 cm, grade 1 or 2) were selected. From these women, only those with a final N-status of pN0(i +) or pN1mi, detected on microscopic examination of the SN and non-SNs (in patients who subsequently underwent ALND) were included. Furthermore, a group of patients with low-risk, node-negative disease, who underwent breast surgery and an SN biopsy with or without an ALND in the period from 2000 through 2001 was included. For all patients, exclusion criteria were neoadjuvant chemotherapy, adjuvant systemic therapy in node-negative disease (pN0(i −)), bilateral breast cancer, or a history of cancer. Detailed information on methods and results were reported by de Boer et al. [2]. All hospitals (n = 113) and pathology laboratories (n = 60) in the Netherlands participated.

pathology

All available original slides from the SNs and of non-SNs obtained in the ALND were reviewed by at least one of three pathologists (CHMvD, PJvD, PB) specialized in breast cancer and experienced in SN pathology. These included slides stained with hematoxylin and eosin (H&E), slides stained for immunohistochemical analysis, and frozen sections. The overall agreement between the review pathologists regarding classification as node negative [pN0/pN0(i +)] or node positive [pN1mi/pN1a] was 98% (κ = 0.96; 95% confidence interval (CI) 0.90–1.02) and the overall agreement in classification according to the sixth AJCC Cancer Staging Manual was 83% (κ = 0.75; 95% CI 0.65–0.85). Almost all participating pathology laboratories used a protocol in which the SN was serially sectioned at least every 150 μm and at a minimum of three levels, with the use of keratin immunohistochemical staining if the H&E staining was negative. The non-SNs were macroscopically sectioned every 2–5 mm and one section per slice was stained with H&E. Additional sections were not obtained, and immunohistochemical staining was not routinely carried out. Classification was done according to the sixth edition of the AJCC Cancer Staging Manual [3].

study objectives

The primary study objective was to assess the difference in the final N-status as determined by central pathology review compared with the final N-status as determined by the local pathologists. The secondary study objective was to assess the 5-year DFS rate according to the original and the revised final N-status, overall and in patients per subgroup (pN0, pN0(i +), pN1mi, and pN1 +). To rule out the influence of adjuvant systemic therapy, this analysis was carried out in patients not treated with adjuvant systemic therapy.

statistical analysis

The chi-square test for trend was used to assess baseline differences between ordinal variables, and analysis of variance was used for continuous variables. To assess the rate of reproducibility between the original and review pathologists, the kappa score was assessed. The Kaplan–Meier method was used to estimate the 5-year rate of DFS in patients who did not receive adjuvant systemic therapy. The period of DFS was defined as the interval from the date of diagnosis to locoregional recurrence, distant metastases, contralateral invasive breast cancer or ductal carcinoma in situ, another malignant condition, or death from any cause, whichever occurred first. For patients who remained alive and disease free, data were censored at the date of the last contact. A Cox proportional hazards model was used to compare the cohorts and to adjust for known prognostic clinical and pathological variables. In this analysis, we included age at diagnosis, tumor size, tumor grade, and hormone receptor status (positive: ER and/or PgR positive). All reported p values are two-sided, and confidence intervals are at the 95% level.

results

patients

At baseline, 2842 breast cancer patients with favorable primary tumor characteristics were selected for this analysis; 135 of

Original articles

Annals of Oncology
these patients were not included in the primary analyses of the MIRROR study because they were classified as having macrometastatic disease (pN1 + ) after central pathology review [2]. Of the 2842 patients, 1082 had originally been classified as pN0, 623 patients as pN0(i + ), and 1137 patients as pN1mi. Baseline characteristics are shown for patients classified according to the review pathology (supplemental Table S1, available at Annals of Oncology online). Overall, there were no major differences in baseline characteristics per nodal category. When patients were classified according to the local pathology also no major differences were seen.

**Final N-status: original versus centrally reviewed**

In Figure S1 (available at Annals of Oncology online), the final N-status as assessed by the local pathologist (original) and after central pathology review, respectively, is shown. Of the 1082 patients with originally pN0, 195 patients (18%) were restaged after central review as pN0(i + ), 29 patients (3%) as pN1mi, and 5 patients (0.5%) as pN1 + . Of the 623 patients with originally pN0(i + ), central review assessed 3 patients (0.5%) as pN0, 158 patients (25%) as pN1mi, and 4 patients (0.6%) as pN1 + . And, of the 1137 patients with originally pN1mi, 166 patients (15%) were considered having pN0(i + ) and 126 patients (11%) pN1 + after review. Overall, central pathology review changed the N-classification in 686 of the 2842 patients (24%, 95% CI 23% to 26%), consisting of downstaging in 6% and upstaging in 18% of the patients. To assess the rate of reproducibility between the local and review pathologists, the kappa score was assessed. The kappa score was 0.69 (95% CI 0.67–0.71).

To rule out the influence of dedication of the local pathologists, we carried out a funnel plot. The number of patients included per laboratory ranged between 1 and 155. Of the 60 involved pathology laboratories, the majority of the laboratories performed comparable with each other, and only a few scored very different showing that the overall discrepancy between the local pathologists and review pathologists was comparable in almost all laboratories and not just caused by a few laboratories (Figure 1).

As the sixth edition of the AJCC Cancer Staging Manual was introduced in 2002 [3], we carried out separate analyses for the patients diagnosed before and since 2003. Of the 1786 patients diagnosed before 2003, 438 (25%, 95% CI 23% to 27%) were restaged after central review, and of the 1056 patients diagnosed since 2003, 248 (23%, 95% CI 21% to 26%) were differentially staged after central pathology review. So, before and after the introduction of the sixth edition of the AJCC Cancer Staging Manual, the percentage of restaged patients remained similar.

**DFS rates**

Overall, the 5-year DFS rate in the patients not having been treated with adjuvant systemic therapy (n = 1751) was 82%.

According to the original final N-status, the unadjusted 5-year DFS rate was 85% for patients with pN0 (n = 1068), 74% for patients with pN0(i + ) (n = 364), and 73% for those with pN1mi (n = 319) (Figure 2). The difference in 5-year DFS rate was statistically different between pN0(i + ) as compared with pN0 (P = 0.0002). The difference in 5-year DFS rate was also statistically different between pN1mi as compared with pN0 (P = 0.0008).

After correction for age at diagnosis, tumor size, tumor differentiation grade, and hormone receptor status an increased risk of disease relapse remained substantial for patients with final pN0(i + ) with an adjusted hazard ratio of 1.70 (95% CI 1.27–2.27) as compared with patients with final pN0 status (Table 1). For patients with final pN1mi, the adjusted hazard ratio was 1.57 (95% CI 1.16–2.13) as compared with final pN0.

Previously, we reported part of the results after central review [2]. After review, the unadjusted 5-year DFS rate was 86% for patients with restaged pN0 (n = 856), 77% for patients with pN0(i + ) (n = 513), and 77% for patients with pN1mi (n = 343) (Figure 3). After correction for age at diagnosis, tumor size, tumor differentiation grade, and hormone receptor status an increased risk for disease-related events was seen for patients with restaged pN0(i + ) with a hazard ratio of 1.50.
As compared with patients with restaged pN0 status. For patients with restaged pN1mi, versus restaged pN0, the hazard ratio was 1.56 (95% CI 1.15–2.13). Patients who were excluded from the previous analysis, because of presence of macrometastases (pN1 +), had an unadjusted 5-year DFS rate of 74% (n = 39). The 5-year DFS rate was lowest in the pN1 + group as expected; however, the number of patients in this group was too low for a formal statistical analysis.

**discussion**

Previously, we reported that final lymph node status pN0(i +) and pN1mi, assessed after central pathology review, in patients with favorable early-stage breast cancer who did not receive adjuvant systemic therapy were associated with a reduced 5-year DFS [2]. In the current analysis, we showed that central pathology review changed the N-classification in 24% of the patients, mainly consisting of upstaging in 18% of the patients. Notably, on a population level, the association of isolated tumor cells and micrometastases with breast cancer outcome remained unchanged. That is, pN0(i +) or pN1mi were also associated with reduced survival when N-status was based on local pathology reports. However, restaging may have clinical relevance for an individual patient as restaging may result in a final N-status making the patient eligible or ineligible for adjuvant systemic therapy with consequences for final outcome.

The discordance among pathologists with respect to breast SN has been shown before in several small studies. Roberts et al. offered 25 slides (containing 1 macrometastatic, 19 micrometastatic, and 5 negative slides) to 10 pathologists with vast experience in breast SNs [4]. Only 3 of the 25 slides (12%) were interpreted in the same way by all 10 pathologists. These included one negative slide and two slides with the most abundant tumor burden. As the number of tumor cells present in an SN gradually decreased, the number of pathologists scoring the case correctly decreased. The European Working Group for Breast Screening Pathology (EWGBSP) assessed the reproducibility of interpreting isolated tumor cells and micrometastases by showing digital images of 50 cases with low-volume SN involvement twice to their members [5]. In the first round, the overall kappa score was only 0.39. The possible causes for this low reproducibility were discussed and methods for improvement were adopted after consultation of experts of the International Committee against Cancer. In the second evaluation round the kappa score improved to a moderate 0.49. Turner et al. showed an equally low kappa score of 0.49 when the digital images of 56 low-volume involved nodes were reviewed by 6 experienced breast pathologists [6]. However, after a training program that included definitions, guidelines, and example images, the kappa score improved to 0.95. The EWGBSP interpretation of AJCC SN staging was however better predictive of non-SN involvement than the Turner et al. interpretation [7, 8]. Adherence to category definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>pN0(i +) versus pN0</th>
<th>pN1mi versus pN0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Node-negative (pN0)</td>
<td>1.00</td>
<td>–</td>
</tr>
<tr>
<td>Isolated tumor cells (pN0(i +))</td>
<td>1.70</td>
<td>1.27–2.27</td>
</tr>
<tr>
<td>Micrometastases (pN1mi)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Age (year)a</td>
<td>1.24</td>
<td>1.03–1.50</td>
</tr>
<tr>
<td>Tumor sizeb</td>
<td>1.00</td>
<td>–</td>
</tr>
<tr>
<td>Tumor grade</td>
<td>1</td>
<td>1.61</td>
</tr>
<tr>
<td>2</td>
<td>2.65</td>
<td>1.52–4.63</td>
</tr>
<tr>
<td>3</td>
<td>1.34</td>
<td>0.87–2.08</td>
</tr>
</tbody>
</table>

a Modeled as continuous variable.
b HR when tumor diameter doubles.

HR, hazard ratio; CI, confidence interval; ITC, isolated tumor cells (pN0(i +)); MI, micrometastases (pN1mi); ER, estrogen receptor; PgR, progesterone receptor.

Figure 3. Kaplan–Meier 5-year disease-free survival curves according to the review N-status.
supplemented by explanatory texts in the staging manual of the new edition of the TNM (tumor–node–metastasis) has shown to lead to moderate (kappa score 0.56 in the isolated tumor cells category) and substantial (kappa score 0.62 in the micrometastases category) reproducibility [9]. The difficulty of assessing low-volume involved lymph nodes was also shown in a larger French study [10]. The slides of 337 operable breast cancer patients with one low-volume involved lymph node were examined by three senior pathologists. It showed that the distinction between isolated tumor cells and micrometastases was often difficult, depending on which definition of measuring size was used (size of area with close clusters or size of largest cluster) and depending on whether the AJCC or International Union Against Cancer (UICC) definitions were used. With the AJCC definition, the classification was uncertain in 40%–45% of patients. The authors further mentioned that there was no significant difference in metastasis-free survival between pN0(i +) and pN1mi patients whatever method of sizing or classification was used. In the MIRROR study, we also showed that the DFS was not significantly different between pN0(i +) and pN1mi, but both subgroups had a worse outcome compared with pN0 (i −) [2].

In our analysis, the discrepancy between the local pathologists and review pathologists was also high resulting in a change of the final N-status in 24% of the patients (kappa score of 0.69). Since our study population comprised patients diagnosed between 1998 and 2005, one might argue that before the introduction of the sixth edition of the AJCC Cancer Staging Manual, it was not mandatory to differentiate between micrometastases and isolated tumor cells [3]. However, in subanalyses considering the patients diagnosed before and since 2003, the discordance between the original and review reports was sustained. Another hypothesis could be the dedication of the pathologists. On the other hand, of the 60 involved pathology laboratories, the majority of the laboratories performed comparable with each other, and only a few scored different (Figure 1). The laboratories with the higher number of patients (ranging between n = 82 and n = 155) had disagreement rates of 12%, 19%, 20%, 21%, 24%, 24%, 26%, 27%, 28%, 34%, and 36%. Although the 36% disagreement rate was in the largest group (n = 155), exclusion of this group still led to an overall disagreement rate of 23.4%.

The overall agreement between our review pathologists regarding classification as node-negative [pN0/pN0(i +)] or node-positive [pN1mi/pN1a] was 98% (κ = 0.96; 95% CI 0.90–1.02) and overall agreement regarding the sixth AJCC Cancer Staging Manual was 83% (κ = 0.75; 95% CI 0.65–0.85). We believe that this high overall agreement can be explained by dedication and training. The 5-year DFS was better for all subgroups considering the classification as assessed by the review pathologists. These differences can be explained by stage migration, also called the Will Rogers phenomenon [11]. The improved outcome indicates that the 24% upstaging as proposed by the review pathologists was indeed correct and of potential clinical relevance.

On a population level, the association of the N-status with clinical outcome remained unchanged. However, on an individual level, the restaging of patients can be of clinical relevance in several ways. When isolated or even no tumor cells in the SN were incorrectly reported, while in fact micro- or macrometastases were present, patients could inappropriate be withheld axillary treatment [12]. Also 15% of the pN1mi patients were downstaged to pN0(i +), a stage in which it is considered acceptable to withhold axillary treatment, indicating that some of these women have a risk of being overtreated including the increased risk of lymph edema and shoulder dysfunction and without gaining further survival benefit. Probably more important, considering our prior report [2], less patients may be treated with systemic therapy because of understaging. In our study population, 18% (n = 517) was upstaged. If the criteria to advise systemic therapy start with an N-status of at least pN0(i +), 8% (n = 229) of the patients would not have received systemic therapy because of understaged N-status. Recently, the criteria indicative for systemic therapy regarding other patient and tumor characteristics were broadened in the Dutch guidelines. However, still about half of these patients would not have received systemic therapy based on the primary tumor characteristics only (data not shown). The recent publication on the finding of occult metastases in the NSABP B-32 trial also showed a small, though statistically significant adverse outcome in patients with occult sentinel lymph node metastases [13]. In this study, it was mentioned that adjuvant systemic therapy was more often administered in the group with occult metastases and it was therefore suggested that perhaps other factors known at the start of therapy probably correlate with the presence of occult metastases. It can also be hypothesized that without the administration of adjuvant systemic therapy the difference in adverse outcome will be larger.

Because of the observation that pathology review altered the N-status in 24% of the patients, we believe that two important recommendations should be made. First, additional studies concerning this issue are eagerly awaited confirming the importance of central review. To interpret forthcoming data reliably, we encourage our colleague investigators to include central pathology review. Our second recommendation regards the quality control of pathology laboratories. We believe that additional training or even specialization of pathologists, in addition to published guidelines, is necessary, as well as inclusion of SN staging in quality control procedures of pathology laboratories. An example of a training instrument with the publication of new guidelines could be an E-learning course comprising digital examples.

acknowledgements

In part presented at San Antonio Breast Cancer Symposium 2010 (PD06-04). We thank Wim A. J. G. Lemmens for his assistance with statistical analyses.

funding

The Netherlands Organization for Health Research and Development (ZonMw 945-06-059).
Correlation of treatment-emergent adverse events and clinical response to endocrine therapy in early breast cancer: a retrospective analysis of the German cohort of TEAM

P. Hadji1,*, D. G. Kieback2, J. Tams3, A. Hasenburg4 & M. Ziller1

1Departments of Gynaecological Endocrinology, Reproductive Medicine and Osteoporosis, University Hospital of Gessen and Marburg, Marburg; 2Department of Obstetrics and Gynaecology, Eliandtkliniken, Meissen; 3Biostatistics Department, ICRC-Weyer GmbH, Berlin; 4Department of Gynaecology and Gynaecologic Oncology, University Hospital Freiburg, Germany

Received 23 August 2011; revised 13 January 2012; accepted 24 January 2012

Background: Previous studies have suggested a correlation between the occurrence of vasomotor or joint symptoms during tamoxifen or aromatase inhibitor treatment and improved clinical response.

Patients and methods: A retrospective analysis of the German cohort of the Tamoxifen Exemestane Adjuvant Multinational (TEAM) trial was carried out to assess disease-free survival (DFS) and overall survival (OS) in patients with and without arthralgia/myalgia and/or menopausal symptoms during adjuvant endocrine treatment.

Results: A total of 1502 patients were included; 739 patients received tamoxifen followed by exemestane and 763 received exemestane. Patients reporting arthralgia/myalgia and patients reporting menopausal symptoms during endocrine treatment had significantly longer OS and DFS than those not reporting these events. The effect on OS was irrespective of treatment. DFS was significantly improved in exemestane-treated patients reporting arthralgia/myalgia or those reporting menopausal symptoms versus those not reporting these events. This effect on DFS was not observed in patients receiving sequential treatment. A combined analysis of patients reporting either

*Correspondence to: Prof. P. Hadji, Department of Gynaecological Endocrinology, Reproductive Medicine and Osteoporosis, Philipps-University of Marburg, Baldingerstrasse, 35033 Marburg, Germany. Tel: +49-6421-586-6485; Fax: +49-6421-586-7070; E-mail: hadji@med.uni-marburg.de

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com.