Increased Lymph Node Evaluation With Colorectal Cancer Resection

Does It Improve Detection of Stage III Disease?

Sachin S. Kukreja, MD; Enrique Esteban-Agusti, MD; Jose M. Velasco, MD; Tina J. Hieken, MD

**Hypothesis:** Evaluation of 12 or more lymph nodes (LNs) with colorectal cancer (CRC) resection may not improve detection of stage III disease.

**Design:** Retrospective review after intervention.

**Setting:** Community teaching hospital.

**Patients:** We evaluated 701 consecutive operative CRC cases ascertained from our Cancer Registry.

**Intervention:** Patients undergoing resection before (n=553) a multidisciplinary initiative emphasizing the importance of LN counts were compared with those undergoing operation afterward (n=148).

**Main Outcome Measures:** Number of LNs evaluated, proportion of patients with stage III disease, and proportion of patients with N1 vs N2 disease.

**Results:** Demographic, tumor, and treatment variables were similar for both groups, except for younger age, fewer white patients, and more laparoscopic resections in the late period. Lymph node counts increased from a mean (SEM [median]) of 12.8 (0.3 [12]) to 17.3 (0.7 [16]) ($P<.001$), with 53.0% of the early vs 71.6% of the late patients having at least 12 LNs examined. The proportion diagnosed as having stage III CRC was 204 of 553 (36.9%) for the early group vs 48 of 148 (32.4%) for the late group ($P=.31$). Among patients with positive LNs, the distribution of N1 and N2 disease was unchanged (early, 50.5% N1 and 49.5% N2; late, 54.2% N1 and 45.8% N2; $P=.54$).

**Conclusions:** Increased LN retrieval does not identify a greater number of patients with stage III CRC nor does it increase the proportion of patients with positive LNs with N2 disease. Our data suggest that harvest of at least 12 LNs as a quality or performance measure appears unfounded.

**Arch Surg.** 2009;144(7):612-617

**COLORECTAL CANCER (CRC) is the third most common cancer among both male and female patients and the third leading cause of cancer-related death in the United States.** Of the estimated 148,000 patients who will be diagnosed as having CRC in 2008, more than 80% will have locoregional disease at the time of diagnosis and will be offered potentially curative surgery. Since publication of the initial staging classification for CRC by Dukes, regional lymph node (LN) status has been recognized as the most powerful prognostic factor for recurrence and overall survival for these patients. Accurate LN staging also is important for determining prognosis and the need for adjuvant chemotherapy. In addition, lymphadenectomy may be therapeutic; several studies have shown a positive association between the number of LNs removed and survival for patients with negative and positive LNs.

The concept of a threshold for a minimum acceptable number of LNs (≥12) with CRC resection was first introduced in 1990 by the World Congress of Gastroenterology. More recently, this benchmark has been adopted as a quality measure for surgical practice by multiple organizations, including the American College of Surgeons Commission on Cancer, the American Society of Clinical Oncology, the College of American Pathologists, and the National Comprehensive Cancer Network, and several health insurance providers. Controversy remains as to whether adoption of this standard will result in the intended improvement in patient outcome and whether it is an appropriate measure of quality of care.

Two recent, large population-based studies found that most patients with CRC (51%-63%) do not undergo an “adequate” resection with evaluation of at least 12 LNs. Using the benchmark of evaluation of at least 12 LNs in at least 75% of...
cases to evaluate hospital performance, another National Cancer Data Base study found adherence to that benchmark in only 38% of institutions in 2004 through 2005. In some studies, evaluation of a greater number of LNs has been associated with an increased likelihood of identifying patients with LN metastases. Other investigators have found no association between the number of LNs evaluated and detection of stage III disease.

We previously evaluated factors associated with variation in LN counts with CRC resections in our institution. In that study, the number of LNs evaluated was independent of the individual surgeon, surgical specialty, case volume, and intent of the procedure (curative vs palliative) but varied with the individual pathologist as well as with patient age, comorbid illness, and tumor stage and location. Other studies also have found that the surgeon is not the main variable with respect to LN yield. These findings prompted us to formulate an institutional initiative to increase reported LN counts with CRC resections. We decided on a multidisciplinary approach spearheaded by our institution’s cancer committee, which has representation from all specialties diagnosing and treating cancer. This led us to investigate whether (1) we could effect such a change and (2) increased LN retrieval would increase the proportion of patients with stage III CRC and/or the number of positive LNs per case, findings with the greatest potential effect on patient prognosis and selection of adjuvant therapy.

### METHODS

In late 2004, we began a multidisciplinary institutional initiative intended to increase the reported number of LNs removed with CRC resections. We discussed unacceptably low LN counts reported with our CRC resections and reviewed the rationale for increased LN evaluation at our multidisciplinary cancer committee meeting. Our cancer liaison physician (T.J.H.) and cancer committee pathologist initiated a program of institutional awareness of these issues. Our pathologists modified their technique of LN assessment after our intervention. They used a formalin-based fat-clearing solution containing glacial acetic acid in an effort to improve LN yield. The LNs larger than 4 mm that were not grossly identified, the blocks were reevaluated.

After approval from our institutional review committee, we then evaluated 701 consecutive operative CRC cases from January 1, 1996, through December 31, 2007, ascertained from our cancer registry. Data were confirmed by individual review of pathology reports, operative notes, and medical records. For the purpose of evaluating the effect of our cancer committee initiative, we compared early (before January 1, 2003, and our institution-wide intervention) and late (January 1, 2003, and later) patient groups. There were 553 patients in the early (preintervention) group and 148 patients in the late (postintervention) group.

Data are presented as proportions (percentages) or as mean (SEM) unless otherwise stated. We used the SAS statistical package (SAS Institute Inc, Cary, North Carolina) for data analysis. We used the χ² or the Fisher exact test for analysis of categorical variables and 1-way analysis of variance, the paired t test (for normally distributed data), and the Mann-Whitney test (for nonparametric data) for continuous variables. We considered P < .05 to be significant.

### RESULTS

The median patient age was 74 (mean, 71.5 [0.5]; range, 26-100) years. Women constituted 51.4% of the patients. Demographic, tumor, and treatment variables were similar for the early and late groups, as shown in Table 1, with the exception of younger age (median age, 70 vs 72 years), fewer white patients (81.1% vs 95.8%), and increased use of laparoscopic techniques (44.6% vs 7.2%) in the late period. For the entire study population, use of a laparoscopic technique (106 cases) was associated with retrieval of more LNs (mean, 16.1 [0.8]) than was an open approach (595 cases; mean, 13.4 [0.3] LNs). The mean number of LNs evaluated in the late period of the study, when most of the laparoscopic resections were performed, was 18.9 (1.1 [median, 22]) for the laparoscopic approach and 16 (1 [median, 22]) for the open approach (P = .06). The difference in the relative proportions of patients by T stage was not significant between the early and late groups.

Overall LN counts increased from a mean of 12.8 (0.3 [median, 12]) to 17.3 (0.7 [median, 16]) (P < .001), with 53.0% of patients in the early period vs 71.6% in the late period having at least 12 LNs examined (P < .001). The data were stratified by T stage for the mean number of LNs evaluated, the number of positive LNs per case, and the proportion of patients with stage III disease and are summarized in Table 2. The proportion of patients with stage III disease was unchanged, with 204 of 553 (36.9%) of the early case patients and 48 of 148 (32.4%) of the late case patients having positive LNs (P = .31). The mean number of positive LNs and the proportion of patients with positive LNs with N2 disease (>3 positive LNs) did
comes.8,31 Although not yet a major factor affecting scrutiny.32,33 Although LN counts with CRC resection seem to be considered mainly a measure of surgical quality, multiple variables contribute to the number of LNs examined. These include variation in pathology department processing procedures, technique, diligence, and patient and tumor factors.8,10,26-28,34,35 In fact, few data exist that justify the use of increased numbers of LNs as a quality measure that will improve cancer staging, use of chemotherapy, or patient survival.15 Surgical resection with en bloc removal of regional mesenteric lymph nodes is the mainstay of treatment of CRC. Adjuvant chemotherapy is advised in the presence of stage III disease owing to the elevated risk of recurrence and in some high-risk patients with stage II disease and adverse prognostic features. Such treatment guidelines emphasize the importance of accurate LN staging.3 Some authors have advised consideration of chemotherapy for patients with stage II disease in whom fewer than 12 LNs were examined.14 Others have suggested excluding patients with CRC and fewer than 12 LNs examined from surgical adjuvant clinical trials.10 In this study, we evaluated the impact of a multidisciplinary initiative on increasing LN evaluation in patients with CRC in an effort to increase identification of patients with LN-positive disease. After this intervention, our pathologists identified a statistically significantly greater number of LNs with our CRC resections. Before our intervention, 53% of patients underwent an adequate resection (≥12 LN). Our baseline adherence was higher than has been reported in review series that used the National Cancer Data Base or the Surveillance, Epidemiology, and End Results program.16-18 The LN count increased significantly after our intervention; nearly three-quarters of our late cases included assessment of at least 12 LNs. Laparoscopic techniques were used more frequently for the late group in our study, not increase from the early to late period (early, 50.5% N1 and 49.5% N2; late, 54.2% N1 and 45.8% N2). By overall tumor stage at diagnosis, the mean LN count increased significantly for each group of patients from the early to the late period. For patients with stage I disease, the mean LN count increased from 9.8 (0.6) to 14.7 (1.2) (P < .001); for patients with stage II disease, from 14.4 (0.7) to 18.9 (1.3) (P = .002); and for patients with stage III disease, from 15.1 (0.7) to 18.4 (1.3) (P = .03).

### Table 2. Comparison of LN Data and Proportion of Patients With Stage III CRC in the Early (Preintervention) and Late (Postintervention) Periods

<table>
<thead>
<tr>
<th>Variable</th>
<th>Early Group (n=553)</th>
<th>Late Group (n=148)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of LNs, mean (SEM)</td>
<td>12.8 (0.3)</td>
<td>17.3 (0.7)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No. of positive LNs, mean (SEM)</td>
<td>1.7 (0.1)</td>
<td>1.5 (0.3)</td>
<td>.59</td>
</tr>
<tr>
<td>Stage III, No. (%)</td>
<td>204/553 (36.9)</td>
<td>48/148 (32.4)</td>
<td>.31</td>
</tr>
<tr>
<td>T1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of LNs, mean (SEM)</td>
<td>8.5 (0.8)</td>
<td>12.7 (1.4)</td>
<td>.01</td>
</tr>
<tr>
<td>No. of positive LNs, mean (SEM)</td>
<td>0.04 (0.02)</td>
<td>0</td>
<td>.35</td>
</tr>
<tr>
<td>Stage III, No. (%)</td>
<td>3/75 (4.0)</td>
<td>0/22 (0)</td>
<td>.21</td>
</tr>
<tr>
<td>T2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of LNs, mean (SEM)</td>
<td>12.2 (0.9)</td>
<td>15.4 (1.9)</td>
<td>.14</td>
</tr>
<tr>
<td>No. of positive LNs, mean (SEM)</td>
<td>0.8 (0.2)</td>
<td>0.9 (0.5)</td>
<td>.75</td>
</tr>
<tr>
<td>Stage III, No. (%)</td>
<td>15/85 (17.6)</td>
<td>5/19 (26.3)</td>
<td>.40</td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of LNs, mean (SEM)</td>
<td>14.3 (0.5)</td>
<td>19.2 (0.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No. of positive LNs, mean (SEM)</td>
<td>2.1 (0.2)</td>
<td>2.2 (0.4)</td>
<td>.84</td>
</tr>
<tr>
<td>Stage III, No. (%)</td>
<td>154/319 (48.3)</td>
<td>38/81 (46.9)</td>
<td>.83</td>
</tr>
<tr>
<td>T4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of LNs, mean (SEM)</td>
<td>13.2 (1.0)</td>
<td>18.5 (1.9)</td>
<td>.02</td>
</tr>
<tr>
<td>No. of positive LNs, mean (SEM)</td>
<td>4.2 (0.7)</td>
<td>2.4 (1.3)</td>
<td>.23</td>
</tr>
<tr>
<td>Stage III, No. (%)</td>
<td>32/45 (71.1)</td>
<td>5/12 (41.7)</td>
<td>.06</td>
</tr>
</tbody>
</table>

Abbreviations: CRC, colorectal cancer; LN, lymph node.

The National Quality Forum recently stated that a minimum of 12 regional LNs (irrespective of tumor stage or location) should be removed and histopathologically examined with CRC resections.30 This benchmark for defining adequate care for patients with CRC has been endorsed by multiple organizations.8-12 Such mandates are intended to increase awareness of the value of LN evaluation; to allow for comparisons between hospitals, programs, surgeons, pathologists, and regions; and to permit surveillance of practice patterns and outcomes.8,31 Although not yet a major factor affecting reimbursement or selection of providers by consumers, these data have been used by health insurance companies to select surgical providers and may be used as a quality improvement measure for monitoring and mediating performance at the hospital and systems levels. Quality measures and outcome data may be used by employers, patients, and insurers for critically evaluating physicians, hospitals, and health plans and for ranking care. Similar quality measures have been used by state health agencies and accrediting bodies to demand that providers publish their performance data for public scrutiny.32,33 Although LN counts with CRC resection

©2009 American Medical Association. All rights reserved.
but this was associated with an increase, not a decrease, in LN counts. The reported data on the association between the laparoscopic vs open surgical approaches and LN yield is variable.\textsuperscript{17,36} Overall, our improvement in LN yield demonstrates the value and impact of communication through a multidisciplinary initiative engaged in adherence to recommended standards and improving quality of care.

As our pathologists examined the resected specimens, presumably with increased scrutiny, they identified a greater number of LNs but not a greater number of positive ones. Various techniques have been described to increase LN harvest. Fat clearance techniques and the use of special fixatives may improve LN yields by as much as 50%.\textsuperscript{37} However, greater scrutiny of mesenteric nodes away from the primary tumor may focus on LNs that are less likely to contain disease. It also has been postulated that, once a positive node is noted, attentiveness to finding more nodes (positive or otherwise) decreases secondary to the perception that staging and treatment may be relatively unaltered.\textsuperscript{13} Use of lymphatic mapping and sentinel node evaluation, although not used regularly in the clinical setting, may allow the pathologist to focus on the LN most likely to harbor metastatic disease, thereby increasing the sensitivity of LN harvest. This may result in improved staging accuracy without identification of more nodes overall, but its role is yet to be fully defined.\textsuperscript{35}

Despite the improved LN analysis in our study, we did not identify more patients with stage III disease. Before our intervention, 36.9% of our patients had LN-positive disease, whereas 32.4% were found to have positive LNs afterward. Analysis by individual T stages also failed to show any statistically significant increase in nodal positivity for the late group. Similarly, we did not see an increase in N2 disease (>3 positive LNs) with increased LN evaluation. These data imply that our success in increasing LN yield did not result in upstaging the disease in our patients.

Some have suggested that increased LN harvest is therapeutic and may improve prognosis by avoidance of understaging and failing to identify those who may benefit from adjuvant systemic therapy.\textsuperscript{3,5,6} This association seems strong in some reviews, but whether it truly improves staging and outcomes remains unclear.\textsuperscript{7,13-15} Some data suggest that very low LN acquisition fails to identify metastatic nodes, resulting in misclassification of stage III cancer as stage II.\textsuperscript{22} Some investigators propose that LN counts much greater than 12 are necessary to prevent missed nodal metastases. Using the INT0089 trial data, Joseph et al\textsuperscript{38} determined that as many as 40 LNs for T3 and 30 for T4 CRC cases need to be evaluated before one can exclude nodal metastases with an 85% certainty. Other investigators have found that adequate LN evaluation is not necessarily associated with identification of more positive nodes and dismiss understaging as a possible cause of the association between LN harvest and survival.\textsuperscript{13,24}

Lymph node yield as a quality measure places a burden on the surgeon to achieve a standard that may not be within his or her control and may not result in improved patient staging or outcome. Adoption of such standards may lead to selective contracting and public reporting. As data regarding performance are compiled, some hospitals or programs may be determined to be “Centers of Excellence” for CRC resection. Often such categorization is opposed by hospitals and surgeons for being unfair and not reflective of the true quality of patient care.\textsuperscript{30} The initiation of an LN minimum as a marker of quality can lead to a varied response by physicians and systems of care. This may create a disparity for future population-based assessments of treatments and outcomes. Should reimbursement patterns change according to adherence to a given standard, efforts to encourage improved outcomes may actually fail. The “carrot or stick” approach has not necessarily been shown to improve performance and may reward only those who had better performance before the initiative.\textsuperscript{40}

The improved outcomes associated with higher LN counts in some studies may not necessarily be a function of clearance of the LNs per se but may be a marker of other aspects of care (eg, other technical aspects of surgical care, pathology reporting, and adjuvant therapy). It would appear advisable to investigate other such potential quality markers. For example, it may be more useful to standardize the assessment by pathologists (the process) and then set an expected standard.\textsuperscript{36} Because prospective randomized trials show that the administration of chemotherapy for stage III disease improves survival, the validity of adjuvant chemotherapy administration to patients with stage III disease as a performance measure seems logical. In contrast, the data correlating LN counts with survival in CRC is observational and may result from unidentified confounding factors.\textsuperscript{33} This implies that the “magic” number of 12 LNs might be reconsidered as the major quality indicator for the treatment of patients with CRC.

To summarize, we successfully used a multidisciplinary initiative to increase LN yield with CRC resections. However, this increased LN analysis did not identify a greater number of patients with stage III disease and did not increase the proportion of patients with positive LNs and N2 disease. Our data suggest that mandatory harvest of a minimum of 12 LNs as a quality indicator or performance measure appears unfounded.

Accepted for Publication: March 22, 2009.

Correspondence: Tina J. Hieken, MD, Department of Surgery, Rush North Shore Medical Center, 9669 N Kenton, Ste 204, Skokie, IL 60076 (thieken@rush.edu).

Author Contributions: Dr Hieken had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Kukreja and Hieken. Acquisition of data: Kukreja, Esteban-Agusti, and Hieken. Analysis and interpretation of data: Kukreja, Velasco, and Hieken. Drafting of the manuscript: Kukreja, Esteban-Agusti, and Hieken. Critical revision of the manuscript for important intellectual content: Kukreja, Velasco, and Hieken. Administrative, technical, and material support: Kukreja, Esteban-Agusti, and Hieken. Study supervision: Velasco and Hieken.

Financial Disclosure: None reported.

©2009 American Medical Association. All rights reserved.
Previous Presentation: This paper was presented at the 116th Annual Meeting of the Western Surgical Association; November 10, 2008; Santa Fe, New Mexico, and is published after peer review and revision. The discussions that follow this article are based on the originally submitted manuscript and not the revised manuscript.

Additional Contributions: Maria Miklin, BA, RHIA, CTR, and Patricia Waldron, CTR, of the Rush North Shore Medical Center Cancer Registry maintained our cancer registry and assisted in retrieval of data and records.

REFERENCES


DISCUSSION

Charles Scoggins, MD, MBA, Louisville, Kentucky: Dr Kukreja and colleagues have provided us with several very interesting observations. First, they clearly demonstrate the phenomenon of pathologist-related LN retrieval. We have all experienced this: when I am not satisfied with the number of nodes reported, I simply call the pathologist and order more nodes. Magically, they are able to find more. This point serves as the main stratification variable in this paper: they “ordered” more nodes and the pathologists responded. This has not, however, led to stage migration or the distribution of N1 and N2 disease. This is simply an observation and appears not to have affected patient care in any meaningful way. In essence, no more patients are eligible for adjuvant chemotherapy because of this change in policy at their hospital.

I have 3 questions for the authors:

1. Given that quality control measures and initiatives are becoming more commonplace, including the real possibility of linkages to reimbursement, what can we as surgeons do to pro-
actively protect our interests, especially when the number of nodes found by the pathologists may not be entirely in our control?

2. Since your hospital was successful in getting your pathologists to find more LNs, how would you propose that the rest of us get our pathologists to likewise change, especially when they do not get better pay for more nodes?

3. Since more nodes did not result in stage migration or more patients who are eligible for adjuvant therapy, what benefit has this policy really provided to your patients?

Dr Velasco: Thank you very much, Dr Scoggins, for your kind comments and your difficult questions. We wanted to look at the quality of care we were providing very early on. In the year 2000, we realized that we needed to get involved as an institution, we needed to take leadership, and we needed to work with other organizations such as the American College of Surgeons, the cancer committee, and the entire institution to be able to determine and define which quality measures were important. Otherwise, someone else would impose them on us. We know Blue Cross/Blue Shield, health maintenance organizations, and preferred provider organizations are active in using quality measures to grade institutions, probably soon will grade individuals, and then will use those parameters for payment for performance. Most of the studies concerning colon resection and number of LNs identified are either retrospective (as ours was), population based, or based on administrative data with information that at times is not all-inclusive. Some of the studies included diverse patients with different stages or had an insufficient number of patients. Some of the databases that have been used do not have enough information about the pathologists or the methodology. We felt that, because of all of these confounding variables, it would be important for us as an institution to look at every case consecutively and then review every chart and every pathology report individually. So the answer to the first question is that we need to run these studies based on data obtained ourselves.

Now, regarding how we can get our pathologists to work with us, in our institution our pathologists are general pathologists; we don’t have any gastrointestinal pathologists. What we needed to do was to involve them in the cancer committee. It really showed them that we all were being scrutinized; therefore, we needed to provide the best patient care and that we needed to examine if we are meeting some of the criteria that have been established as important. We know LN count is important and that the presence of positive LNs confers a negative prognosis. In our institution 16 surgeons perform all procedures, yet 89% of the operations were done by 6 surgeons: 1 colorectal surgeon, 1 surgical oncologist, and 4 general surgeons. Our pathologists wanted to participate; they are proud of their work. The best way is for us to have a collegial relationship and to make them part of the process. In our study, we did not find a migration of nodes into higher stages as the number of nodes identified increased.

Financial Disclosure: None reported.

Call for Papers

Maintenance of Certification

During 2009, we are soliciting manuscripts on maintenance of certification. The editorial board has selected this topic as our year-long theme. Accepted manuscripts will be given priority publication. Please indicate in your cover letter that your manuscript is to be considered for this theme.