Lymph Node Counts as an Indicator of Quality at the Hospital Level in Colorectal Surgery

Jan H Wong, MD, FACS, Sharon S Lum, MD, FACS, John W Morgan, DRPH

BACKGROUND: Substantial evidence suggests that the number of lymph nodes examined in colorectal cancer (CRC) is a powerful predictor of outcomes. However, the lymph node count as a benchmark of quality in CRC is controversial. We sought to examine the impact of lymph node counts on disease-specific survival (DSS) of CRC patients at the hospital level.

STUDY DESIGN: This study used data obtained between 1994 and 2003 from Region 5 of the California Cancer Registry. Hospitals in Region 5 of the California Cancer Registry were stratified according to the median number of nodes examined and grouped according to the median number of nodes examined, <7, 7 to 9, and ≥10. These hospital groups were then evaluated for the frequency of meeting the 12-node threshold, frequency of positive lymph nodes, and DSS at the hospital level.

RESULTS: Median number of nodes examined in group A was 4 (mean 5.6, SD 5.9), in group B was 8 (mean 9.7, SD 8.5), and in group C was 10 (mean 11.3, SD 9.2). In group A, 13.7%, in group B 32.8%, and in group C, 42.8% met the 12-node threshold. The frequency of N1 and N2 disease for group A was 20.7% and 9.1%, 19.7% and 11.1% for group B, and 20.1% and 11.3% for group C (p = 0.12). Five-year DSS was 72.7% for group A, 73.7% for group B, and 76.7% for group C (p = 0.002). DSS survival of N0 patients for group A was 78.6%, 81.5% for group B, and 85.1% for group C (p < 0.0001). There was no statistically significant difference in DSS for N1 (p = 0.18) or N2 (p = 0.90) between the 3 groups.

CONCLUSIONS: Lymph node counts can have value as a benchmark of surgical/pathologic quality in node-negative CRC. These results question the value of lymph node counts as a benchmark of surgical/pathologic quality for node-positive CRC. (J Am Coll Surg 2011;213:226–230. © 2011 by the American College of Surgeons)

En bloc resection of regional lymph nodes is recognized as a critical component of the surgical care of the patient with colorectal cancer (CRC), not only for regional control, but to identify individuals with nodal metastases who might benefit from adjuvant systemic therapy. Considerable evidence now demonstrates that when the patient is the unit of analysis, the number of nodes harvested is a powerful predictor of outcomes in CRC. Patients from whom higher numbers of nodes are harvested have lower late mortality from CRC than patients with fewer nodes examined. The number of nodes harvested is a particularly powerful predictor of outcomes in the node-negative (N0) population.

In 2005, the National Quality Forum convened a technical panel meeting on the Quality of Cancer Care Measures project to review potential CRC care consensus standards. The National Quality Forum used a process to address the strength of the evidence of the quality indicator; the importance of the measure; whether the measure was scientifically sound based on assessment of its specifications, reliability, validity, and adjustments; and the usability and feasibility of the quality indicator, and recommended that a 12-node minimum be used for quality improvement and community-level surveillance.

It remains unknown whether the 12-node minimum recommended by the National Quality Forum for quality improvement and community-level surveillance will im-
prove outcomes in patients with CRC. A recent report by Wong and colleagues evaluating hospital lymph node examination rates and survival after resection for colon cancer. At the patient level, the 12-node threshold was associated with improved survival relative to less than 12 nodes examined, but the number of lymph nodes examined at the hospital level was not associated with staging or patient survival, raising the question of the value of the 12-node threshold as a measure of quality in colon cancer at the hospital level.

Large databases have demonstrated profound variability in the number of lymph nodes currently being harvested. A minority of patients are actually staged with the 12-node threshold that is recommended in CRC. We sought to examine the validity of the node threshold concept as a measure of quality at the hospital level.

METHODS

Data source
This study used data obtained between January 1994 and December 2003 within Region 5 of the California Cancer Registry (R5 CCR), part of the Surveillance Epidemiology and End Results program of the United States. R5 CCR is 1 of 10 population-based cancer registries that encompass the entire California population and serves Mono, Inyo, San Bernadino, and Riverside Counties, with a population of approximately 4 million individuals, the majority of whom reside in urban population centers that surround the cities of Riverside and San Bernadino. The Desert Sierra Cancer Surveillance Program provides cancer surveillance for these 4 counties.

Patient selection
To minimize the potential for confounding factors, we excluded patients with in situ disease and those with non-regional metastatic disease (M1) on the presumption that these individuals might have had less than a complete lymphadenectomy and nodal counts are less relevant to this population of patients. Those patients in whom the actual number of nodes was not documented or was unknown were excluded from analysis. However, individuals in whom the number of nodes identified was recorded as “zero” were considered to have had nodal staging.

Hospital nodal examination rates
Hospitals in R5 CCR that performed a minimum of 1 resection annually for CRC were identified. Hospitals were then characterized according to the median number of nodes harvested. Because only 2 hospitals had a median number of nodes harvested that exceeded the 12-node threshold, we stratified hospitals into those with a median number of nodes harvested of <7, 7 to 9, and ≥10.

Analysis
All patients were restaged using current American Joint Committee on Cancer criteria. One of the primary outcomes was disease-specific survival (DSS). DSS was calculated by Kaplan-Meier method, with significance assessed by the log-rank test. Patients in whom the cause of death was not documented to be CRC were censored. Chi-square was used where appropriate. Statistical analysis was performed using NCSS Statistical Software Systems. The protocol was determined to be exempt from institutional review board approval by the Institutional Review Board of Loma Linda University.

RESULTS
A total of 33 hospitals that performed at least 1 CRC resection in R5 CCR were included in this analysis. Table 1 summarizes the clinical characteristics of patients in each hospital tercile. Patients in the first tercile tended to be older than in the other 2 terciles, patients treated in hospitals in the second tercile had the most advanced (T3 and T4) tumors, and patients in the first tercile had the most early (T1 and T2) tumors.

There was substantial variability in the number of nodes examined at the hospital level. Hospitals in the first tercile examined, on average, only 5.6 nodes, and hospitals in the third tercile examined, on average, 11.3 nodes (Table 1). There was a highly statistically significant difference in the percent of patients who met the 12-node threshold within the hospital terciles. Only 13.7% of patients in the lowest tercile hospitals met the 12-node threshold within the hospital terciles. Only 13.7% of patients in the lowest tercile hospitals met the 12-node threshold for the number of nodes examined, 32.8% of the patients in the second hospitals tercile met the 12-node threshold, and 42.8% of patients in the highest tercile hospitals met the 12 node threshold (p < 0.001).

Despite the variability of the number of nodes examined in these hospitals terciles, there was no statistically significant difference in percentage of node-positive patients across the hospital terciles (Table 1). However, the number of nodes examined at the hospital level was predictive of DSS (Fig. 1). Hospitals in the first tercile had a 5-year DSS of 72.7% and those in the third tercile had a 5-year DSS of 76.7% (p = 0.002). This survival advantage, however, was

Abbreviations and Acronyms
CRC = colorectal cancer
DSS = disease-specific survival
R5CCR = Region 5 of the California Cancer Registry
confined to only those with N0 disease (Fig. 2). We did not observe any difference in DSS survival between hospitals terciles and their node-positive patients. Because of the lower frequency of more advanced primary tumors in hospitals in the first tercile, we examined the outcomes for T3 primary tumors. There was a trend toward a survival advantage in hospitals in the third tercile ($p = 0.057$).

**DISCUSSION**

The number of nodes retrieved is an important prognostic indicator of CRC.\textsuperscript{1,2,7-9} For patients undergoing curative resection of CRC, the DSS is worse for those apparently node-negative individuals in whom fewer lymph nodes are examined. These survival differences are not insignificant and can approximate the improved survival that has been attributed to adjuvant systemic chemotherapy.\textsuperscript{2} However, substantial variability in the number of nodes harvested at the hospital level frequently occurs in CRC. For these reasons, it is not unreasonable to expect that the number of nodes harvested in CRC might be an appropriate measure of the surgical and pathologic quality of care that is delivered.

<table>
<thead>
<tr>
<th></th>
<th>First tercile, &lt;7 nodes</th>
<th>Second tercile, 7–9 nodes</th>
<th>Third tercile, $\geq$10 nodes</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1,840</td>
<td>3,638</td>
<td>3,043</td>
<td></td>
</tr>
<tr>
<td>No. of hospitals</td>
<td>12</td>
<td>12</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>932 (50.7)</td>
<td>1,916 (52.7)</td>
<td>1,598 (52.5)</td>
<td>0.33</td>
</tr>
<tr>
<td>Female</td>
<td>908 (49.3)</td>
<td>1,722 (47.3)</td>
<td>1,445 (47.5)</td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>72.7</td>
<td>69.7</td>
<td>69.3</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>20–99</td>
<td>18–103</td>
<td>23–100</td>
<td></td>
</tr>
<tr>
<td>No. of nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>5.6 (5.9)</td>
<td>9.7 (8.5)</td>
<td>11.3 (9.2)</td>
<td></td>
</tr>
<tr>
<td>T stage, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>370 (20.1)</td>
<td>619 (17.0)</td>
<td>504 (16.6)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>T2</td>
<td>297 (16.1)</td>
<td>558 (15.3)</td>
<td>545 (17.9)</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>885 (48.0)</td>
<td>1855 (51.0)</td>
<td>1574 (51.7)</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>291 (15.8)</td>
<td>606 (16.7)</td>
<td>420 (13.8)</td>
<td></td>
</tr>
<tr>
<td>12-Node threshold, %</td>
<td>13.7</td>
<td>32.8</td>
<td>42.8</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Frequency of positive nodes, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>20.8</td>
<td>19.8</td>
<td>20.1</td>
<td>0.12</td>
</tr>
<tr>
<td>N2</td>
<td>9.1</td>
<td>11.1</td>
<td>11.3</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.** Actuarial disease-specific survival of patients according to hospital lymph node counts.

**Figure 2.** Actuarial disease-specific survival N0 patients according to hospital lymph node counts.
In this report, we examined data from R5 CCR to determine the impact of the number of nodes examined on staging and outcomes at the hospital level. Our results demonstrate that increasing node examinations are not associated with increasing levels of identification of node positivity, but are associated with improved survival at the hospital level. This improved survival can be attributed entirely to the node-negative population.

In an examination of the Surveillance, Epidemiology, and End Results/Medicare-linked database of patients undergoing colectomy for nonmetastatic colon cancer, similar findings were reported by Wong and colleagues.5 Similar to our results, these authors noted outcomes differences only in stage II patients treated in hospitals with the highest lymph node examination rates and compared with hospitals with lower lymph node examination rates.

The finding that the survival outcomes benefit at the hospital level appears to be confined to the node-negative population alone is not unanticipated. Much of the patient data that has demonstrated improved outcomes with increasing numbers of lymph nodes examined has been confined to an examination of the node-negative population of nonmetastatic colon and rectal cancer. Substantially less data are available on the impact of the number of nodes on staging accuracy in the node-positive population, although recent reports suggest that the ratio of positive to total nodes harvested has prognostic relevance.6,11

It is not unreasonable to anticipate that those individuals in whom only a small number of lymph nodes are examined might exhibit worse outcomes because of metastases that are unidentified. However, this was not the case in our analysis. Once again, these results are in keeping with those of Wong and colleagues.4 These authors speculated that the extent of the surgical procedure, the diligence in the management of the surgical specimen in pathology, and the extent that lymph nodes undergo examination for metastatic disease might all have a part in explaining these results. Sentinel node staging experience suggests that nodes, in proximity to the primary tumor, are the most likely site for nodal metastases.12 In this setting, even a relatively limited surgical resection can result in the identification of positive nodes and any more extensive lymphadenectomy is not likely to identify substantially greater numbers of positive nodes. That surgeons can modify the conduct of an operation is well-documented. In a prospective randomized trial in gastric cancer, surgeons commonly modified the extent of lymphadenectomy to encompass nodes that were clinically suspicious.13

Differences in the efficiency of identifying metastatic disease have also been suggested to explain these findings. Particularly skilled individuals might be more adept at palpat ing nodes that are most likely to harbor metastatic disease and can focus on harvesting only those nodes presumed to be at greatest risk for harboring metastatic disease. This seems less likely, however, because lymph node size is not a reliable predictor of metastatic disease in colon cancer, and lymph node metastases frequently occur in lymph nodes <5 mm in size.14,15 The extent of metastatic lymph node involvement in R5 CCR was not available for analysis.

Left-sided and rectal tumors are less likely to have an adequate lymph node evaluation than right-sided tumors.16 This, however, does not explain these findings, as there was no significant difference in the distribution of primary tumors between the hospital terciles (p = 0.72). Because, preoperative chemoradiation affects the number of nodes identified in rectal specimens,17 the inclusion of the subset of patients might confound these results. However, only 63% of rectal cancer patients received radiation and >90% of these patients received postoperative radiation. The importance of the number of nodes examined in rectal cancer has been demonstrated to have prognostic impacts similar to those in colon cancer.18

These findings are in stark contrast to the repeated observations that increasing the numbers of nodes examined at the individual patient level increases the percentage of node-positive specimens.2,19-21 In our study population, 42.8% of patients met the 12-node threshold in hospitals within the highest tercile, and hospitals in the lowest tercile had only 13.7% of patients meeting the 12-node threshold. It is not surprising that when <50% of patients in a hospital system meet the 12-node threshold, any impact of understaging on outcomes at the hospital-level threshold is likely to be obscured. As a measure of quality at the hospital level, a considerable majority of patients cared for in the facility might need to meet the 12-node threshold as a benchmark of quality. The most appropriate hospital benchmark remains to be determined, but likely exceeds 60% of patients cared for within the hospital system.

Achieving these benchmarks is not likely to be easy. When compliance with the 12-node threshold is defined as ≥75% of specimens having ≥12 nodes examined, only 15% of hospitals between 1996 and 1997 that participated in the National Cancer Data Base were in compliance with the 12-node threshold.22 The rate of compliance increased to 38% between 2004 and 2005,22 suggesting that despite increasing acceptance of the 12-node threshold, substantial numbers of patients are likely to be understaged.23 Whether 75% of specimens meeting the 12-node threshold is an appropriate benchmark that translates into improved staging and outcomes at the hospital level remains to be determined.
Numerous factors can contribute to outcomes in CRC, including socioeconomic status and race and ethnicity of the patient. In addition, the inherent limitations of a registry database that cannot provide the details of the clinical care of these patients, such as the type and extent of resection performed and use of neoadjuvant chemotherapy and radiotherapy, particularly in patients with rectal cancer, and is not easily discernable from large registry databases might confound the observed outcomes results.

CONCLUSIONS

It is clear that lymph nodes must be examined to accurately stage colon and rectal cancer. Although numerous possible reasons can explain lymph node count variability, having zero nodes examined, as has occurred in approximately 6.5% of patients diagnosed with localized invasive CRC in the Surveillance, Epidemiology, and End Results Cancer registry16 and nearly 16% in R5 CCR with localized invasive CRC, is clearly unacceptable. Continued efforts to maximize node harvesting and examination are warranted. We believe that median lymph node count likely represent an appropriate measure of quality of care at the hospital level in node-negative colon and rectal cancer. Other measures of quality at the hospital level for node-positive colon and rectal cancer need to be identified.

Author Contributions

Study conception and design: Wong, Lum, Morgan
Acquisition of data: Wong, Morgan
Analysis and interpretation of data: Wong, Lum, Morgan
Drafting of manuscript: Wong
Critical revision: Wong, Lum, Morgan

REFERENCES