The American Journal of Surgery xxx (xxxx) xxx

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Original Research Article

Predictors of nodal positivity in clinically under-staged patients with colon cancer: A National Cancer Database study and proposal of a predictive scoring system

Justin Dourado [a], Peter Rogers [a], Sameh Emile [a,b], Anjelli Wignakumar [a], Brett Weiss [a], Nir Horesh [a,c], Zoe Garoufalia [a], Pauline Aeschbacher [d], Steven Wexner [a,*]

[a] Ellen Leifer Shulman and Steven Shulman Digestive Disease Center, Cleveland Clinic Florida, Weston, FL, USA
[b] Colorectal Surgery Unit, General Surgery Department, Mansoura University Hospitals, Mansoura, Egypt
[c] Department of Surgery and Transplantation, Sheba Medical Center, Ramat-Gan, Israel
[d] Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland

ARTICLE INFO

Keywords:
Frequency
Predictors: clinical under-staging
Nodal status
Colon cancer
National cancer database

ABSTRACT

Background: Colon cancer pathological and clinical staging may be disoncordant. This study assessed patients with colon cancer in whom the nodal status was clinically understaged.
Methods: Patients with stage I-III clinical node-negative colon cancer from the National Cancer Database were included. Regression analyses were conducted to elucidate risk factors for clinical nodal understaging and a scoring system was developed to identify high-risk patients.
Results: The study included 94,945 patients with 78.4 % of patients correctly staged and 21.6 % clinically understaged. The predictors of nodal positivity in clinically understaged patients were age <65 (OR 1.43), left-sided tumors (OR 1.41), elevated CEA (OR 2.03), moderately (OR 1.81) or poorly/undifferentiated tumors (OR 3.76), T1 tumors (OR 1.29), signet-ring cell histology (OR 2.26), and microsatellite-stable tumors (OR 1.4).
Conclusion: Patients with colon cancer and the above factors are more likely to have their nodal status clinically understaged. A scoring system has been developed to identify high-risk patients.

1. Introduction

Colorectal cancer (CRC) is the third most common cancer in the world[1] with 126,240 new cases reported in the United States in 2020, accounting for 51,869 deaths.[2] When counselling patients on their treatment and prognosis, accurate staging is of paramount importance. Disease staging is accomplished using the TNM staging system which encompasses depth of invasion (T stage), lymph node (LN) involvement (N stage), and metastatic evaluation (M stage).[3] A variety of modalities are utilized to achieve this staging, including computer tomography (CT) scans, magnetic resonance imaging (MRI), positron emission tomography (PET) scan, endoscopy, and ultrasound.[4]

While the indications for neoadjuvant therapy in colon cancer are currently limited to T4b or "bulky nodal" disease,[5] these guidelines will likely expand in the future. A recent study by the FOxTROT collaboration demonstrated that neoadjuvant chemotherapy in operable colon cancer resulted in fewer incomplete resections, and better 2-year disease control.[6] Additionally, knowledge of an accurate pre-operative stage will have implications for patient counselling and prognostication. Localized colon cancer has a 5-year survival rate of 91 %, whereas spread of disease to the regional LNs or nearby structures further reduces 5-year survival to 72 %.[7] The assessment of nodal status is particularly challenging. A recent study has shown that the agreement between clinical and pathologic staging of nodal status in rectal cancer is suboptimal.[8] Indeed, the sensitivity of CT scan in assessment of nodal disease can be as low as 71 %.[9]

Higher rates of minimally invasive approaches, lower complications, lower reoperation, and lower costs are associated with treatment at high-volume centers.[10] This finding was supported in elderly stage I-III colon cancer patients across 465 hospitals where selection of hospital improved

* Corresponding author. Ellen Leifer Shulman and Steven Shulman Digestive Disease Center, Cleveland Clinic Florida, 2950 Cleveland Clinic Blvd, Weston, FL, 33331, USA.
E-mail address: wexners@ccf.org (S. Wexner).
@SWexner (S. Wexner)

https://doi.org/10.1016/j.amjsurg.2024.115777
Received 25 March 2024; Received in revised form 16 May 2024; Accepted 23 May 2024

length of stay and in-hospital mortality.[11] Notably, in rectal cancer, referral to a specialized center has been linked to reduced local recurrence and prolonged overall survival.[12] This finding was supported in a large cohort that looked at centralization of care for multiple cancer types including colon cancer.[13]

Our study aimed to assess the frequency and predictors of clinical understaging of nodal disease in patients with colon cancer. We hypothesized that certain patient-related, tumor-related, and treatment-related factors may affect the accuracy of clinical nodal staging in colon cancer. Learning about the characteristics that predispose to nodal understaging may help identify patients who could benefit from review of their scans or additional assessment before treatment.

2. Methods

2.1. Study design

We undertook a retrospective case-control study of patients with clinical node-negative colon cancer from the NCDB from 2010 to 2019. Patients were subdivided into cases (those upstaged from cN0 to pN+) and controls who were correctly staged (pN0) according to the pathologic N stage after resection.

2.2. Data source

The NCDB includes data from over 1500 Commission on Cancer (CoC)-accredited hospitals across the United States. The NCDB is a joint project of the CoC of the American College of Surgeons and the American Cancer Society. "The de-identified data used in the study are derived from the NCDB and its participating hospitals that are not responsible for the statistical validity of the analysis or the conclusions of the study". Given that de-identified data were derived from a public database, local ethical approval was not required.

2.3. Study population

Patients included in this study were adults of either sex with stage I-III colonic adenocarcinoma who underwent colectomy with curative intent. Only patients with clinical N0 stage were included. Patients with appendiceal cancers, rectal cancer, histologic types other than adenocarcinomas, unknown clinical stage or stage IV disease, patients who underwent local excision or non-specified type of surgery, patients with cN1-2 stage, and patients with unknown cN or pN stage.

2.4. Data collected

Data points included in this study were age, sex, race, clinical and pathologic TNM stage, histological subtype, clinical assessment of tumor size, tumor location, facility subtype, insurance status, Charlson score, nodal harvest, carcinoembryonic antigen (CEA) levels, surgical approach, facility type, microsatellite instability (MSI) status, residence area, and surgical approach.

2.5. Study outcomes

The main outcome of this study was the predictive factors associated with clinical understaging of nodal affection in colon cancers [cN0 that had pN1-2 (N+) status on pathologic examination].

3. Statistical analysis

Statistics were analysed using EZR (version 1.55) and R (version 4.1.2).[14] Categorical data was expressed as absolute numbers and percentages, with analysis performed using the Fisher-exact test or Chi-square test. Based on normality of distribution of the continuous data, they were expressed as a mean and standard deviation or median and inter-quartile range (IQR) with analyses performed using the student t-test or Mann-Whitney test as appropriate. To identify factors associated with nodal positivity in clinical understaged patients, a univariable analysis was performed. To identify the independent factors associated with nodal understaging, a multivariable binary logistic regression analysis was performed on preoperative factors that reached statistical significance ($p < 0.05$) in the univariable analysis, with similarity among health systems and practice environments (by the exclusion of insurance type or hospital classification). A sensitivity multivariable analysis was conducted including only patients who had ≥12 lymph nodes harvested at surgery. A Receiver Operator Curve (ROC) was generated with the area under the curve (AUC) calculated to determine the discriminatory ability of the model used. Variance Inflation Factor (VIF) was calculated for each factor in our multivariable analysis to assess multicollinearity. A VIF of 5–10 indicated a moderate degree of collinearity whereas VIF >10 indicated extensive collinearity.

3.1. Development of predictive score

A predictive risk score for clinical nodal understaging was developed using the odds ratios (OR) of the significant predictors of nodal understaging identified by the multivariable regression analysis. After insignificant factors were excluded, the weighted odds ratio for each predictor was determined by dividing it by the smallest significant OR. The weighted ORs were then rounded to integer points and each variable was allocated score points, as was previously described.[15] Cutoff points for the score were calculated using ROC curve analyses and patients were divided into three risk groups. The incidence of nodal understaging in each risk group was calculated and compared. The diagnostic value of the score was expressed as sensitivity, specificity, and accuracy.

4. Results

4.1. Description of the cohort

The present study included 94,945 patients with colon cancer who were initially staged with clinical node-negative disease (Fig. 1). The cohort had a mean age of 68.6 ± 13.1 years; 48.9 % of the cohort were male, 84.3 % were White, 11.5 % were Black and 3 % were of Asian descent. The mean follow-up was 57.9 months. 32.4 % of patients had a Charlson-Deyo score of ≥1. Medicare insurance accounted for 59.3 % of patients, followed by private insurance (32.5 %). 43.4 % were treated in a comprehensive community cancer program, and 27.2 % at an academic research program. All patients were surgically treated and the mean number of examined regional nodes of 18 (IQR: 14–24). The number of LNs harvested in the cohort ranged from 0 (n = 416) to 90 (n = 130). All patients with 0 nodes recorded underwent colectomy and this is likely a transcription error. Most patients had clinical stage 1 disease (59.0 %), followed by stage 2 (39.7 %), and stage 3 (1.3 %). Overall 5-year survival was 66.6 %. A summary of the characteristics of the study cohort is shown in Table 1.

4.2. Comparison of correctly staged and understaged patients

Within the cohort there were 74,420 (78.4 %) patients who were correctly staged in clinical assessment and had node-negative disease after pathologic evaluation. Conversely 20,525, 21.6 % (95%CI: 21.4–21.9) patients were understaged and had a node-positive disease on pathology. Factors associated with clinical understaging on univariate analysis were race ($p < 0.001$), age ($p < 0.001$), Charlson-Deyo score ($p < 0.001$), hospital setting ($p < 0.001$), and Medicare status ($p < 0.001$). Area of residence ($p = 0.291$) and sex ($p = 0.981$) were not associated with clinical node understaging (Table 2). Tumor characteristics associated with clinical nodal understaging were CEA level ($p < 0.001$), grade ($p < 0.001$), histology ($p < 0.001$), MSI status ($p < 0.001$), tumor sidedness ($p < 0.001$), and clinical tumor size ($p < 0.001$) (Table 3).

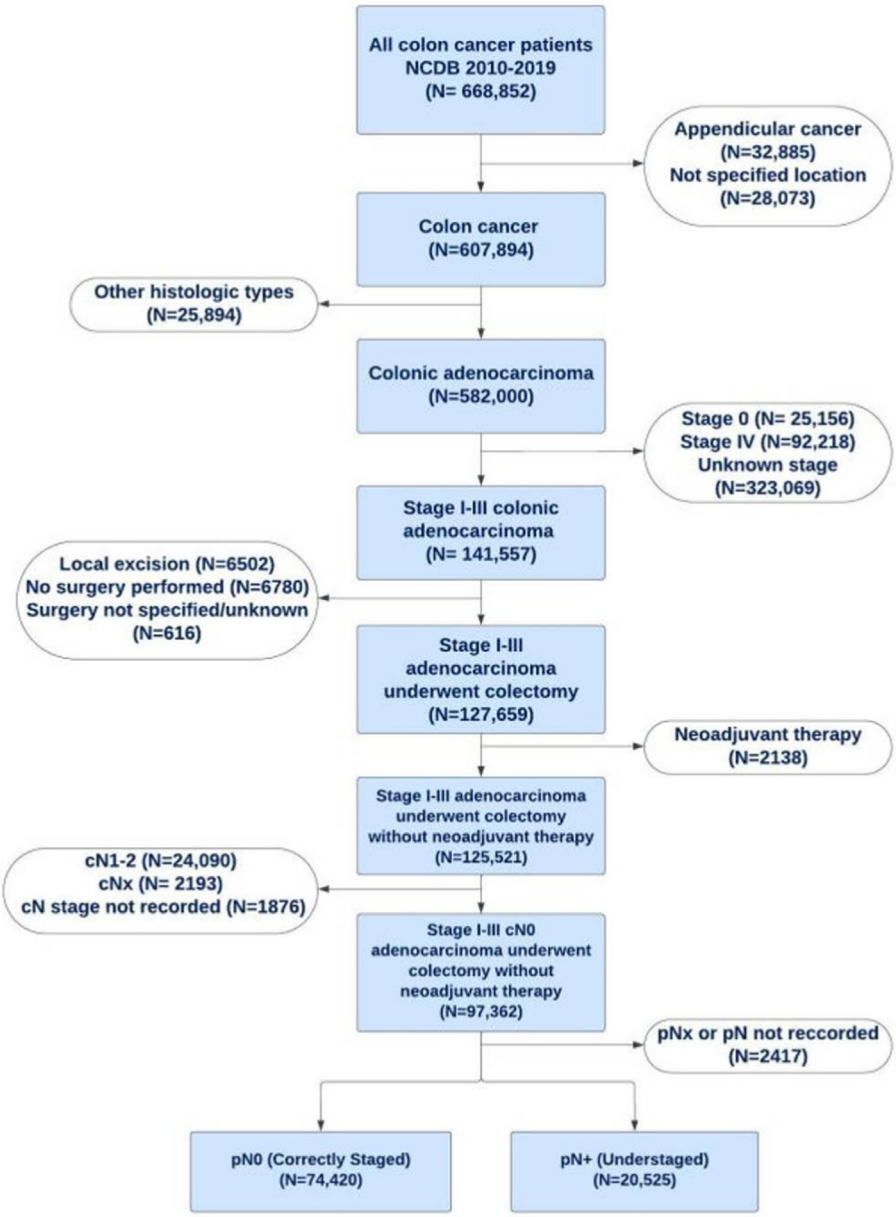

Fig. 1. Identification of the cohort.

4.3. Multivariable analyses

Multivariable regression analysis showed that the significant independent predictors of nodal positivity in clinically understaged patients were left-sided tumors (OR 1.41, 95 % CI: 1.27–1.57, $p < 0.001$), elevated CEA levels (OR 2.03, 95 % CI: 1.05–3.94, $p = 0.036$), moderately differentiated tumors (OR 1.81, 95 % CI: 1.51–2.17, $p < 0.001$), poorly/undifferentiated tumors (OR 3.76, 95 % CI: 3.08–4.60, $p < 0.001$), T1 tumors (OR 1.29, 95 % CI: 1.08–1.55, $p = 0.006$), microsatellite stability (OR 1.4, 95 % CI: 1.27–1.57, $p < 0.001$), and signet-ring cell carcinoma (OR 2.26, 95 % CI: 1.46–3.50, $p < 0.001$). Younger age is additionally associated with a higher OR of clinical understaging (OR 1.430, 95 % CI: 1.29–1.57, $p < 0.001$) (Table 4). The cutoff age of 65 years was identified on ROC analysis (sensitivity = 44 %, specificity = 62.8 %, AUC = 0.546). The AUC of the model was 0.65 (95 % CI: 0.637–0.662), which indicates acceptable discriminatory ability (Supplementary Fig. 1). The above analysis was then undertaken including only those patients who had at least 12 LNs harvested during surgery. Only CEA level became insignificant with all other factors remaining significant and no additional significant factors were identified (Supplementary Table 1). The AUC of the secondary model remained similar indicating acceptable discriminatory ability and similarity to the prior analysis. Additionally, the VIF of all predictors within both analyses was <3 implying the absence of collinearity (Supplementary Table 2).

4.4. Risk scoring system

A scoring system was developed to predict which patients were at risk for nodal understaging (Table 5). A total of 12,799 patients, who had data for all factors were included in the scoring system. A ROC analysis was performed to identify suitable cutoffs for the three risk groups: low-risk (0–3), intermediate risk group (4–5), and high-risk group (≥6) (Supplementary Fig. 2). There were then 1792 patients in the high-risk group, 5260 in the intermediate risk, and 5746 in the low-risk group. The rates of nodal understaging were 42.8 % (IQR 40.5–45.1 %) with a median score of 6 (IQR 6–7) in the high-risk group, 27.5 % (IQR

Table 1
Description of cohort.

Factor	Group	Overall
n		94,945
Age		68.58 (13.07)
Sex	**Female**	48,513 (51.1)
	Male	46,432 (48.9)
Race	**White**	79,414 (84.3)
	Black	10,833 (11.5)
	Asian	2833 (3.0)
	Other	882 (0.9)
	American Indian	293 (0.3)
Follow up in Months		57.95 [31.05, 85.52]
Charlson Deyo Score	**0**	64,182 (67.6)
	1	20,622 (21.7)
	2	6471 (6.8)
	3	3670 (3.9)
Insurance	**Medicaid**	4466 (4.8)
	Medicare	55,536 (59.3)
	Not insured	2411 (2.6)
	Other government	795 (0.8)
	Private insurance	30,499 (32.5)
Facility Type	**Academic/Research Program**	25,320 (27.2)
	Community Cancer Program	8997 (9.7)
	Comprehensive Community Cancer Program	40,358 (43.4)
	Integrated Network Cancer Program	18,371 (19.7)
Histology Summary	**Adenocarcinoma**	85,923 (90.5)
	Mucinous adenocarcinoma	8209 (8.6)
	Signet-ring cell carcinoma	813 (0.9)
Clinical T Stage	**0**	118 (0.1)
	1	39,161 (41.6)
	2	16,252 (17.2)
	3	30,575 (32.5)
	4	6713 (7.1)
	x	1399 (1.5)
Clinical TNM Stage	**1**	56,041 (59.0)
	2	37,658 (39.7)
	3	1246 (1.3)
Examined Regional Nodes		18.00 [14.00, 24.00]
Overall Survival	**Alive**	59,908 (66.6)
	Dead	30,053 (33.4)

Data presented as n (%) or median [IQR].

Table 2
Univariate analyses of patient factors.

Factor	Group	Negative (correctly staged)	Positive (understaged)	p-value
n		74,420	20,525	
Race	**American Indian**	226 (0.3)	67 (0.3)	<**0.001**
	Asian	2085 (2.8)	748 (3.7)	
	Black	8334 (11.3)	2499 (12.3)	
	Other	702 (1.0)	180 (0.9)	
	White	62,533 (84.6)	16,881 (82.9)	
Age		69.06 (12.84)	66.85 (13.70)	<**0.001**
Charlson Deyo Score	**0**	49,947 (67.1)	14,235 (69.4)	<**0.001**
	1	16,378 (22.0)	4244 (20.7)	
	2	5192 (7.0)	1279 (6.2)	
	3	2903 (3.9)	767 (3.7)	
Facility Type	**Academic/ Research Program**	19,531 (26.7)	5789 (29.1)	<**0.001**
	Community Cancer Program	7261 (9.9)	1736 (8.7)	
	Comprehensive Community Cancer Program	31,860 (43.6)	8498 (42.7)	
	Integrated Network Cancer Program	14,472 (19.8)	3899 (19.6)	
Insurance	**Medicaid**	3340 (4.6)	1126 (5.5)	<**0.001**
	Medicare	44,406 (60.5)	11,130 (54.8)	
	Other government	613 (0.8)	182 (0.9)	
	Private insurance	23,195 (31.6)	7304 (35.9)	
	Not insured	1832 (2.5)	579 (2.8)	
Residence Area	**Metro**	62,382 (85.8)	17,100 (85.4)	0.291
	Rural	1146 (1.6)	322 (1.6)	
	Urban	9186 (12.6)	2612 (13.0)	
Sex	**Female**	38,024 (51.1)	10,489 (51.1)	0.981
	Male	36,396 (48.9)	10,036 (48.9)	

Data presented as n (%) or mean (SD). Bold text in p value column indicates statistical significance.

26.3–28.7 %) with a median score of 4 (IQR 4–5) in the intermediate risk group, and 16.6 % (IQR 15.6–17.6 %) with a median score of 3 (IQR 2–3) in the low-risk group ($p < 0.001$). The IQRs did not have any overlap among the 3 risk groups. Intermediate risk patients had a 90 % higher rate of nodal understaging (OR 1.9, 95 % CI: 1.7–2.1) and high-risk patients had approximately a four-fold higher rate of nodal understaging (OR 3.8, 95 % CI: 3.3–4.2) as compared to low-risk patients (Table 6). The specificity of the scoring system when comparing high to low-risk patients was 82.4 %, sensitivity 44.6 %, negative predictive value 83.4 %, and positive predictive value 42.8 % with an accuracy of 73.8 %.

5. Discussion

The present study found that, within the NCDB, 21.6 % of patients with clinically node negative colon cancer were subsequently found to be node positive on pathology. Factors found to be associated with clinical nodal understaging were age younger than 65 years, T1 tumors, left-sided tumors, elevated CEA levels, moderately or poorly/undifferentiated tumors, MSI negative status, and signet-ring cell carcinoma. This finding of understaging was maintained with the exception of CEA when evaluating only patients who had an adequate nodal harvest at surgery. Additionally, a scoring system was created and internally tested to attempt to predict which patients are at highest risk of being clinically understaged in terms of nodal status.

The clinical assessment of colon cancer aims to identify the extent of malignant progression. Based on pre-operative assessments, we formulate a plan of treatment best suited to the individual patient. The decisions on treatment, frequently discussed at a multidisciplinary team (MDT) meeting, should be based on the most accurate high-quality information. Inaccurate information at this planning stage will inevitably impede optimal patient care. Within this study, we investigated a cohort of patients whose lymph node status was clinically understaged when compared to the final stage revealed by pathological examination of the specimen.

Unlike rectal cancer, clear indications for neoadjuvant therapy in colon cancer are not as robust. The NCCN guidelines,[5] based on the FoxTroT study,[16] describe “bulky nodal disease’ and “T4b disease” as indicators for neoadjuvant chemotherapy, but these decisions are nuanced and usually decided during the MDT discussion. Despite nodal disease not being an absolute indication for neoadjuvant treatment according to the NCCN, it may have implications when planning preoperative therapy. Additionally, in a recent prospective multicenter study, complete mesocolic excision (CME) for right sided colon cancer was shown to have an overall survival benefit (HR 0.52, $p = 0.01$) in stage 3 disease.[17] As nodal positivity upgrades a patient from stage 2 to stage 3, and in light of the above two trials, proper staging of these patients has increasing clinical significance. We propose that the identified risk factors, and the scoring system that we developed could help identify high-risk patients for whom either neoadjuvant therapy or CME may be considered. Contingent upon external validation of the new scoring system, prospective studies assessing the benefits of neoadjuvant

Table 3
Univariate analyses of tumor factors.

Factor	Group	Negative (correctly staged)	Positive (understaged)	p-value
n		74,420	20,525	
CEA levels	**Normal**	26,849 (68.4)	6815 (57.1)	<**0.001**
	Elevated	12,150 (31.0)	5033 (42.2)	
	Borderline	235 (0.6)	81 (0.7)	
Grade	**Well differentiated**	10,090 (14.4)	1497 (7.5)	<**0.001**
	Moderately differentiated	50,778 (72.5)	13,360 (67.0)	
	Poorly/ Undifferentiated	9200 (13.1)	5082 (25.5)	
Histology	**Adenocarcinoma**	67,737 (91.0)	18,186 (88.6)	<**0.001**
	Mucinous adenocarcinoma	6290 (8.5)	1919 (9.3)	
	Signet-ring cell carcinoma	393 (0.5)	420 (2.0)	
T-stage	**0**	97 (0.1)	21 (0.1)	<**0.001**
	1	31,039 (42.0)	8122 (39.8)	
	2	13,337 (18.1)	2915 (14.3)	
	3	23,571 (31.9)	7004 (34.4)	
	4	4717 (6.4)	1996 (9.8)	
	x	1070 (1.4)	329 (1.6)	
MSI Status	**Negative**	12,421 (70.9)	3966 (76.4)	<**0.001**
	Positive	5092 (29.1)	1226 (23.6)	
Tumor Location	**Right**	33,395 (50.4)	8438 (45.7)	<**0.001**
	Transverse colon	8200 (12.4)	2229 (12.1)	
	Left	23,818 (35.9)	7554 (40.9)	
	Overlapping lesion	897 (1.4)	232 (1.3)	
Median tumor size in mm		37.00 [20.00, 56.00]	45.00 [30.00, 60.00]	<**0.001**

Data presented as n (%) or mean (SD). Bold text in p value column indicates statistical significance.

Table 4
Multivariable regression analysis of nodal positivity in clinically understaged patients in the entire cohort.

Variable	Group	Odds Ratio	Lower 95%CI	Upper 95%CI	p value
Age < 65		1.430	1.290	1.570	<**0.001**
Race	**White**	**Ref**			
	Black	1.100	0.945	1.270	0.229
	Asian	1.280	0.995	1.640	0.055
	Other	1.340	0.894	2.010	0.156
	American Indian	1.380	0.680	2.790	0.373
Charlson Deyo Score	**0**	**Ref**			
	1	0.950	0.845	1.070	0.394
	2	0.864	0.708	1.060	0.154
	3	1.150	0.890	1.490	0.285
Clinical T stage	**4**	**Ref**			
	3	0.930	0.779	1.110	0.426
	2	0.849	0.695	1.040	0.109
	1	1.290	1.080	1.550	**0.006**
Tumor Location	**Right**	**Ref**			
	Transverse colon	1.050	0.909	1.220	0.483
	Left	1.410	1.270	1.570	<**0.001**
	Overlapping lesion	1.000	0.662	1.510	0.997
Tumor Size in mm		1.000	0.999	1.000	0.328
CEA levels	**Borderline**	**Ref**			
	Normal	1.280	0.659	2.470	0.469
	Positive/elevated	2.030	1.050	3.940	**0.036**
Grade	**Well**	**Ref**			
	Moderate	1.810	1.510	2.170	<**0.001**
	Poorly/ undifferentiated	3.760	3.080	4.600	<**0.001**
Histology	**Adenocarcinoma**	**Ref**			
	Mucinous adenocarcinoma	1.010	0.858	1.180	0.939
	Signet-ring cell carcinoma	2.260	1.460	3.500	<**0.001**
MSI Status [Negative]		1.400	1.260	1.570	<**0.001**

CEA, carcinembryonic antigen; MSI, microsatellite instability; CI, confidence interval. Bold text in p value column indicates statistical significance.

Table 5
Scoring system for prediction of nodal positivity in clinically understaged colon cancer patients.

Group	Odds Ratio	Weighted OR	Score points
Age <65	1.43	1.11	1
Elevated CEA	2.03	1.57	2
T1 stage	1.29	1	1
Negative MSI status	1.4	1.1	1
Grade 2	1.81	1.4	1
Grade 3	3.76	2.9	3
Signet ring cell carcinoma	2.26	1.75	2
Left colon cancer	1.41	1.1	1

MSI, microsatellite instability; CEA, carcinoembryonic antigen; OR, odds ratio.
*Total score ranges from 0 to 11.
**Risk groups: Low risk (0–3), Intermediate risk (4–5), High risk (6 or higher).

Table 6
Analysis of scoring system within the entire cohort.

		Risk Group			
Factor	Group	High	Intermediate	Low	P value
N (%)		1792	5260	5746	<**0.001**
Understaged?	No	1025 (57.2)	3813 (72.5)	4792 (83.4)	
	Yes	767 (42.8, 40.5–45.1)	1447 (27.5, 26.3–28.7)	954 (16.6, 15.6–17.6)	<**0.001**
Total Score		6.00 [6.00, 7.00]	4.00 [4.00, 5.00]	3.00 [2.00, 3.00]	

*Values are represented as (%, IQR) or median [IQR]. Bold text in the p value column indicates statistical significance.

chemotherapy and/or CME in high-risk patients are needed.

The most pertinent finding of this study is that 21.6 % of colon cancer patients who were clinically node negative were node positive in the surgical specimen. This finding implies that clinical assessment was inaccurate for almost one-quarter of the patients when their management plan was formulated. This inaccuracy may have altered surgical and oncological decision-making that could potentially adversely impact oncological outcomes. What this result demonstrates is that we must endeavor to improve our pre-operative staging ability. CT scanning is the benchmark staging investigation used in colon cancer, but accuracy is limited. A study by Rafaelson et al.,[18] showed the sensitivity and specificity of the CT scan in determining nodal status in colon cancer was 65 % and 50 %, respectively. Interestingly, MRI fared worse, with a sensitivity and specificity of 58 % and 50 %, respectively.

Factors independently associated with clinical understaging of nodal status are patient-related, tumor-related, and treatment-related. The patient-related factor identified was younger age. Younger age was significantly associated with clinical node under-staging. It is well documented that younger patients diagnosed with colon cancer tend to have higher rates of nodal positivity.[19–21] Factors that have been implicated in this association are minimal screening for younger patients who are therefore more likely to present with advanced stage disease,[21] more aggressive operations given age therefore larger nodal sampling,[22] and differing tumor biology.[23]

Left-sided disease was another factor identified to increase risk of nodal under-staging by 83 % compared to right-sided disease. Accurate nodal staging can be more difficult when the primary lesion is in the left colon, compared to the right side. A prospective multicentre study by Dehal et al.,[24] in 2019 described that the sidedness of cancer effects the size of surgical nodal harvest required to accurately stage clinical T3N0 disease. The exact mechanism of this is unclear, but certain biological and epidemiological differences have been described between left and right sided colon cancer.[25,26] This finding could potentially lead to variability in clinical and pathological nodal detection.

Tumor related factors, such as elevated CEA levels, T1 tumors, moderately or poorly differentiated carcinomas, and signet-ring cell

histology, were associated with an increased risk of clinical nodal understaging. CEA levels, a preoperative tumor marker, were associated with 61 % higher odds of having their nodal status under-staged. CEA is a glycoprotein that is produced in the embryonic period, its secretion is reduced after birth and is hardly measurable in normal adulthood. It is commonly used in the detection of colorectal cancer recurrence and progression.[27,28] The pathophysiology of CEA elevation being implicated in nodal under-staging is not clear, but a study by Zhang et al.[29] circa 2023 suggested an 18-node harvest should be undertaken in CEA elevated colon cancer instead of the typical 12-node harvest. This observation implies that more nodes should be sampled to achieve an accurate pathological nodal staging compared to those cancers with normal CEA levels. Furthermore, T-stage has been associated with both the number of nodes that need to be examined to have an adequate sensitivity for detecting nodal metastases[30] and with more nodal metastases.[31] In terms of increasing grade, there is evidence that increasing grade leads to a higher rate of nodal positivity in colon cancer.[31] Finally, signet-ring cell histology has previously been related to increasing nodal metastases and poorer prognosis.[32–35] Our study supports the fact that the above factors are not only associated with nodal metastasis, but the risk of clinically understaging these patients' nodal status.

Microsatellite instability was found to significantly decrease the odds of nodal clinical under-staging by 27 %. It is well known that tumour biology, pathological, and clinical characteristics are altered when colon cancer displays MSI.[36–38] A study by Kang et al.[39] in 2018 demonstrated that MSI-high patients had more lymph nodes found, earlier stage tumors, more advanced T stage, and poorer differentiation in colorectal cancer than did patients with MSI stable tumors. Kim et al.[26] in 2022, demonstrated that patients who are MSI-high and have lymph node metastasis demonstrate longer measured nodes than MSI stable tumors, therefore a higher specificity for predicting N0 on CT assessment. This could explain the lower rate of clinical understaging. Secondly, the authors speculate that those who are MSI high are referred to specialist centers. Here, they have access to higher fidelity imaging and expert radiologist interpretation, reducing the rate of clinical under-staging.

The factors identified within this study that increase the odds of clinical nodal under-staging need to be considered in the management of cN0 colon cancer. Patients <65 years of age, left-sided tumors, elevated CEA, moderately and poorly/undifferentiated tumors, and MSI stable status were all factors significantly associated with nodal under-staging in our NCDB study. A scoring system has been developed with good specificity for nodal understaging which can lead to early identification of these patients with reasonably high accuracy. It is logical to suggest that these patients should be referred for more detailed evaluation, such as higher quality imaging modalities, referral to a specialized centre, multidisciplinary tumor board discussion, or imaging re-review by an expert radiologist. Through correctly identifying patients at high risk for occult nodal disease, better patient counselling on oncologic outcomes can be undertaken in addition to consideration of neoadjuvant therapy or complete mesocolic excision.

There are some significant limitations to this study. Firstly, it is retrospective in nature. There is potential misclassification and data entry errors that are common with databases. Moreover, there is a lack of staging data in many patients. Additionally, the clinical assessment method is often not known and the criteria of defining cN+ is unknown. Finally, some of the pathologic predictors of under-staging are only known after resection. However, this very large dataset allowed for performing a sensitive analysis and the development of a scoring system to predict patients at high risk of nodal understaging in colon cancer that may have clinical implications.

6. Conclusion

Patients <65 years of age, with left-sided tumors, T1 tumors, elevated CEA levels, moderately or poorly/undifferentiated carcinomas, signet-ring cell histology, and MSI stable status are more likely to have their nodal status understaged during the clinical assessment of colon cancer. A scoring system has been developed to identify which high-risk patients should undergo more detailed clinical assessment to avoid missing an otherwise undetected nodal disease. Expert radiological interpretation or referral to a specialist center may be warranted in the high-risk patients to correctly identify under-staged nodal disease.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Justin Dourado: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Peter Rogers:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sameh Emile:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Anjelli Wignakumar:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Brett Weiss:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Nir Horesh:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Zoe Garoufalia:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Pauline Aeschbacher:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Steven Wexner:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

None of the authors has any relevant financial disclosures. Dr. Wexner is a consultant for Baxter, Becton, Dickinson and Co, Glaxo Smith Kline, Intuitive Surgical, Livsmed, Medtronic, OstomyCure, Stryker, Takeda, Virtual Ports, is a member of the Data Safety Monitoring Board of JSR/WCG/ACI (chair), Polypoid (chair), and Boomerang and receives royalties from Intuitive Surgical, Karl Storz Endoscopy America Inc., and Unique Surgical Solutions, LLC.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amjsurg.2024.115777.

References

1. World Health Organization: WHO & World Health Organization: WHO. Colorectal cancer. https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer; 2023, July 11. Accessed December 12, 2023.
2. Colorectal cancer statistics. Center for Disease Control and Prevention. Available on https://gis.cdc.gov/Cancer/USCS/#/Trends/. Accessed on November 29, 2023.
3. Hari DM, Leung AM, Lee JH, et al. AJCC Cancer Staging Manual 7th edition criteria for colon cancer: do the complex modifications improve prognostic assessment? *J Am Coll Surg*. 2013;217:181–190.
4. Shkurti J, van den Berg K, van Erning FN, Lahaye MJ, Beets-Tan RGH, Nederend J. Diagnostic accuracy of CT for local staging of colon cancer: a nationwide study in The Netherlands. *Eur J Cancer*. 2023;193:113314. https://doi.org/10.1016/j.ejca.2023.113314.
5. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) for Colon Cancer. Version 4.2023*. November 16, 2023. © National Comprehensive Cancer Network, Inc. 202X. All rights reserved. Accessed 11/28/2023.
6. Morton D, Seymour M, Magill L, et al. Preoperative chemotherapy for operable colon cancer: mature results of an international randomized controlled trial. *J Clin Oncol*. 2023;41:1541–1552. https://doi.org/10.1200/JCO.22.00046. Epub 2023 Jan 19. PMID: 36657089; PMCID: PMC10022855.
7. Survival rates for colorectal cancer. American Cancer Society. Available on https://www.cancer.org/cancer/types/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html. Accessed on November 29, 2023.
8. Emile SH, Silva-Alvarenga E, Horesh N, Freund MR, Garoufalia Z, Wexner SD. Concordance between clinical and pathologic assessment of T and N stages of rectal adenocarcinoma patients who underwent surgery without neoadjuvant therapy: a

National Cancer Database analysis. *Eur J Surg Oncol*. 2023;49:426–432. https://doi.org/10.1016/j.ejso.2022.09.014.
9. Nerad E, Lahaye MJ, Maas M, et al. Diagnostic accuracy of CT for local staging of colon cancer: a systematic review and meta-analysis. *AJR Am J Roentgenol*. 2016;207:984–995. https://doi.org/10.2214/AJR.15.15785.
10. Damle RN, Macomber CW, Flahive JM, et al. Surgeon volume and elective resection for colon cancer: an analysis of outcomes and use of laparoscopy. *J Am Coll Surg*. 2014;218:1223–1230. https://doi.org/10.1016/j.jamcollsurg.2014.01.057. Epub 2014 Mar 12. PMID: 24768291; PMCID: PMC4467094.
11. Zheng Z, Hanna N, Onukwugha E, Bikov KA, Mullins CD. Hospital center effect for laparoscopic colectomy among elderly stage I-III colon cancer patients. *Ann Surg*. 2014;259(5):924–929. https://doi.org/10.1097/SLA.0b013e31829d0468. PMID: 23817508.
12. Khani MH, Smedh K. Centralization of rectal cancer surgery improves long-term survival. *Colorectal Dis*. 2010;12:874–879. https://doi.org/10.1111/j.1463-1318.2009.02098.x. PMID: 19878515.
13. Sheetz KH, Dimick JB, Nathan H. Centralization of high-risk cancer surgery within existing hospital systems. *J Clin Oncol*. 2019;37:3234–3242. https://doi.org/10.1200/JCO.18.02035. Epub 2019 Jun 28. PMID: 31251691; PMCID: PMC7351344.
14. Kanda Y. Investigation of the freely available easy-to-use software "EZR" for medical statistics. *Bone Marrow Transplant*. 2013;48:452–458.
15. Chua SL, Chow WL. *Development of predictive scoring model for risk stratification of no-show at a public hospital specialist outpatient clinic*. *Proceedings of Singapore Healthcare*. 2019:96–104. https://doi.org/10.1177/2010105818793155.
16. Foxtrot Collaborative Group. Feasibility of preoperative chemotherapy for locally advanced, operable colon cancer: the pilot phase of a randomised controlled trial. *Lancet Oncol*. 2012;13:1152–1160. https://doi.org/10.1016/S1470-2045(12)70348-0. Epub 2012 Sep 25. PMID: 23017669; PMCID: PMC3488188.
17. Benz SR, Feder IS, Vollmer S, et al. Complete mesocolic excision for right colonic cancer: prospective multicentre study. *Br J Surg*. 2022;110:98–105. https://doi.org/10.1093/bjs/znac379. PMID: 36369986; PMCID: PMC10364501.
18. Rafaelsen SR, Dam C, Vagn-Hansen C, et al. CT and 3 tesla MRI in the tn staging of colon cancer: a prospective, blind study. *Curr Oncol*. 2022;29:1069–1079. https://doi.org/10.3390/curroncol29020091. PMID: 35200590; PMCID: PMC8870524.
19. Alexander MS, Lin J, Shriver CD, McGlynn KA, Zhu K. Age and lymph node positivity in patients with colon and rectal cancer in the US military health system. *Dis Colon Rectum*. 2020;63:346–356. https://doi.org/10.1097/DCR.0000000000001555. PMID: 31842166; PMCID: PMC7021218.
20. Khan H, Olszewski AJ, Somasundar P. Lymph node involvement in colon cancer patients decreases with age; a population based analysis. *Eur J Surg Oncol*. 2014;40:1474–1480. https://doi.org/10.1016/j.ejso.2014.06.002. Epub 2014 Jul 3. PMID: 25027280.
21. Saraste D, Martling A, Nilsson PJ, Blom J, Törnberg S, Janson M. Screening vs. non-screening detected colorectal cancer: differences in pre-therapeutic work up and treatment. *J Med Screen*. 2017;24:69–74.
22. Wang L, Hollenbeak CS, Stewart DB. Node yield and node involvement in young colon cancer patients: is there a difference in cancer survival based on age? *J Gastrointest Surg*. 2010;14:1355–1361.
23. Ahnen DJ, Wade SW, Jones WF, et al. The increasing incidence of young-onset colorectal cancer: a call to action. *Mayo Clin Proc*. 2014;89:216–224. https://doi.org/10.1016/j.mayocp.2013.09.006. Epub 2014 Jan 4. PMID: 24393412.
24. Dehal AN, Nelson D, Chang SC, Dahel A, Bilchik AJ. Accuracy of nodal staging is influenced by sidedness in colon cancer. *J Gastrointest Oncol*. 2019;10:902–909. https://doi.org/10.21037/jgo.2019.08.01.
25. Benedix F, Kube R, Meyer F, et al. Comparison of 17,641 patients with right- and left-sided colon cancer: differences in epidemiology, perioperative course, histology, and survival. *Dis Colon Rectum*. 2010;53:57–64 [Crossref] [PubMed].
26. Kim SR, Song N, Yothers G, et al. Tumour sidedness and intrinsic subtypes in patients with stage II/III colon cancer: analysis of NSABP C-07 (NRG Oncology). *Br J Cancer*. 2018;118:629–633.
27. Kuroki M, Haruno M, Arakawa F, et al. Reaction profiles of seven enzyme immunoassay kits for carcinoembryonic antigen (CEA) analyzed with purified preparations of CEA and related normal antigens. *Clin Biochem*. 1992;25:29–35.
28. Sell SS. *Serological Cancer Markers*. Humana Press; 1992.
29. Zhang H, Wang C, Liu Y, et al. The optimal minimum lymph node count for carcinoembryonic antigen elevated colon cancer: a population-based study in the SEER set and External set. *BMC Cancer*. 2023;23:100. https://doi.org/10.1186/s12885-023-10524-y.
30. Wu Z, Qin G, Zhao N, Jia H, Zheng X. Assessing the adequacy of lymph node yield for different tumor stages of colon cancer by nodal staging scores. *BMC Cancer*. 2017;17:498. https://doi.org/10.1186/s12885-017-3491-2. PMID: 28743236; PMCID: PMC5526283.
31. Guan X, Chen W, Li S, et al. Alterations of lymph nodes evaluation after colon cancer resection: patient and tumor heterogeneity should be taken into consideration. *Oncotarget*. 2016;7:62664–62675. https://doi.org/10.18632/oncotarget.11633. PMID: 27577077; PMCID: PMC5308756.
32. Piredda ML, Ammendola S, Sciammarella C, et al. Colorectal cancer with microsatellite instability: right-sided location and signet ring cell histology are associated with nodal metastases, and extranodal extension influences disease-free survival. *Pathol Res Pract*. 2021;224:153519. https://doi.org/10.1016/j.prp.2021.153519. Epub 2021 Jun 8. PMID: 34119815.
33. Nitsche U, Zimmermann A, Späth C, et al. Mucinous and signet-ring cell colorectal cancers differ from classical adenocarcinomas in tumor biology and prognosis. discussion 782-3 *Ann Surg*. 2013;258:775–782. https://doi.org/10.1097/SLA.0b013e3182a69f7e. PMID: 23989057; PMCID: PMC3888475.
34. Gopalan V, Smith RA, Ho YH, Lam AK. Signet-ring cell carcinoma of colorectum–current perspectives and molecular biology. *Int J Colorectal Dis*. 2011;26:127–133. https://doi.org/10.1007/s00384-010-1037-z. Epub 2010 Aug 5. PMID: 20686774.
35. Sung CO, Seo JW, Kim KM, Do IG, Kim SW, Park CK. Clinical significance of signet-ring cells in colorectal mucinous adenocarcinoma. *Mod Pathol*. 2008;21:1533–1541. https://doi.org/10.1038/modpathol.2008.170. Epub 2008 Oct 10. PMID: 18849918.
36. Nojadeh JN, Behrouz Sharif S, Sakhinia E. Microsatellite instability in colorectal cancer. *EXCLI J*. 2018;17:159–168. https://doi.org/10.17179/excli2017-948. PMID: 29743854; PMCID: PMC5938532.
37. Saridaki Z, Souglakos J, Georgoulias V. Prognostic and predictive significance of MSI in stages II/III colon cancer. *World J Gastroenterol*. 2014;20:6809–6814. https://doi.org/10.3748/wjg.v20.i22.6809. PMID: 24944470; PMCID: PMC4051919.
38. Lin A, Zhang J, Luo P. Crosstalk between the MSI status and tumor microenvironment in colorectal cancer. *Front Immunol*. 2020;11:2039. https://doi.org/10.3389/fimmu.2020.02039. PMID: 32903444; PMCID: PMC7435056.
39. Kang S, Na Y, Joung SY, Lee SI, Oh SC, Min BW. The significance of microsatellite instability in colorectal cancer after controlling for clinicopathological factors. *Medicine (Baltim)*. 2018;97:e0019. https://doi.org/10.1097/MD.0000000000010019. PMID: 29489646; PMCID: PMC5851768.