



Published in final edited form as:

Arch Surg. 2011 October ; 146(10): 1128–1134. doi:10.1001/archsurg.2011.141.

Receipt of Appropriate Surgical Care for Medicare Beneficiaries with Cancer

Caprice C. Greenberg, MD MPH^{1,2}, Stuart Lipsitz, ScD¹, Bridget Neville, MPH², Haejin In, MD MBA^{1,2}, Nathanael Hevelone, MPH¹, Stacy Porter³, Christine Weeks, MD¹, Ashish Jha, MD⁴, Atul Gawande, MD MPH^{1,4}, Deborah Schrag, MD², and Jane C. Weeks, MD MSc²

¹ Center for Surgery and Public Health, Department of Surgery, Brigham and Women's Hospital, Boston MA

² Center for Outcomes and Policy Research, Department of Medical Oncology, Dana-Farber Cancer Institute, Boston MA

³ Harvard Medical School, Boston MA

⁴ Department of Health Policy and Management, Harvard School of Public Health, Boston, MA

Abstract

Context—Appropriate care is often not defined and when defined, is often not uniformly provided across institutions or demographic populations in the current American healthcare system.

Objective—We sought to investigate receipt of appropriate surgical care in Medicare beneficiaries with cancer.

Design and Setting—Retrospective cohort study using national Surveillance Epidemiology and End Results (SEER) registry linked to Medicare claims data.

Patients—Fee-for-service Medicare patients aged 65 years or older who underwent a definitive surgical resection for breast, colon, gastric, rectal, or thyroid cancers that were diagnosed between 2000 and 2005. Claims data were available from 1999 through 2007.

Main Outcome Measures—Receipt of care concordant with established practice guidelines in surgical oncology in the aggregate and by hospital.

Results—Concordance with guidelines was > 90% for seven of eleven measures. All guidelines regarding adjuvant therapy had concordance rates >90%. Only two of five measures for nodal management had concordance rates >90%. At least 50% of hospitals provided guideline-concordant care to 100% of their patients for 6 of 11 guidelines. Patients receiving appropriate care tended to be younger, healthier, white, more affluent, have less advanced disease, and live in the midwest.

Corresponding Author: Caprice C. Greenberg, MD, MPH Brigham and Women's Hospital Center for Surgery and Public Health One Brigham Circle 1620 Tremont Street, 4-020 Boston, MA 02115 Phone: 617-525-7300 Fax: 617-525-7723 ccgreenberg@partners.org.

Authors' contributions: conception and study design (Greenberg, Lipsitz, Jha, Gawande, J. Weeks), data acquisition (Greenberg, Neville, Hevelone), data analysis (Greenberg, Lipsitz, Neville, Hevelone, Porter, In, C. Weeks, Schrag), data interpretation (all authors), drafting of manuscript (Greenberg, Porter, C. Weeks, In), critical review of manuscript (Lipsitz, Neville, Hevelone, Jha, Gawande, Schrag, J. Weeks), final approval of published version (all authors). Dr. Greenberg had full access to all of the data in the study and takes personal responsibility for the integrity of the data and the accuracy of the data analysis. The authors report no financial conflicts of interest.

The authors would like to acknowledge the contribution of Scott Regenbogen, MD, MPH to study conception and design and Katherine Corso MPH for her assistance in manuscript preparation.

Conclusions—We found a high level of concordance with guidelines in some domains of surgical oncology care, but far less so in others, particularly for gastric and colon nodal management. Given the current national focus on improving the quality of healthcare, surgeons must focus on generating data to define appropriate care and translating that data into everyday practice.

INTRODUCTION

There is currently a major focus on improving the quality of healthcare in America. Quality healthcare means delivering the right care to the right patient at the right time. In order to ensure such care is provided, we must first know what “right” or appropriate care is and then be able to determine whether that care was provided. Currently, it is well-documented that the practice of healthcare in America varies widely across both institutions and demographic populations; this is true across specialties and disease sites.¹⁻² This variation may reflect two things: (1) a lack of knowledge about the optimal approach to care or (2) a lack of acceptance regarding currently defined standards of care.

We sought to use practice guidelines to determine whether appropriate surgical care was provided to Medicare beneficiaries with a new diagnosis of cancer. We investigated factors associated with likelihood of receiving guideline-concordant care at a patient-level. We then analyzed guideline concordance at an aggregate patient level to determine the degree of variation in hospital-level performance to study whether each guideline has been accepted into routine practice.

METHODS

Identification of Existing Clinical Practice Guidelines

We organized our study around the three areas in which a surgeon treating cancer must be proficient: 1) surgical management of the primary tumor; 2) evaluation and treatment of regional nodal basins; and 3) appropriate referral for multidisciplinary adjuvant therapy. We first identified existing disease-specific guidelines in each of these areas for five common cancers in which surgery plays an important role, including breast, colon, gastric, rectal, and thyroid cancers, based on the following criteria:

1. The guideline must be endorsed by a professional organization or society whose members are considered experts in that disease, addressed in an NIH Consensus Statement, or included as a recommendation in the AJCC Cancer Staging Manual.³
2. Cancer registry and/or administrative data must be sufficient to define concordance, including both the numerator (patients who received the recommended care) and the denominator (patients meeting inclusion criteria for the recommended care).
3. All other treatments must be considered inappropriate care.

In order to classify the strength of a given guideline we used the classification system published by the National Comprehensive Cancer Network.⁴ The NCCN clinical practice guidelines are readily available, widely used, span multiple cancer sites and classify the strength of both the evidence and panel members’ consensus supporting a recommendation. The NCCN identifies four guideline categories:

- 1 – The recommendation is based on high level of evidence and there is uniform NCCN consensus
- 2A – The recommendation is based on lower level of evidence and there is uniform NCCN consensus

- 2B** – The recommendation is based on lower level of evidence and there is non-uniform NCCN consensus (but no major disagreement)
- 3** – The recommendation is based on any level of evidence but reflects major disagreement.

Identification of the Overall Cohort

The Surveillance Epidemiology and End Results (SEER) program of the National Cancer Institute collects detailed data including stage at diagnosis, tumor morphology, first course of treatment and demographic variables for persons with cancer living in a SEER region.⁵ SEER data has been linked to Medicare enrollment and claims data in order to facilitate evaluation of treatment and outcomes of care across the cancer continuum.⁶

Patients diagnosed with breast, colon, gastric, rectal, or thyroid cancers as their first or only cancer between the years 2000 and 2005 were identified and linked to Medicare claims data from 1999 through 2007. We required that all patients have at least one cancer-specific surgical procedure between 30 days before through 365 days after diagnosis and be associated with a non-missing hospital identifier. Patients were also required to be continuously enrolled in Medicare and have no HMO enrollment during this time period.

Between 2000 and 2005, there were a total of 181,534 patients with breast cancer, 116,383 with colon cancer, 22,625 with gastric cancer, 41,857 with rectal cancer, and 10,453 with thyroid cancer identified. Criteria used to specify the cohort are shown in Figure 1. We excluded patients under 65 years of age at the time of diagnosis and those with end stage renal disease or disability as their qualifying condition. Patients with Stage IV or unknown stage (except for thyroid cancer) or a discrepancy in the death date reported in SEER and Medicare of greater than 3 months were also excluded.

Definition of Specific Concordance Measures

Additional specific inclusion and exclusion criteria for the denominator and numerator for each guideline were applied and are available from the authors upon request. For situations where there was a discrepancy or ambiguity between identified guidelines regarding appropriate criteria, we performed sensitivity analyses varying the inclusion criteria. Tumor size criteria for receipt of total thyroidectomy were varied between 1 and 2cm. As conclusions did not change, results are only presented for a threshold of 1.5cm. Because this study was focused specifically on the surgical practitioner, guidelines that related to the receipt of adjuvant therapy were expanded to include a visit with a medical oncologist or radiation oncologist regardless of whether the patient received treatment. Medical oncologists were identified by an associated Medicare specialty code for medical oncology or hematology/oncology or having submitted a claim for chemotherapy.⁷⁻⁹ Radiation oncologists were defined by a Medicare specialty code for radiation oncology or having submitted a claim for radiation planning or administration.

Analysis

Concordance is measured as the proportion of patients who met the numerator criteria among all patients who met the denominator criteria for each guideline. The concordance rate for each guideline was determined at the aggregate patient level, by considering all patients with the relevant condition and at the hospital level, by considering all hospitals that treated at least 5 patients with the relevant condition. This was done to ensure stability in the estimate of the hospital concordance rate. While patients treated at hospitals with fewer than 5 patients were excluded from the hospital-level analysis, they were included in the aggregate patient-level analysis. We report aggregate patient-level concordance rates with

95% confidence intervals. Generalized estimating equations were used to account for clustering at the hospital level when calculating 95% confidence intervals.¹⁰ For the hospital level analysis, we determined the proportion of hospitals for which 100% of treated patients received guideline-concordant care, as well as exact binomial 95% confidence intervals. For each of the guidelines meeting our eligibility criteria, logistic regression models were built to investigate covariates associated with receipt of concordant care. Generalized estimating equations were used to estimate the logistic regression parameters, which allowed us to investigate both patient and known institutional factors while accounting for clustering at the institutional level. Because the logistic regression analyses were exploratory, we did not adjust the type I error rate to account for multiple comparisons; thus, p values should be interpreted cautiously. The number of significant associations ($p < 0.05$) in these analyses is greater than expected by chance if there were no associations in the data. All tests were two-tailed. All data management and statistical analyses were performed using SAS v. 9.2 (Cary, NC).

RESULTS

Identification of Guidelines and Cohort

Eleven guidelines were identified and are listed in Table 1, showing the original source, year it was proposed, and whether the National Comprehensive Cancer Network grades them as based on a high level of evidence (Grade 1) or uniform consensus but a lower level of evidence (Grade 2A).^{4, 11-18} Given that we are using guidelines to represent appropriate care, we did not include any guidelines for which there was non-uniform consensus or major disagreement. The 11 guidelines include five for breast cancer, two for colon cancer, one for gastric cancer, one for rectal cancer, and two for thyroid cancer. They represent the following domains of surgical oncology care: 1 measure for surgery directed at the primary cancer, 5 measures for nodal management and 5 measures related to adjuvant therapy.

Table 2 provides a description of the cohort for each disease site included in this analysis.

Aggregate Patient-Level Concordance Rates

Concordance with guidelines was greater than 90% for seven of the eleven measures (see Table 3). All guidelines regarding evaluation for or receipt of adjuvant therapy had concordance rates $>90\%$. The measures that evaluated breast cancer management had the highest concordance rates. These included radiation therapy or evaluation following breast conserving surgery (BCS) (99.2%), chemotherapy or medical oncology evaluation for estrogen receptor (ER) negative breast cancer (98.1%), axillary dissection for node positive breast cancer (96.7%), post-mastectomy radiation therapy (PMRT) or evaluation for >4 positive nodes, stage III, node positive and $T > 5\text{cm}$ (94.9%) cancers and nodal evaluation for invasive breast cancer (91.8%). The lowest concordance rates were seen for central neck dissection for node-positive thyroid papillary cancer (72.7%), colon cancer nodal evaluation of at least 12 nodes (48.5%), and gastric nodal evaluation of at least 15 nodes (32.5%).

Hospital-Level Analysis

As shown in Figure 2, there was wide variation in the proportion of hospitals providing uniformly guideline-concordant care, i.e., treating 100% of their patients according to the guideline recommendation. Fewer than 1% of institutions met this standard for colon cancer nodal evaluation or central neck dissection while 93% provided either evaluation for or treatment with radiation for all patients following BCS. More institutions had uniformly concordant care for guidelines concerning adjuvant therapy than for guidelines dealing with surgery directed at the primary cancer or nodal management.

Predictors of Receipt of Appropriate Care

Table 4 summarizes the factors associated with increased or decreased likelihood of receiving concordant care by guideline. For a number of the guidelines, younger age, less aggressive disease, white race, higher income level, being married, and care at a hospital that participates in an oncology group were independently associated with higher guideline concordance. Patients treated in the Midwest were more likely than patients treated in the Northeast to receive guideline-concordant care for all measures; the direction of the effect for the South and West (as compared to the referent Northeast) varied by measure.

DISCUSSION

Using practice guidelines to define appropriate care, we found that over 90% of all patients received recommended care for 7 of the 11 guidelines examined at an aggregate patient level. We also identified 6 measures for which at least half the institutions were concordant with the guidelines 100% of the time. These high concordance rates suggest that the factors influencing clinical decision-making are adequately captured in the current guidelines and surgeons recognize the importance of these therapies. As a result, we were able to demonstrate that Medicare beneficiaries are highly likely to receive appropriate care and that this finding is consistent across hospitals.

For several of the guidelines related to nodal management, however, concordance rates were low and few hospitals provided appropriate care to all patients. Specifically, we found that published guidelines recommending examination of a minimum number of lymph nodes in colon and gastric cancer have not resulted in routine adoption of this practice in elderly Americans despite the inclusion of these recommendations in the AJCC staging manual. It is possible that despite apparent expert consensus regarding the importance of evaluation of a minimum number of nodes, the lack of definitive evidence supporting a particular threshold has left many practitioners (surgeons and/or pathologists) unconvinced. They may not believe the potential benefits of more extensive lymph node harvests outweigh the added operative risks, perhaps particularly in older patients. The limitations of the evidence base supporting nodal evaluation and the difficulty in defining an appropriate nodal threshold for use in guidelines or quality measurement are well-described in two recent reviews, a report from the Cochrane Collaboration regarding extent of lymph node dissection for gastric cancer and a meta-analysis on colon cancer nodal evaluation.¹⁹⁻²⁰

We observed a somewhat different pattern of concordance for central neck dissection for node positive papillary thyroid cancer, with a higher aggregate concordance rate of 71.6%. In the hospital-level analysis, but no institutions at all performing neck dissections in all of their patients. Instead, most institutions performed them in approximately 80% of patients. One possible explanation for these findings is that there is agreement that a central neck dissection constitutes appropriate care for most but not all elderly patients with node positive papillary thyroid cancer, and that the factors relevant to selecting patients for the procedure are not adequately captured by the current guideline inclusion criteria, at least in this population of elderly Americans. Such factors might include the presence of a macro-rather than micrometastasis, or nodes detectable pre-operatively or at the time of surgery, rather than only on post-operative pathology review.

Several other explanations could account for the patterns observed in this study. Prior research on the Hospital Quality Alliance measures has suggested an association between the length of time guidelines have been in place and concordance rates.²¹ We found high concordance rates and minimal variation for the breast measures, guidelines that were among the first to be developed in cancer care.

It's also possible that because referral for adjuvant therapy is a dichotomous decision, which facilitates both compliance and measurement, it's easier for institutions to achieve consistent and high concordance rates on these measures than on guidelines measuring a continuous outcome such as number of lymph nodes examined. Alternatively or in addition, the fact that referrals result in reimbursement for the provider or the health system while lymph node harvests do not may contribute to the higher concordance we observed for guidelines related to adjuvant therapy.

Finally, there is a suggestion in our data that the level of evidence on which a guideline is based is associated with the level of concordance. Six of the 8 guidelines with concordance rates > 90% were based on high level evidence and all 3 of the guidelines with concordance rates <90% were based on lower level evidence (Table 3). Similarly, 5 of the 6 guidelines where more than 50% of institutions provided concordant care to 100% of their patients were graded as 1, while 4 of the 5 guidelines for which less than half of the institutions provided fully concordant care were graded 2 (Figure 2). The role that level of evidence may play in acceptance of guidelines deserves further investigation.

This study has the usual limitations associated with analyses of large national databases, including incomplete capture of cases, loss of follow up, and missing data as well as the limitations of claims data such as variation in billing practices and coding inaccuracies. Additionally, because Medicare provides the only consistent and comprehensive national data source on medical services delivered, our analysis was limited to patients 65 and older. A focus on older patients did allow us to study a particularly vulnerable population and one that makes up the majority of patients with gastric and colorectal cancer. However, breast and thyroid cancer are common in younger Americans, so our reliance on Medicare data represents a more significant limitation for these diagnoses.

We found a high level of concordance with guidelines in some domains of surgical oncology care, but far less so in others, especially those that are associated with nodal management. Five of the six measures with wide acceptance into practice relate to appropriate referral for or receipt of adjuvant therapy. Given the current national focus on quality in healthcare, there is increasing pressure to develop measures to determine whether patients are getting appropriate care; however, within the surgical disciplines, there is a paucity of data to support what constitutes appropriate care. It is critical that surgeons focus on the generation of the data necessary to inform clinical decision-making and promote quality surgical care.

Acknowledgments

This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of the Applied Research Program, NCI; the Office of Research, Development and Information, CMS; Information Management Services (IMS), Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER-Medicare database.

This work was supported in part by a grant from the American Surgical Association Foundation. The sponsor did not have any role in design or conduct of the study or manuscript preparation or review.

REFERENCES

1. Jencks SF, Huff ED, Cuedon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. *JAMA*. Jan 15; 2003 289(3):305-312. [PubMed: 12525231]
2. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. *N Engl J Med*. Jun 26; 2003 348(26):2635-2645. [PubMed: 12826639]
3. Greene, FL.; Page, DL.; Fleming, ID. American Joint Commission on Cancer. *AJCC Cancer Staging Manual*. Sixth ed.. Lippincott Raven Publishers; Philadelphia, PA: 2002. al. e

4. National Comprehensive Cancer Network. NCCN Categories of Evidence and Consensus.; Accessed January. 2010. p. 2011http://www.nccn.org/professionals/physician_gls/categories_of_consensus.asp
5. [June 15, 2010] Surveillance Epidemiology and End Results.. <http://seer.cancer.gov>.
6. [June 15, 2010] SEER-Medicare: Brief Description of the SEER-Medicare Database.. <http://healthservices.cancer.gov/seermedicare/overview/>.
7. Baldwin LM, Adamache W, Klabunde CN, Kenward K, Dahlman C, J LW. Linking physician characteristics and medicare claims data: issues in data availability, quality, and measurement. *Med Care*. Aug; 2002 40(8 Suppl):IV-82-95. [PubMed: 12187173]
8. Keating NL, Landrum MB, Ayanian JZ, Winer EP, Guadagnoli E. Consultation with a medical oncologist before surgery and type of surgery among elderly women with early-stage breast cancer. *J Clin Oncol*. Dec 15; 2003 21(24):4532-4539. [PubMed: 14673040]
9. Pollack LA, Adamache W, Ehemann CR, Ryerson AB, Richardson LC. Enhancement of identifying cancer specialists through the linkage of Medicare claims to additional sources of physician specialty. *Health Serv Res*. Apr; 2009 44(2 Pt 1):562-576. [PubMed: 19207588]
10. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. Mar; 1986 42(1):121-130. [PubMed: 3719049]
11. American Joint Commission on Cancer. Manual for Staging of Cancer. Fourth ed.. Beahrs, OH.; Henson, DE.; Hutter, RVP.; Kennedy, BJ., editors. J.B. Lippincott Company; Philadelphia: 1992.
12. Beahrs, OH.; Carr, DT.; Rubin, P. Manual for Staging of Cancer 1977. American Joint Committee; Chicago: 1977.
13. Fleming, ID.; Cooper, JS.; Henson, DE., et al. AJCC Cancer Staging Manual. Fifth ed.. Lippincott-Raven; Philadelphia: 1997.
14. Adjuvant Therapy for Patients with Colon and Rectum Cancer. NIH Consensus Statement Online. Apr 16-18; 1990 8(4):1-25.
15. Treatment of Early-Stage Breast Cancer. NIH Consensus Statement Online. June 18-21; 1990 8(6): 1-19.
16. Cooper DS, Doherty GM, Haugen BR, et al. Management Guidelines for Patients with Thyroid Nodules and Differentiated Thyroid Cancer. The American Thyroid Association Guidelines Task Force. *Thyroid*. 2006; 16(2):109-142. [PubMed: 16420177]
17. National Comprehensive Cancer Network: Oncology Practice Guidelines. *Oncology*. 1996; 10(Supplement 11):3-317.
18. Carlson, RW. Update: Breast Cancer Guidelines NCCN Practice Guidelines Version 1.98.. Paper presented at: The National Comprehensive Cancer Network Third Annual Conference; 1998.
19. Chang GJ, Rodriguez-Bigas MA, Skibber JM, Moyer VA. Lymph node evaluation and survival after curative resection of colon cancer: systematic review. *J Natl Cancer Inst*. Mar 21; 2007 99(6): 433-441. [PubMed: 17374833]
20. McCulloch P, Nita ME, Kazi H, Gama-Rodrigues J. Extended versus limited lymph nodes dissection technique for adenocarcinoma of the stomach. *Cochrane Database Syst Rev*. 2003; (4):CD001964. [PubMed: 14583942]
21. Jha AK, Li Z, Orav EJ, Epstein AM. Care in U.S. hospitals--the Hospital Quality Alliance program. *N Engl J Med*. Jul 21; 2005 353(3):265-274. [PubMed: 16034012]

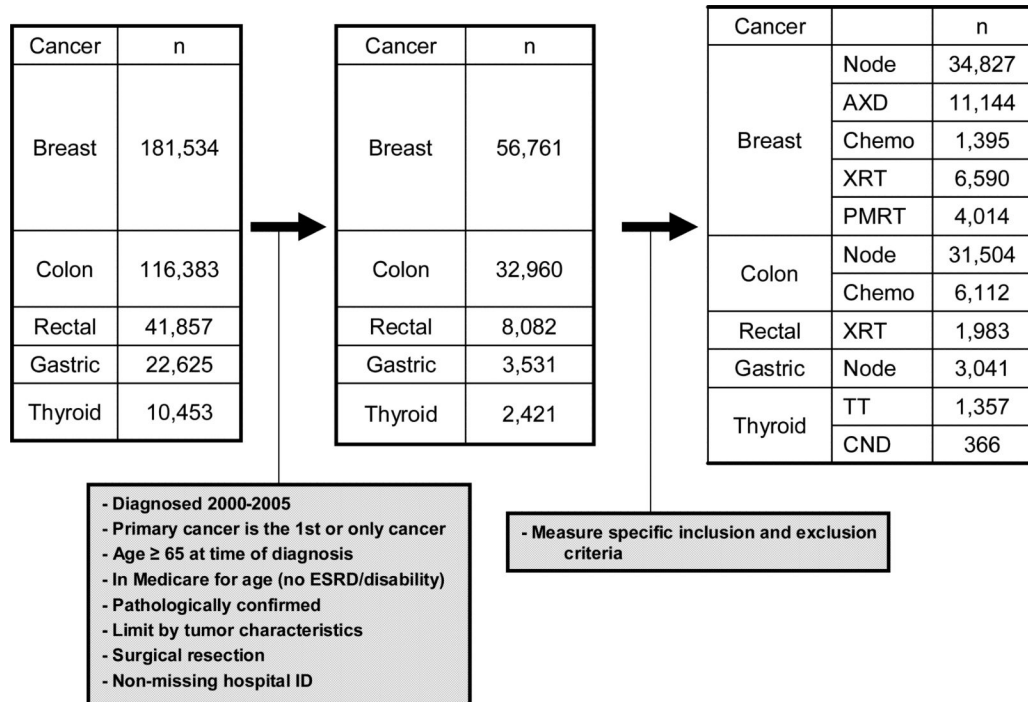
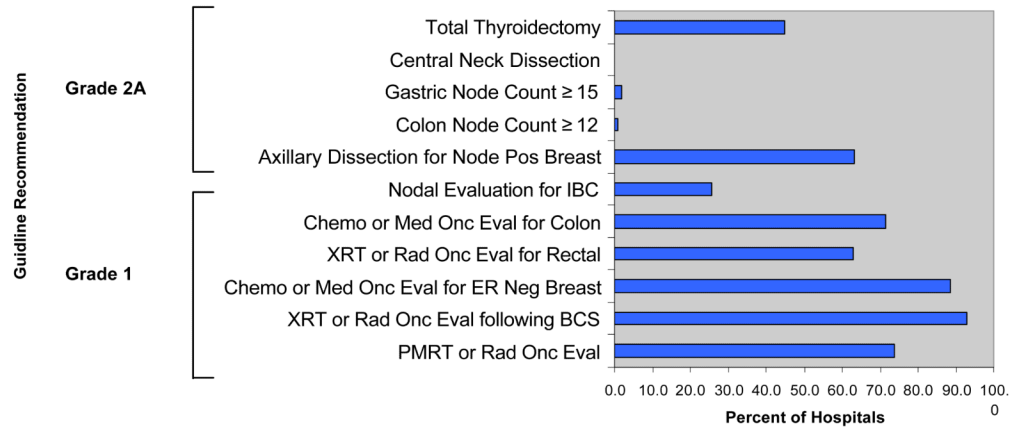


Figure 1. Cohort Identification

The first box represents all cases of cancer identified within the dataset for each disease site. The second box depicts the cohort for each disease site after general restrictions. The final box represents the cohort for each guideline following guideline-specific criteria. These guideline-specific cohorts are not mutually exclusive and may overlap since more than one guideline may be applicable to a given patient.

Proportion of Hospitals with 100% Guideline Concordant Care



Abbreviations: Pos – positive; IBC – invasive breast cancer; Chemo – chemotherapy; Med onc eval – medical oncology evaluation; XRT – radiation therapy; Rad onc eval – radiation oncology evaluation; ER – estrogen receptor; Neg – negative; BCS – breast conserving surgery; PMRT – post-mastectomy radiation therapy

Figure 2. Hospital Distribution of 100% Concordance for Each Guideline

The graph depicts the percentage of hospitals with ≥5 case volume providing guideline concordant care to 100% of eligible patients treated.

Table 1
Current Guidelines for Surgical Oncology Care

Guidelines are grouped according to the domain of surgical oncology expertise represented. The original source of the guideline as well as the year it was first suggested and whether the 2009 NCCN categories grade the guideline as based on high level evidence (1) or expert consensus with lower level evidence (2A) are presented.

Domain	Guideline Recommendation	Original Source	Original Year	NCCN Grade
<i>Surgery Directed at Primary Cancer</i>	<i>Total thyroidectomy for papillary cancer \geq 1.5cm or node positive</i>	ATA Guidelines	1996	2A
<i>Nodal Management</i>	<i>Central neck dissection for node positive papillary cancer</i>	ATA Guidelines	1996	2A
	<i>Gastric node count \geq 15</i>	AJCC Cancer Staging Manual 5 th Edition	1997	2A
	<i>Colon node count \geq 12</i>	AJCC Cancer Staging Manual 4 th Edition	1992	2A
	<i>Axillary Dissection for Node Positive Breast Cancer</i>	NCCN Guideline	1998	2A
	<i>Nodal evaluation for invasive breast cancer</i>	AJCC Cancer Staging Manual 1 st Edition	1977	1
<i>Referral for/Receipt of Adjuvant Therapy</i>	<i>Chemo or med onc evaluation for Stage III colon cancer (age <80)</i>	NIH Consensus	1990	1
	<i>XRT or rad onc evaluation for T4 or Stage III rectal cancer (age <80)</i>	NIH Consensus	1990	1
	<i>Chemo or med onc eval for ER negative breast cancer (age <70)</i>	NCCN Guideline	1996	1
	<i>XRT or rad onc evaluation following BCS (age <70)</i>	NIH Consensus	1990	1
	<i>PMRT or rad onc evaluation for >4 + nodes, Stage III, node positive and T>5cm</i>	NCCN Guideline	1996	1

Abbreviations: Chemo – chemotherapy; Med onc – medical oncology; XRT – radiation therapy; Rad onc – radiation oncology; ER – estrogen receptor; BCS – breast conserving surgery; PMRT – post-mastectomy radiation therapy

Table 2

Patient and hospital characteristics for each cohort

This table includes patients that were included in at least one of the 11 measures by site of disease. Because patients may have been included in the cohort for more than one measure, the number represented here does not necessarily represent the sum of the final cohorts depicted in the last box of Figure 1.

	Breast	Colon	Rectal	Gastric	Thyroid
Patient Characteristics					
Total (n)	37,941	31,647	1,983	3,037	1,357
Median Diagnosis Age	74.8	77.9	72.4	76.5	73.4
Female	37,941 (100%)	18,638 (59%)	907 (46%)	1,320 (43%)	951 (70%)
Race					
White	34,048 (90%)	27,802 (88%)	1,723 (87%)	2,138 (70%)	1,171 (86%)
Black	2,515 (6%)	2,352 (7%)	136 (7%)	333 (11%)	63 (5%)
API	1,166 (3%)	1,352 (4%)	114 (5%)	546 (18%)	123 (9%)
Other	212 (1%)	141 (1%)	10 (1%)	20 (1%)	N/A
Married	16,533 (44%)	15,422 (49%)	1,199 (60%)	1,742 (57%)	755 (56%)
Charlson					
0	25,715 (68%)	18,683 (59%)	1,353 (68%)	1,644 (54%)	901 (66%)
1	8,139 (21%)	7,674 (24%)	428 (22%)	828 (27%)	298 (22%)
2	2,239 (6%)	2,744 (9%)	121 (6%)	281 (9%)	90 (7%)
3+	1,848 (5%)	2,546 (8%)	81 (4%)	284 (9%)	68 (5%)
Median Income	\$38,518 [†]	\$36,863 [†]	\$37,918 [†]	\$37,142 [†]	\$40,305
Median Percent College	28.2%	27.5%	27.8%	26.6%	27.5%
Stage					
I	17,962 (47%)	7,654 (24%)	N/A	1,123 (37%)	537 (40%)
II	16,405 (43%)	13,281 (42%)	173 (9%)	1,914 (63%)	689 (51%)
III	3,574 (9%)	10,712 (34%)	1,810 (91%)	N/A	N/A
IV	N/A	N/A	N/A	N/A	131 (10%)

	Breast	Colon	Rectal	Gastric	Thyroid
Hospital Characteristics					
Total (n)	1,314	1,182	593	612	476
Region					
Northeast	219 (17%)	187 (16%)	111 (19%)	118 (19%)	90 (19%)
South	329 (25%)	273 (23%)	123 (21%)	127 (21%)	91 (19%)
Midwest	223 (17%)	210 (18%)	79 (13%)	59 (10%)	57 (12%)
West	543 (41%)	512 (43%)	280 (47%)	308 (50%)	238 (50%)
Urban	918 (70%)	826 (70%)	494 (83%)	532 (87%)	418 (88%)
NCI Center	25 (2%) †	27 (2%) †	25 (4%)	24 (4%)	23 (5%)
Oncology Group	533 (41%)	487 (41%)	335 (56%)	347 (57%)	288 (60%)
Control	†	†			
Non-profit	764 (58%)	696 (59%)	401 (68%)	413 (67%)	330 (69%)
Proprietary	260 (20%)	227 (19%)	103 (17%)	114 (19%)	80 (17%)
Governmental	288 (22%)	257 (22% 0)	89 (15%)	85 (14%)	66 (14%)
Teaching	460 (36%) †	434 (38%) †	260 (44%) †	291 (48%) †	228 (49%) †
Hospice	567 (43%) †	520 (44%) †	264 (45%)	260 (42%)	224 (47%)

†Less than 1% of data missing

Table 3
Population-level concordance rate for each guideline in surgical oncology care

The guidelines are arranged according to the domain of surgical care to which they apply. For each guideline, the NCCN grade is depicted. The aggregate patient level results show the proportion and overall concordance rate for all patients in the cohort. At the hospital level, the proportion of hospitals providing concordant care to 100% of eligible patients is depicted in the first column followed by the range of observed institutional concordance rates.

Domain	Guidelines		Aggregate Patient-Level Analysis			Hospital Level Analysis*	
	Recommendation	NCCN Grade	Proportion	Overall Concordance (%)	95% CI	Proportion with 100% Concordance	Range of Concordance Rates
<i>Surgery Directed at Primary Cancer</i>	<i>Total Thyroidectomy</i>	2A	1213/1357	89.4%	87.7 - 91.0%	35/78	60.0 - 100%
	<i>Central Neck Dissection</i>	2A	266/366	72.7%	67.8 - 77.6%	<11 [‡]	20.0 - 91.7%
	<i>Gastric Node Count ≥ 15</i>	2A	987/3041	32.5%	29.2 - 35.7%	<11 [‡]	0.0 - 100%
<i>Nodal Management</i>	<i>Colon Node Count ≥ 12</i>	2A	15266/31504	48.5%	46.7 - 50.3%	<11 [‡]	0.0 - 100%
	<i>Axillary Dissection for Node Pos IBC</i>	2A	10779/11144	96.7%	96.3 - 97.1%	339/536	66.7 - 100%
	<i>Nodal Evaluation for IBC</i>	1	31979/34827	91.8%	91.3 - 92.4%	203/794	50.0 - 100%
	<i>Chemo or Med Onc Eval for Colon</i>	1	5917/6112	96.8%	96.4 - 97.3%	300/419	66.7 - 100%
<i>Referral for/Receipt of Adjuvant Therapy</i>	<i>XRT or Rad Onc Eval for Rectal</i>	1	1812/1983	91.4%	89.9 - 92.8%	92/146	55.6 - 100%
	<i>Chemo or Med Onc Eval for ER Neg Breast</i>	1	1369/1395	98.1%	97.4 - 98.9%	70/79	71.4 - 100%
	<i>XRT or Rad Onc Eval following BCS</i>	1	6535/6590	99.2%	98.9 - 99.4%	345/372	80.0 - 100%
	<i>PMRT or Rad Onc Eval</i>	1	3808/4014	94.9%	94.0 - 95.7%	222/301	0.0 - 100%

Abbreviations: Pos – positive; IBC – invasive breast cancer; Chemo – chemotherapy; Med one eval – medical oncology evaluation; XRT – radiation therapy; Rad one eval – radiation oncology evaluation; ER – estrogen receptor; Neg – negative; BCS – breast conserving surgery; PMRT – post-mastectomy radiation therapy

* Among hospitals with ≥5 cases included in the cohort under analysis.

‡ Data reported limited by SEER-Medicare confidentiality rules.

Table 4
Results of Multivariable Analysis of Factors Associated with Receipt of Appropriate Care

The grid below presents the direction of odds ratios that were significant ($p < 0.05$) in the logistic regression model. An arrow directed downward represents an odds ratio below 1 (indicating that variable is associated with a lower rate of concordance) and an arrow upward represents an odds ratio above 1 (indicating that variable is associated with a higher rate of concordance). Region is the only categorical variable presented. Northeast region was used as the reference group (as indicated by REF in the table).

	Total Thyroidectomy	Central Neck Dissection	Gastric Node Count \geq 15	Colon Node Count \geq 12	Axillary Dissection for Node Pos Breast	Nodal Evaluation for IBC	Chemo or Med Onc Eval for Colon	XRT or Rad Onc Eval for Rectal	Chemo or Med Onc Eval for ER Neg Breast	XRT or Rad Onc Eval following BCS	PMRT or Rad Onc Eval
Patient Characteristics											
Increasing Age	↓		↓	↓	↓	↓	↓	↓			↓
Male			↓								
White			↓				↑		↑	↑	
Married				↑		↑	↑	↑			
Charlson Score Zero			↑	↑		↑					
Income	↑		↑	↑	↓					↑	
Education											
Increasing Stage	↓	↓	↓	↓	↓	↓					
Hospital Characteristics											
Region											
Northeast	Ref					Ref	Ref	Ref		Ref	Ref
South	↑					↑	↓	↓		↓	↓
Midwest	↑					↑	↑	↑		↑	↑
West	↑					↑	↓	↓		↓	↑
Urban											↑
NCI Center		↑	↑	↑							
Oncology Group			↑	↑				↑		↑	↑
Non-profit Status			↑	↑							
Teaching Hospital	↑		↑	↑	↓	↓	↓				
Hospice											

Abbreviations: XRT – radiation therapy; Rad onc eval – radiation oncology evaluation; Chemo – chemotherapy; Med onc eval – medical oncology evaluation; IBC – invasive breast cancer; Pos – positive; BCS – breast conserving surgery; PMRT – post-mastectomy radiation therapy; ER – estrogen receptor; Neg – negative