Published in final edited form as:

Cancer Causes Control. 2014 November ; 25(11): 1503–1512. doi:10.1007/s10552-014-0451-x.

Postoperative 30-day mortality in patients undergoing surgery for colorectal cancer: development of a prognostic model using administrative claims data

S. de Vries, MSc.¹, D.B. Jeffe, Ph.D.^{1,2}, N. O. Davidson, M.D.^{2,3}, A.D. Deshpande, Ph.D.^{1,2}, and M. Schootman, Ph.D.^{2,4}

¹Washington University School of Medicine, Department of Medicine, Division of Health Behavior Research, Saint Louis, MO

²Alvin J. Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine, Saint Louis, MO

³Washington University School of Medicine, Department of Medicine, Division of Gastroenterology, Saint Louis, MO

⁴Saint Louis University, College for Public Health and Social Justice, Department of Epidemiology, St. Louis, MO

Abstract

Purpose—To develop a prognostic model to predict 30-day mortality following CRC surgery using the Surveillance, Epidemiology, and End Results (SEER)-Medicare linked data, and to assess whether race/ethnicity, neighborhood, and hospital characteristics influence model performance.

Methods—We included patients aged 66 years and older from the linked 2000–2005 SEER-Medicare database. Outcome included 30-day mortality, both in-hospital and following discharge. Potential prognostic factors included tumor, treatment, sociodemographic, hospital, and neighborhood characteristics (census-tract-poverty rate). We performed a multilevel logistic regression analysis to account for nesting of CRC patients within hospitals. Model performance was assessed using the area under the receiver operating characteristic curve (AUC) for discrimination and the Hosmer-Lemeshow goodness-of-fit test for calibration.

Results—In a model that included all prognostic factors, important predictors of 30-day mortality included age at diagnosis, cancer stage and mode of presentation. Race/ethnicity, census-tract-poverty rate and hospital characteristics were independently associated with 30-day mortality, but they did not influence model performance. Our SEER-Medicare model achieved moderate discrimination (AUC=0.76), despite suboptimal calibration.

Conclusions—We developed a prognostic model that included tumor, treatment, sociodemographic, hospital, and neighborhood predictors. Race/ethnicity, neighborhood, and

Conflict of interest

The authors state that they have no conflict of interest.

hospital characteristics did not improve model performance compared with previously developed models.

Keywords

colorectal cancer; mortality; prognostic model; administrative claims; SEER-Medicare

Introduction

Colorectal cancer (CRC) is the second most common malignancy in the United States, and the majority of patients will receive surgical treatment (1). Providing high-quality care to CRC patients is important, but extensive variability in mortality rates exists among hospitals (2, 3). Public reporting of 30-day mortality measures may assist hospitals in their quality improvement efforts, and increase transparency of hospital care. For such audit purposes, prognostic models can be used to compare hospitals, in which 30-day mortality has been commonly used to capture quality of care (4, 5). Accurate comparison of the quality of care for CRC patients among hospitals requires an accurate model for risk stratification, including proper adjustments for differences in case mix (6).

Several models have been developed to predict postoperative mortality in CRC patients, including the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) model (7) and its Portsmouth (P-POSSUM)(8) and colorectal (CR-POSSUM)(9) modifications, the Association of Coloproctology of Great Britain and Ireland (ACPGBI) model (10), and the Cleveland Clinic Foundation Colorectal Cancer Model (CCF-CCM) (11). Several studies have shown that the POSSUM models either under- or overestimated true mortality risk following CRC surgery (12–15). Another limitation of the POSSUM models is the frequent lack of available data about the physiological predictors; incomplete data may result in inaccurate mortality predictions (13). The ACPGBI model was developed in response to the limitations posed by the POSSUM models and includes only a limited number of variables (i.e., age, American Society of Anesthesiologists [ASA] grade, Dukes cancer stage, and operative urgency). The ACPGBI model achieved the same performance as the CR-POSSUM model without including any physiological predictors (16). The CCF-CCM model assesses operative mortality in CRC patients, but used data gathered over three decades, which may underestimate improvements in operative management and surgical procedures over time (17). Moreover, all these aforementioned models omit important sociodemographic risk factors (i.e. race/ethnicity) associated with adverse outcomes after CRC surgery. These sociodemographic factors are important considerations in translating the models to different CRC patient populations, especially in the US, where healthcare is not universally available and varies substantially by race/ethnicity and geography.

Recent evidence suggests that both race/ethnicity and neighborhood characteristics are associated with adverse short-term outcomes following CRC surgery. Compared with white Medicare patients, black Medicare patients had increased in-hospital mortality risk (18). Postoperative 30-day mortality risk also was reported to be influenced by social and economic factors that characterize geographic areas (19, 20), and hospital characteristics,

such as hospital volume and caseload (2, 3) To our knowledge, no prognostic models of postoperative 30-day mortality risk in CRC patients have been developed by taking a more holistic and multilevel approach to include neighborhood-level social and economic, and hospital-level characteristics beyond tumor and treatment characteristics that are part of existing models. Accordingly, the aim of this study was to develop a prognostic model to predict 30-day mortality following CRC surgery using the Surveillance, Epidemiology, and End Results (SEER) data linked with Medicare data, and to assess whether individual-level sociodemographic, neighborhood-level, and hospital-level characteristics improve the discrimination and calibration of a prognostic model for CRC patients in the US. In addition, we compared our SEER-Medicare model to a model including similar prognostic factors as the revised ACPGIBI model (21), and assessed the influence of the abovementioned characteristics in a model with few (patient and tumor-related) predictors. With the increasing availability of claims data, there are more opportunities to include these data in developing prognostic models for use in routine medical care (22).

Methods

Data were obtained from an existing linkage of 2000–2005 National Cancer Institute's SEER program data with 1999–2005 Medicare claims. Linked SEER-Medicare data provide a rich source of information on Medicare patients included in SEER, a nationally representative collection of population-based cancer registries (23). Ninety-four percent of cancer patients reported to SEER aged 65 years or older have been successfully linked with Centers for Medicare and Medicaid (CMS) data (23). In this study, data from 12 tumor registries were included, covering approximately 14% of the United States population.

Study population

We selected patients aged 66 years of age or older with a first primary in-situ or invasive colon or rectal cancer diagnosis between 2000 and 2005, with both Medicare parts A and B coverage, and who underwent surgical treatment for CRC. We excluded patients who were identified at autopsy or from death certificate only, patients for whom the month of diagnosis was not available, and patients who were members of a Health Maintenance Organization because claims data are not available. We excluded patients who were only eligible for Medicare because they were disabled or had end-stage renal disease (i.e., patients younger than 65 years of age). The age restriction of 66 years of age or older allowed for one year of complete claims data prior to diagnosis to determine comorbidity before CRC diagnosis. The final sample size included 44,941 patients.

We searched inpatient, outpatient, and carrier claims for surgical treatments, using previously identified Healthcare Common Procedure Coding System (HCPCS) and/or International Classification of Diseases, ninth revision (ICD-9) codes (24). In the event of multiple surgical interventions in one patient record, the most extensive surgery performed and associated date was used.

Prognostic factors

Potential prognostic factors for the development of the first model were chosen based on previous prognostic models and expert opinion (16, 18, 25–28). Tumor characteristics included American Joint Commission on Cancer (AJCC) stage, tumor grade, tumor location, and histology. Tumor location was classified as proximal colon (cecum, ascending); transverse colon (hepatic flexure, transverse colon, splenic flexure); distal colon (descending and sigmoid colon); or rectosigmoid junction or rectum. Histology was defined as mucinous adenocarcinoma (e.g., Mucinous adenocarcinoma; Mucin-producing adenocarcinoma), other adenocarcinoma (e.g., adenocarcinoma, Not Elsewhere Specified or insitu; Adenocarcinoma in tubulo-villous adenoma), and non-adenocarcinoma (e.g., Signet cell carcinoma). Information regarding tumor biology, including microsatellite instability (MSI) status was not available. Treatment characteristics included mode of presentation (emergency or elective surgery), and type of surgery (total, subtotal or partial colectomy). Surgical interventions within 24 hours after emergency admission were defined as emergency surgery; all other surgical treatments were classified as elective. Patient sociodemographic characteristics included sex, age, race/ethnicity, and comorbidity. We searched the inpatient and carrier claims for multiple comorbid chronic diseases occurring 1–12 months prior to diagnosis and computed a comorbidity score using the Klabunde adaption of the Charlson comorbidity index score (29). We further categorized comorbidity as a score of zero, one, or two or more. Neighborhood socioeconomic condition was measured by the census-tract-poverty rate, defined by the percentage of the population living in poverty at the census-tract level of the patient's residence at the time of diagnosis and was derived from the 2000 US Census. Characteristics of the hospital where the colorectal surgery was performed included hospital volume (number of hospital beds), and surgeon volume (case load per surgeon during the study period). Data regarding hospital characteristics were obtained from the Healthcare Cost Report and the Provider of Service files from CMS.

Statistical analyses

Univariable analyses were performed to assess the association between 30-day mortality following CRC surgery and each prognostic factor. The outcome included in-hospital and post-discharge 30-day mortality regardless of cause. Vital status at 30 days after definitive CRC surgery was based on Medicare data, because SEER data only includes the month and year of death. In all analyses, we used multilevel logistic regression to account for nesting of CRC patients within hospitals, using a two-level model with a random intercept. We developed a model by adding several blocks of prognostic factors to a base model. The base model included only tumor-related predictors since those are typically used by clinicians, and we subsequently added treatment, sociodemographic, hospital, and neighborhood characteristics to examine the improvement in model performance. The final model included (blocks of) predictors based on model improvement as assessed by log-likelihood-ratio tests. Additionally, we fitted a model using predictors derived from the revised ACPGIBI model. This model included age, comorbidity score, AJCC stage, mode of presentation and type of surgery. We added race/ethnicity, hospital, and neighborhood characteristics to this ACPGIBI-derived model to determine their influence on model performance. Adjusted odds ratios (ORs) were calculated with 95% confidence intervals (CIs). We decided not to include any interactions among covariates based on the purpose of our study, the potential difficulty

in implementing the model in clinical care that included interactions, and the sparse data in some of the cells, the latter reflecting the structure whereby patients are clustered within hospitals. Statistical analyses were performed using Stata (version 12.0, Stata Corp, College Station, Texas, USA).

Model performance

The performance of the models was evaluated in terms of calibration and discrimination (30). Calibration refers to the agreement between observed outcomes and expected probabilities (31) and is highly relevant when trying to predict the expected mortality for a group of patients (32). Calibration of the models was tested using the Hosmer-Lemeshow χ^2 goodness-of-fit test (33). We ranked patients into 10 equal groups of ascending mortality risk based on the predicted probability of 30-day mortality in the models and statistically tested observed and expected number of deaths in each tenth. Discrimination refers to the ability of a model to distinguish between patients with and without the outcome (31), and was assessed by the area under the receiver operating characteristic curve (AUC) (34). Models with an AUC between 0.7 and 0.8 are considered to provide moderate discrimination, whereas models with an AUC exceeding 0.8 are considered to provide good-to-excellent discrimination (32).

We examined the influence of the blocks of predictors on model performance, specifically assessing the (incremental) effect of sociodemographic, hospital characteristics, and census-tract poverty rate, and compared the performance of our SEER-Medicare model to the ACPGIBI-derived model. Additionally, sensitivity analyses were performed to assess stability of our SEER-Medicare model across several subgroups. Prognostic model performance measures were calculated separately for colon and rectal cancer patients and for elective and emergency procedures as part of sensitivity analyses.

Results

In our SEER-Medicare cohort, patients were on average 78 years of age, the majority of patients were female, and over 85% of patients were white (Table 1). Over half of patients had at least one comorbid condition, with 27% having a comorbidity score of two or more. Sixteen percent of patients resided in census tracts with a poverty rate over 20%. Most patients had colon cancer, with only about 18% rectal cancer, and over half of patients were diagnosed with stage II or III cancer. The vast majority of patients underwent partial or subtotal colectomy, and 24% underwent emergency surgery. Almost half of the patients were treated in a hospital with 350 beds or more. Of the 44,941 patients in our cohort, 2,658 (5.9%) died within 30 days of definitive surgery; over half of these patients died from CRC, and most died during their hospital stay. Furthermore, all included predictors were associated with 30-day mortality in univariable analysis.

Multivariable multilevel models

The base model containing only tumor-related predictors showed poor discrimination (AUC=0.69) and calibration was less than adequate ($\chi^2=58.59$, 8 df $p<0.001$). Next, we added blocks of predictors (treatment, sociodemographic, neighborhood, and hospital

characteristics) to this base model, which resulted in a full model containing all variables based on improvements of model fit assessed by log-likelihood-ratio tests. All predictors in our final multivariable multilevel regression model were statistically associated with 30-day mortality following CRC surgery, with strongest predictors including age at diagnosis, AJCC stage, and mode of presentation (Table 2). Adding treatment and sociodemographic characteristics to the model resulted in improvements in model discrimination, either when added separately (AUC=0.72 and AUC=0.74, respectively) or when combined (AUC=0.76). In these models, calibration improved, but was less than ideal (i.e., $p < 0.05$). Furthermore, when subsequently adding hospital and neighborhood characteristics to the model, improvements in calibration were observed, but discrimination of the model did not change (Table 2). No differences in model performance measures were observed when excluding race/ethnicity from the sociodemographic factors in both the base and full model (data not shown). The most optimal combination of performance measures was observed for the model including all predictors (AUC=0.76 and $\chi^2=28.83$, 8 df $p < 0.001$) (Table 2). Although a statistically significant p-value was observed for the full model when using the Hosmer-Lemeshow test, the calibration plot indicated that the model seemed to assign the correct probability of the outcome to patients across the risk deciles (Figure 1), and overall observed and expected 30-day mortality ratio did not differ substantially (5.9 vs. 5.8).

In addition, we fitted a multivariable multilevel prognostic model using predictors derived from the revised ACPGBI model (data not shown). All predictors included in the model (i.e., age at diagnosis, comorbidity, AJCC stage, mode of presentation and type of surgery) were significantly associated with an increased 30-day mortality risk following CRC surgery. Discrimination of the ACPGBI-derived model (AUC=0.76) was comparable to our SEER-Medicare model (AUC=0.76), but calibration of the former model was less than optimal ($\chi^2=53.32$, 8 df $p < 0.001$; data not shown). Adding race/ethnicity, census-tract poverty rate, or hospital characteristics to the ACPGBI-derived model resulted in slight improvements of calibration, but no changes in discrimination were observed (data not shown).

Finally, we performed sensitivity analyses by applying our SEER-Medicare model to subgroups of patients (i.e., to colon and rectal patients and to elective and emergency procedures). The model showed moderate discrimination among the different subgroups (Table 3). Among colon cancer patients, the SEER-Medicare model achieved an AUC of 0.76, the same value of the AUC was observed for rectal cancer patients. For elective and emergency procedures, AUC values of 0.75 and 0.72 were observed. In terms of calibration, the Hosmer-Lemeshow test indicated that our model was only well-calibrated when applied to rectal cancer patients; however, observed and expected mortality rates did not differ appreciably in the other subgroups (Table 3). Because of the sparseness of some of the data in our multilevel approach, the stage-specific multilevel models did not converge. This sparseness also includes the lack of data regarding microsatellite instability status. However, discrimination was similar for the stage-specific models in single-level analysis. These single-level models were well calibrated except for stage 0/I where $p=0.0325$ (data not shown). We also examined two interactions (stage-histology and stage-tumor site), but they did not improve model performance and were therefore not included.

Discussion

The aim of this study was to improve upon the existing models to predict 30-day mortality following CRC surgery using SEER-Medicare data, and to assess the influence of race/ethnicity, neighborhood and hospital characteristics on model performance. Both our SEER-Medicare model and the ACPGIBI-derived model performed equally well in our cohort but, perhaps surprisingly, the incremental predictive ability of the models was only minimally influenced by the additions of race/ethnicity, neighborhood and hospital characteristics.

Previously developed models aimed at predicting 30-day mortality following CRC surgery (CR-POSSUM, ACPGIBI and CCF-CCM) included a variety of prognostic variables (9–11), but all of these models omitted race/ethnicity, neighborhood, and hospital characteristics. Race/ethnicity, census-tract poverty rate, and hospital characteristics were associated with 30-day mortality risk, but SEER-Medicare overall model performance did not improve substantially. Treatment and (other) sociodemographic characteristics, however, did improve the discrimination of the base model that included only tumor-related predictors.

Additionally, when we added race/ethnicity, neighborhood or hospital characteristics to the ACPGIBI-derived model, no improvement was observed either in terms of discrimination or calibration. Although improvements in the AUC depend on the AUC of the base model, our SEER-Medicare base model appears to have sufficient opportunity for improvement since it had poor to moderate discrimination (AUC=0.69). Because the effect sizes of race/ethnicity, neighborhood and hospital characteristics were not very large and larger effects may raise the AUC more than predictors with only moderate or weak effects (35), this may indicate why we did not observe large improvements in the AUC when adding these predictors to the SEER-Medicare and ACPGIBI-derived models. Furthermore, adding an important risk factor to a model does often not result in improvements in both discrimination and calibration of the model (36). Although model discrimination did not markedly improve when adding race/ethnicity, neighborhood and hospital characteristics to the model, slight improvements in calibration were observed.

Compared to previously developed prognostic models to predict 30-day mortality following CRC surgery, our SEER-Medicare model (AUC=0.76) achieved discrimination comparable to the ACPGIBI model (AUC=0.78), but the CR-POSSUM and CCF-CCM models yielded more favorable discrimination (AUC=0.90 and AUC=0.80, respectively). (9–11) Upon external validation, however, both CR-POSSUM and ACPGIBI underestimated mortality, and showed poor discrimination (16). Furthermore, although the CCF-CCM model exhibited good discriminative performance, it lacked calibration when assessed in CRC patients from a single university clinic (37). Our model showed fair calibration using the Hosmer-Lemeshow goodness-of-fit test. Although the SEER-Medicare model seemed to assign the correct probability of the outcome to patients across risk deciles based on the calibration plot, p-values <0.05 were observed, which could be explained by the large sample size of our study (38, 39). In sensitivity analyses, our SEER-Medicare model performed equally well in colon and rectal cancer patients, but less favorably for emergency than elective procedures. The observed and expected mortality rates did not differ appreciably, indicating again that overall, the model seemed to assign the correct outcome to patients in our study sample.

Use of the SEER-Medicare registry was a strength of this study, as the registry includes a large sample size and adequate population-based representation of cancer cases in the US (31). These data facilitated development of the current prognostic model in a diverse patient population and across different hospitals/regions in the US. By contrast, both the CR-POSSUM and ACPGBI models were developed on UK patients, which limit their utility in a US population by virtue of the intrinsic differences between US and UK health care systems (i.e., higher percentage of elderly patients, greater frequency of emergency surgeries and greater proportion of patients with more advanced cancer in the UK than in the US) (13).

Among the limitations of our study, we recognize that the linked SEER-Medicare data do not include information on performance status (e.g., American Society of Anesthesiologists [ASA] grade), an important factor for medical decision-making and predicting survival (31). Although ASA grade was a relevant variable in the ACPGBI model, information on this predictor was missing for 40% of the patients (10), limiting its utility. Additionally, both the Charlson comorbidity index and ASA grade are important measures to assess comorbidity and comorbidity is an independent risk factor for adverse outcome following CRC surgery, irrespective of which measure was used (40). Also, microsatellite instability was not available. Although important, smaller or rural hospitals may not be able to test for MSI or other markers, which would reduce the utility of developing models with missing data for many patients. Furthermore, SEER-Medicare data lack information on physiological measures, which may increase performance of future models. Although the ACPGBI model achieved similar performance as the CR-POSSUM model without incorporating physiological predictors, the CCF-CCM included hematocrit as a physiological measure to predict 30-day mortality after CRC surgery. It is possible that adding other clinical and physiological prognostic factors (including patient comorbidities) could enhance the performance of our model, especially since both our SEER-Medicare model including all prognostic factors and the ACPGBI-derived model achieved only moderate discrimination in our SEER-Medicare cohort. However, key physiological predictors may not be available for all patients, especially those presenting in an emergency setting. Thus, additional efforts may be required to obtain these measures from other data sources if not available in administrative claims data (38, 41). In addition, predictors in prognostic models should reflect the prevailing circumstances at the time of admission, rather than during or after the hospital stay (42). It is worth noting that our model included only those predictors that reflected the status and environment of patients at the time of admission and surgery.

In conclusion, our SEER-Medicare model achieved moderate discrimination and fair calibration in our population of CRC patients. Race/ethnicity, neighborhood, and hospital characteristics did not appreciably increase model performance. Since prognostic models are increasingly used to compare hospital quality and increase transparency of care, using administrative claims data may be especially useful with regard to the availability of the data. Either our prognostic model, or a model similar to the ACPGBI, could eventually aid in quality improvements for CRC patients, but further improvement in model performance and external validation is needed.

Acknowledgments

This work was supported by grants from the National Cancer Institute at the National Institutes of Health (grant number CA112159); and the Health Behavior, Communication and Outreach Core; the Core is supported in part by the National Cancer Institute Cancer Center Support Grant (grant number P30 CA91842) to the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital in St. Louis, Missouri. Dr. Davidson was supported in part through grants HL-38180, DK-56260, and Digestive Disease Research Core Center DK-52574. We gratefully acknowledge James Struthers for his data management and programming services. We thank the Alvin J. Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine in St. Louis, Missouri, for the use of the Health Behavior, Communication, and Outreach Core. 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.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA: a cancer journal for clinicians*. 2013; 63:11–30. [PubMed: 23335087]
2. Billingsley KG, Morris AM, Dominitz JA, et al. Surgeon and hospital characteristics as predictors of major adverse outcomes following colon cancer surgery: understanding the volume-outcome relationship. *Arch Surg*. 2007; 142:23–31. [PubMed: 17224497]
3. Iversen LH, Harling H, Laurberg S, Wille-Jørgensen P. Influence of caseload and surgical speciality on outcome following surgery for colorectal cancer: a review of evidence. Part 1: short-term outcome. *Colorectal Dis*. 2007; 9:28–37. [PubMed: 17181843]
4. Rosen L, Stasik JJ Jr, Reed JF 3rd, Olenwine JA, Aronoff JS, Sherman D. Variations in colon and rectal surgical mortality. Comparison of specialties with a state-legislated database. *Dis Colon Rectum*. 1996; 39:129–135. [PubMed: 8620777]
5. Sagar PM, Hartley MN, MacFie J, Taylor BA, Copeland GP. Comparison of individual surgeon's performance. Risk-adjusted analysis with POSSUM scoring system. *Dis Colon Rectum*. 1996; 39:654–658. [PubMed: 8646952]
6. Sagar PM, Hartley MN, Mancey-Jones B, Sedman PC, May J, Macfie J. Comparative audit of colorectal resection with the POSSUM scoring system. *Br J Surg*. 1994; 81:1492–1494. [PubMed: 7820482]
7. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg*. 1991; 78:355–360. [PubMed: 2021856]
8. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. *Br J Surg*. 1998; 85:1217–1220. [PubMed: 9752863]
9. Tekkis PP, Prytherch DR, Kocher HM, et al. Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). *Br J Surg*. 2004; 91:1174–1182. [PubMed: 15449270]
10. Tekkis PP, Poloniecki JD, Thompson MR, Stamatakis JD. Operative mortality in colorectal cancer: prospective national study. *BMJ*. 2003; 327:1196–1201. [PubMed: 14630754]
11. Fazio VW, Tekkis PP, Remzi F, Lavery IC. Assessment of operative risk in colorectal cancer surgery: the Cleveland Clinic Foundation colorectal cancer model. *Dis Colon Rectum*. 2004; 47:2015–2024. [PubMed: 15657649]
12. Richards CH, Leitch FE, Horgan PG, McMillan DC. A systematic review of POSSUM and its related models as predictors of post-operative mortality and morbidity in patients undergoing surgery for colorectal cancer. *J Gastrointest Surg*. 2010; 14:1511–1520. [PubMed: 20824372]
13. Senagore AJ, Warmuth AJ, Delaney CP, Tekkis PP, Fazio VW. POSSUM, p-POSSUM, and Cr-POSSUM: implementation issues in a United States health care system for prediction of outcome for colon cancer resection. *Dis Colon Rectum*. 2004; 47:1435–1441. [PubMed: 15486738]
14. Al-Homoud S, Purkayastha S, Aziz O, et al. Evaluating operative risk in colorectal cancer surgery: ASA and POSSUM-based predictive models. *Surg Oncol*. 2004; 13:83–92. [PubMed: 15572090]

15. Teeuwen PH, Bremers AJ, Groenewoud JM, van Laarhoven CJ, Bleichrodt RP. Predictive value of POSSUM and ACPGBI scoring in mortality and morbidity of colorectal resection: a case-control study. *J Gastrointest Surg.* 2011; 15:294–303. [PubMed: 20936370]
16. Ferjani AM, Griffin D, Stallard N, Wong LS. A newly devised scoring system for prediction of mortality in patients with colorectal cancer: a prospective study. *Lancet Oncol.* 2007; 8:317–322. [PubMed: 17395105]
17. Fichera A. How accurate is the Cleveland Clinic Foundation model in predicting operative risk in colorectal cancer patients? *Nat Clin Pract Oncol.* 2005; 2:258–259. [PubMed: 16075798]
18. Schneider EB, Haider AH, Hyder O, Efron JE, Lidor AO, Pawlik TM. Assessing short- and long-term outcomes among black vs white Medicare patients undergoing resection of colorectal cancer. *Am J Surg.* 2013
19. Bharathan B, Welfare M, Borowski DW, et al. Impact of deprivation on short- and long-term outcomes after colorectal cancer surgery. *Br J Surg.* 2011; 98:854–865. [PubMed: 21452190]
20. Schootman M, Lian M, Pruitt SL, et al. (under review) Variability in 30-day mortality following colorectal cancer surgery. *Health Services Research.*
21. Smith JJ, Tekkis PP. ACPGBI Colorectal Cancer Model. 2010
22. Alvarez C, Clark C, Zhang S, et al. Predicting out of intensive care unit cardiopulmonary arrest or death using electronic medical record data. *BMC Medical Informatics and Decision Making.* 2013; 13:28. [PubMed: 23442316]
23. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care.* 2002; 40 IV-3-18.
24. Cooper GS, Virnig B, Klabunde CN, Schussler N, Freeman J, Warren JL. Use of SEER-Medicare data for measuring cancer surgery. *Medical Care.* 2002; 40 IV-43-48.
25. Archampong D, Borowski D, Wille-Jorgensen P, Iversen LH. Workload and surgeon's specialty for outcome after colorectal cancer surgery. *The Cochrane database of systematic reviews.* 2012; 3:Cd005391. [PubMed: 22419309]
26. Chien LC, Schootman M, Pruitt SL. The modifying effect of patient location on stage-specific survival following colorectal cancer using geosurvival models. *Cancer Causes Control.* 2013; 24:473–484. [PubMed: 23306551]
27. Kolfshoten NE, Marang van de Mheen PJ, Gooiker GA, et al. Variation in case-mix between hospitals treating colorectal cancer patients in the Netherlands. *Eur J Surg Oncol.* 2011; 37:956–963. [PubMed: 21944049]
28. Morris EJ, Taylor EF, Thomas JD, et al. Thirty-day postoperative mortality after colorectal cancer surgery in England. *Gut.* 2011; 60:806–813. [PubMed: 21486939]
29. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol.* 2000; 53:1258–1267. [PubMed: 11146273]
30. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ.* 2009; 338:b604. [PubMed: 19336487]
31. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer. 2009
32. Steyerberg EW, Neville BA, Koppert LB, et al. Surgical mortality in patients with esophageal cancer: development and validation of a simple risk score. *J Clin Oncol.* 2006; 24:4277–4284. [PubMed: 16963730]
33. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med.* 1997; 16:965–980. [PubMed: 9160492]
34. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996; 15:361–387. [PubMed: 8668867]
35. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *American journal of epidemiology.* 2012; 176:473–481. [PubMed: 22875755]
36. Smulders YM, Thijs A, Twisk JW. New cardiovascular risk determinants do exist and are clinically useful. *European Heart Journal.* 2008; 29:436–440. [PubMed: 18276616]

37. Dogrul AB, Kilic YA, Celebi AE, et al. External validation of Cleveland Clinic Foundation colorectal cancer model in a University Clinic in terms of predicting operative mortality. *Tech Coloproctol*. 2010; 14:9–12. [PubMed: 19953287]
38. Tabak YP, Sun X, Johannes RS, Hyde L, Shorr AF, Lindenauer PK. Development and Validation of a Mortality Risk-Adjustment Model for Patients Hospitalized for Exacerbations of Chronic Obstructive Pulmonary Disease. *Med Care*. 2013
39. Hosmer DW, Lemeshow S. Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics: Theory and Methods*. 1980; 9:1043–1069.
40. Dekker JW, Gooiker GA, van der Geest LG, et al. Use of different comorbidity scores for risk-adjustment in the evaluation of quality of colorectal cancer surgery: does it matter? *Eur J Surg Oncol*. 2012; 38:1071–1078. [PubMed: 22703758]
41. Bann SD, Sarin S. Comparative audit: the trouble with POSSUM. *J R Soc Med*. 2001; 94:632–634. [PubMed: 11733590]
42. Ash AS, Fienberg SE, Louis TA, Normand SL, Stukel TA, Utts J. Statistical issues on assessing hospital performance. The COPSS-CMS White Paper Committee. 2011

Figure 1.
Calibration plot of predicted 30-day mortality following CRC surgery by SEER-Medicare model

Descriptive characteristics and univariable associations between prognostic factors and 30-day mortality following colorectal cancer surgery among n=44,941 patients

Table 1

Predictor	Category	Number of patients (%)	Number of deaths (%)	OR*	95% CI
<i>Tumor characteristics</i>					
<i>AJCC stage</i>	0/I	10,700 (23.8)	316 (3.0)	Ref.	--
	II	14,662 (32.6)	844 (5.8)	2.00	1.76–2.29
	III	12,407 (27.6)	655 (5.3)	1.83	1.59–2.09
	IV	5,539 (12.4)	726 (13.1)	4.95	4.31–5.68
	Unknown	1,633 (3.6)	117 (7.2)	2.52	2.02–3.15
<i>Tumor grade/differentiation</i>					
	Well	3,769 (8.4)	180 (4.8)	Ref.	--
	Moderate	29,348 (65.3)	1,544 (5.3)	1.13	0.96–1.33
	Poor	8,928 (19.9)	726 (8.1)	1.85	1.56–2.19
	Undifferentiated	493 (1.1)	55 (11.2)	2.52	1.82–3.48
	Unknown	2,403 (5.3)	153 (6.4)	1.37	1.09–1.71
<i>Tumor location</i>					
	Rectal	8,426 (18.7)	357 (4.2)	Ref.	--
	Proximal colon	17,824 (39.7)	1,054 (5.9)	1.42	1.26–1.61
	Transverse colon	6,776 (15.1)	502 (7.4)	1.82	1.58–2.09
	Distal colon	11,915 (26.5)	745 (6.3)	1.50	1.32–1.71
<i>Tumor histology</i>					
	Mucinous adenocarcinoma	38,454 (85.6)	2,195 (5.7)	Ref.	--
	Other adenocarcinoma	5,779 (12.8)	381 (6.6)	1.17	1.04–1.31
	Non-adenocarcinoma	708 (1.6)	82 (11.6)	2.16	1.71–2.74
<i>Treatment characteristics</i>					
<i>Mode of presentation</i>					
	Elective	34,133 (75.9)	1,419 (4.2)	Ref.	--
	Emergency	10,808 (24.1)	1,239 (11.5)	2.98	2.75–3.23
<i>Type of surgery</i>					
	Local tumor excision	951 (2.1)	49 (5.2)	Ref.	--
	Partial colectomy	18,594 (41.4)	1,041 (5.6)	1.09	0.81–1.47

Predictor	Category	Number of patients (%)	Number of deaths (%)	OR*	95% CI
	Subtotal (hemi)colectomy	22,496 (50.1)	1,358 (6.0)	1.18	0.88–1.59
	Total (procto)colectomy	1,775 (3.9)	105 (5.9)	1.14	0.81–1.62
	Colectomy NOS	941 (2.1)	90 (9.6)	1.94	1.35–2.80
	Other surgery	184 (0.4)	15 (8.2)	1.66	0.91–3.04
Sociodemographics					
Age, years	66–74	16,400 (36.5)	562 (3.4)	Ref.	--
	75–84	20,398 (45.4)	1,132 (5.6)	1.67	1.50–1.85
	85+	8,143 (18.1)	964 (11.8)	3.84	3.45–4.28
Sex	Male	19,288 (42.9)	1,190 (6.2)	Ref.	--
	Female	25,653 (57.1)	1,468 (5.7)	0.92	0.85–1.00
Race/Ethnicity	White	38,613 (85.9)	2,313 (6.0)	Ref.	--
	African American	3,374 (7.5)	217 (6.4)	1.06	0.91–1.23
	Other	2,954 (6.6)	128 (4.3)	0.72	0.60–0.87
Comorbidity	0	19,501 (43.4)	906 (4.7)	Ref.	--
	1	13,138 (29.2)	767 (5.8)	1.26	1.14–1.39
	2+	12,302 (27.4)	985 (8.0)	1.77	1.61–1.94
Hospital characteristics					
Hospital volume (no. of beds)	1–199	11,326 (25.2)	759 (6.7)	Ref.	--
	200–349	12,616 (28.1)	764 (6.1)	0.90	0.80–1.02
	350–499	10,852 (24.1)	605 (5.6)	0.82	0.72–0.94
	500+	10,147 (22.6)	530 (5.2)	0.74	0.64–0.85
Hospital surgeon volume	1–20	10,420 (23.2)	720 (6.9)	Ref.	--
	21–38	9,935 (22.1)	611 (6.2)	0.87	0.77–0.97
	39–65	9,762 (21.7)	573 (5.9)	0.81	0.72–0.91
	66+	9,994 (22.2)	442 (4.4)	0.61	0.54–0.70
	Unknown	4,830 (10.8)	312 (6.5)	0.93	0.80–1.06

Predictor	Category	Number of patients (%)	Number of deaths (%)	OR*	95% CI
<i>Neighborhood characteristic</i>					
<i>Poverty rate</i>	<10%	26,232 (58.4)	1,394 (5.3)	Ref.	--
	10–19%	11,473 (25.5)	745 (6.5)	1.23	1.12–1.35
	20%	7,236 (16.1)	519 (7.2)	1.36	1.22–1.52

* Univariable odds ratios are adjusted for clustering of patients within hospitals

Table 2

Multivariable multilevel logistic regression analyses to predict 30-day mortality in patients undergoing CRC surgery

Predictor	Category	Model 1 – Base			Model 2			Model 3			Model 4			Model 5 – Final		
		OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI	
<i>Tumor characteristics</i>																
<i>AJCC stage</i>	0/I	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	
	II	1.92	1.68–2.20	1.72	1.50–1.97	1.62	1.41–1.86	1.61	1.40–1.85	1.61	1.40–1.85	1.61	1.40–1.85	1.61	1.40–1.84	
	III	1.72	1.49–1.98	1.49	1.29–1.72	1.52	1.31–1.76	1.51	1.30–1.75	1.50	1.30–1.75	1.50	1.30–1.74	1.50	1.30–1.74	
	IV	4.50	3.90–5.18	3.70	3.20–4.28	4.15	3.59–4.81	4.12	3.56–4.77	4.10	3.56–4.77	4.10	3.54–4.75	4.10	3.54–4.75	
	Unknown	2.06	1.62–2.62	1.80	1.41–2.29	1.81	1.41–2.31	1.78	1.39–2.28	1.77	1.39–2.28	1.77	1.38–2.26	1.77	1.38–2.26	
<i>Tumor grade/differentiation</i>	Well	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	
	Moderate	0.94	0.80–1.10	0.94	0.80–1.11	0.94	0.79–1.11	0.94	0.79–1.11	0.93	0.79–1.11	0.93	0.79–1.10	0.93	0.79–1.10	
	Poor	1.28	1.08–1.53	1.27	1.06–1.52	1.26	1.05–1.51	1.27	1.06–1.53	1.27	1.06–1.53	1.27	1.06–1.52	1.27	1.06–1.52	
	Undifferentiated	1.59	1.13–2.22	1.61	1.15–2.26	1.62	1.15–2.29	1.63	1.16–2.31	1.63	1.16–2.31	1.63	1.15–2.30	1.63	1.15–2.30	
	Unknown	1.16	0.92–1.47	1.14	0.90–1.45	1.15	0.90–1.46	1.14	0.89–1.45	1.13	0.89–1.44	1.13	0.89–1.44	1.13	0.89–1.44	
<i>Tumor location</i>	Rectal	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	
	Proximal colon	1.31	1.15–1.48	1.28	1.10–1.49	1.14	0.97–1.33	1.11	0.95–1.30	1.12	0.95–1.30	1.12	0.96–1.31	1.12	0.96–1.31	
	Transverse colon	1.66	1.44–1.91	1.59	1.35–1.86	1.42	1.21–1.67	1.39	1.18–1.63	1.40	1.18–1.63	1.40	1.19–1.65	1.40	1.19–1.65	
	Distal colon	1.44	1.26–1.64	1.35	1.17–1.55	1.30	1.13–1.50	1.26	1.10–1.46	1.27	1.10–1.46	1.27	1.10–1.46	1.27	1.10–1.46	
<i>Tumor histology</i>	Mucinous adenocarcinoma	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	
	Other adenocarcinoma	1.02	0.91–1.15	1.04	0.93–1.17	1.01	0.90–1.14	1.01	0.90–1.14	1.01	0.90–1.14	1.01	0.90–1.14	1.01	0.90–1.14	
	Non-adenocarcinoma	1.63	1.24–2.15	1.56	1.18–2.07	1.61	1.21–2.14	1.63	1.22–2.16	1.64	1.22–2.16	1.64	1.24–2.18	1.64	1.24–2.18	
<i>Treatment characteristics</i>																
<i>Mode of presentation</i>	Elective	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	
	Emergency	2.64	2.43–2.86	2.39	2.20–2.60	2.35	2.20–2.60	2.35	2.16–2.56	2.34	2.16–2.56	2.34	2.16–2.55	2.34	2.16–2.55	
<i>Type of surgery</i>	Local tumor excision	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	Ref.	Ref.	--	

Predictor	Category	Model 1 – Base		Model 2		Model 3		Model 4		Model 5 - Final	
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
	Partial colectomy			1.15	0.84–1.57	1.37	1.00–1.88	1.40	1.01–1.94	1.38	1.00–1.92
	Subtotal (hemi)colectomy			1.05	0.76–1.44	1.27	0.92–1.75	1.30	0.94–1.82	1.29	0.92–1.79
	Total (procto)colectomy			1.41	0.98–2.04	1.72	1.18–2.50	1.75	1.20–2.57	1.73	1.18–2.54
	Colectomy NOS			1.48	1.01–2.16	1.87	1.27–2.75	1.92	1.29–2.84	1.89	1.27–2.81
	Other surgery			1.28	0.69–2.39	1.53	0.81–2.88	1.55	0.82–2.93	1.56	0.83–2.95
Sociodemographics											
Age, years	66–74					Ref.	--	Ref.	--	Ref.	--
	75–84					1.69	1.52–1.88	1.69	1.52–1.88	1.70	1.53–1.90
	85+					3.75	3.34–4.20	3.75	3.35–4.21	3.79	3.38–4.25
Sex	Male					Ref.	--	Ref.	--	Ref.	--
	Female					0.79	0.73–0.86	0.79	0.73–0.86	0.79	0.72–0.86
Race/Ethnicity	White					Ref.	--	Ref.	--	Ref.	--
	African American					1.02	0.87–1.19	1.03	0.88–1.21	0.93	0.79–1.09
	Other					0.72	0.59–0.87	0.70	0.57–0.84	0.66	0.55–0.81
Comorbidity	0					Ref.	--	Ref.	--	Ref.	--
	1					1.21	1.09–1.34	1.21	1.09–1.34	1.20	1.08–1.33
	2+					1.65	1.49–1.82	1.65	1.49–1.82	1.63	1.48–1.80
Hospital characteristics											
Hospital volume (no. of beds)	1–199							Ref.	--	Ref.	--
	200–349							0.86	0.76–0.98	0.88	0.78–1.00
	350–499							0.80	0.70–0.92	0.83	0.72–0.95
	500+							0.75	0.65–0.87	0.78	0.67–0.90
Hospital surgeon volume	1–20							Ref.	--	Ref.	--
	21–38							0.91	0.81–1.02	0.92	0.81–1.03
	39–65							0.88	0.77–0.99	0.88	0.78–1.00

Predictor	Category	Model 1 – Base		Model 2		Model 3		Model 4		Model 5 - Final	
		OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
	66+							0.75	0.66–0.86	0.76	0.67–0.87
	Unknown							0.94	0.81–1.09	0.94	0.81–1.09
<i>Neighborhood characteristic</i>											
<i>Poverty rate</i>	<10%									Ref.	--
	10–19%									1.24	1.12–1.36
	20%									1.31	1.16–1.47
Area under ROC			0.687		0.721		0.763		0.762		0.763
Hosmer-Lemeshow χ^2 , p-value *		55.04	0.000	47.58	0.000	41.27	0.000	36.58	0.000	28.83	0.001

Abbreviations: OR=odds ratio; 95% CI=95% confidence interval; Ref=reference category; ROC: receiver operating curve

* P-value based on χ^2 value using 8 degrees of freedom

Table 3

Performance of SEER-Medicare model among different subgroups of patients

Subgroups	Number of patients	Number of deaths	Observed % 30-day mortality	Expected % 30-day mortality	AUC
All patients	44,941	2,658	5.9	5.8 (p<0.001) *	0.76
Elective procedures	34,133	1,419	4.2	4.0 (p<0.000)	0.75
Emergency procedures	10,808	1,239	11.5	11.2 (p=0.013)	0.72
Colon cancer	36,515	2,301	6.3	6.2 (p<0.001)	0.76
Rectal cancer	8,426	357	4.2	4.2 (p=0.763)	0.76

* P-values obtained from Hosmer-Lemeshow test for comparison of predicted and observed mortality