



HHS Public Access

Author manuscript

Med Care. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:

Med Care. 2019 April ; 57(4): 286–294. doi:10.1097/MLR.0000000000001073.

Comparison of Medicare claims-based proxy measures of poor function and associations with treatment receipt and mortality in older colon cancer patients.

Sophie E. Mayer, MS,

Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

Hung-Jui Tan, MD, MSHPM,

Department of Urology, School of Medicine, University of North Carolina at Chapel Hill

Sharon Peacock Hinton, MPA,

Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

Hanna K. Sanoff, MD, MPH,

Division of Hematology/Oncology, School of Medicine & Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill

Til Stürmer, MD, PhD,

Department of Epidemiology Gillings School of Global Public Health, University of North Carolina at Chapel Hill

Laura L. Hester, PhD, ScM,

Janssen Research and Development, Titusville, NJ

Keturah R. Faurot, PhD,

Sophie E. Mayer, MS, 135 Dauer Drive, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599-7435, 773-860-7207, Hung-Jui Tan, MD, MSHPM, 170 Manning Drive, CB #7235, Chapel Hill, NC 27599, 919-966-2571, 919-966-0098, ray_tan@med.unc.edu. Sharon Peacock Hinton, MPA, 101 E Weaver St, 2nd Floor, Carrboro, NC 27610, (919) 962-5371, sharon_peacock@med.unc.edu. Hanna K. Sanoff, MD, MPH, Division of Hematology/Oncology, 170 Manning Drive, CB #7305, Chapel Hill, NC 27599, (919) 966-4431, hanna_sanoff@med.unc.edu. Til Stürmer, MD, PhD, 135 Dauer Drive, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599-7435, (919) 966-7433, (919) 966-2089, sturmer@unc.edu. Laura L. Hester, PhD, ScM, 1125 Trenton-Harbourton Rd, Titusville, NJ, 08560, (609) 730-6927, lhester@its.jnj.com. Keturah R. Faurot, PhD, 336 Emergency Room Dr. Wing D-191, CB #7200, Chapel Hill, NC 27599, (919) 966-8586, (919) 843-5452, faurot@med.unc.edu. Michele Jonsson Funk, PhD, 104B Market Street, Campus Box 7521, Chapel Hill, NC 27599, (919) 843-0384, (919) 843-7364, mfunk@unc.edu. Jennifer L. Lund, PhD, 135 Dauer Drive, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599-7435, (919) 966-7440, jennifer.lund@unc.edu. **Corresponding author:** Sophie E. Mayer, 135 Dauer Drive, 2101 McGavran-Greenberg Hall, CB #7435, Chapel Hill, NC 27599-7435, Phone: 773-860-7207, Fax: 919-966-4914, semayer@unc.edu.

Potential conflicts of interest:

- SEM, HJT, SPH, LLH, KRF, JLL: No potential conflicts of interest.
- MJF receives research funding via UNC from AstraZeneca, and salary support from the Center for Pharmacoepidemiology in the Department of Epidemiology, Gillings School of Global Public Health (current members: GlaxoSmithKline, UCB BioSciences, Merck). MJF is a member of the Scientific Steering Committee (SSC) for a post-approval safety study of an unrelated drug class funded by GSK. All compensation for services provided on the SSC is invoiced by and paid to UNC.
- TS receives salary support as Director of the Comparative Effectiveness Research Strategic Initiative, Clinical & Translational Science Award (UL1TR001111) and as Director of the Center for Pharmacoepidemiology and research support from pharmaceutical companies (Amgen, AstraZeneca) to the Department of Epidemiology, Gillings School of Global Public Health at UNC. He owns stock in Novartis, Roche, BASF, AstraZeneca, and Novo Nordisk.
- HKS received research funding paid to UNC from Bayer, Merck, and Precision Biologics.

Department of Physical Medicine and Rehabilitation, School of Medicine, University of North Carolina at Chapel Hill

Michele Jonsson Funk, PhD, and

Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

Jennifer L. Lund, PhD

Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

Abstract

Background: Multiple claims-based **proxy measures of** poor function have been developed to address confounding in observational studies of drug effects in older adults. We evaluated agreement between these measures and their associations with treatment receipt and mortality in a cohort of older colon cancer patients.

Methods: Medicare beneficiaries age 66+ diagnosed with stage II-III colon cancer were identified in the Surveillance, Epidemiology, and End Results-Medicare database (2004–2011). Poor function was operationalized by: (1) summing the **total** poor function indicators for each model and (2) estimating predicted probabilities of poor function at diagnosis. Agreement was evaluated using Fleiss' kappa and Spearman's correlation. Associations between **proxy** measures and (1) laparoscopic vs. open surgery, (2) chemotherapy vs. none, (3) 5-fluorouracil (5FU) +oxaliplatin (FOLFOX) vs. 5FU monotherapy, and (4) one-year mortality were estimated using log-binomial regression, controlling for age, sex, stage, and comorbidity. Survival estimates were stratified by functional group, age, and comorbidity.

Results: Among 29,687 eligible colon cancer patients, 67% were 75+ years and 45% had stage III disease. Concordance across the poor function indicator counts was moderate (κ : 0.64) and correlation of predicted probability measures varied (ρ : 0.21–0.74). Worse function was associated with lower chemotherapy and FOLFOX receipt, and higher one-year mortality. Within age and comorbidity strata, poor function remained associated with mortality.

Conclusions: While agreement varied across the claims-based **proxy measures**, each demonstrated anticipated associations with treatment receipt and mortality independent of comorbidity. Claims-based comparative effectiveness studies in older populations should consider applying one of these models to improve confounding control.

Keywords

Frailty; poor function; algorithm; administrative claims; Medicare; colon cancer

Introduction

Large administrative healthcare claims databases are increasingly used to describe patterns and quality of medical care and to assess the effectiveness and safety of medical products and interventions. However, these data sources do not include important covariates known to influence treatment decisions and impact clinical outcomes (e.g., body mass index, smoking

status, disease severity), potentially compromising the validity of comparative effectiveness research (CER) conducted in claims databases.¹ **In particular, frailty, disability, physical and cognitive function, and dependency are major considerations** in clinical decision-making in older adult populations, making uncontrolled confounding by **these factors** a major threat to the validity of CER studies conducted in populations with diseases of aging, such as cancer.

In the clinical literature, frailty is defined as “a state of increased vulnerability to poor resolution of homeostasis after a stressor event”² that arises due to depleted reserves and age-related declines in physiologic function.³ Stressor events in frail individuals can trigger disproportionate changes in mobility and lucidity that are difficult to capture in administrative data. **Disability, a clinical phenotype distinct from frailty, is defined as difficulty or dependency in performing daily tasks required for independent living, often due to impaired physical or cognitive function.**³ Frailty and disability are **understood to** act independently of comorbidity and other measures of health status to increase the risk of poor health outcomes.³ **Though frailty and disability represent distinct clinical constructs, herein they and the related concepts of dependency and poor function will be referred to collectively as measures of poor function.** To highlight the potential for unmeasured confounding in CER studies among older adult populations, we aimed to demonstrate the associations of several claims-based **proxy measures of poor function** with treatment receipt and mortality in a large population of older adults undergoing colon cancer treatment.

There are a number of validated measures of poor function commonly used in clinical and research settings, such as Eastern Cooperative Oncology Group (ECOG) performance status,⁴ which is used to guide treatment decisions and assess patients’ ability to care for themselves following a cancer diagnosis. In contrast to comorbidity, for which there are multiple widely-used indices that can be easily implemented using healthcare claims (e.g., Charlson Comorbidity Index,⁵ Gagne combined comorbidity score⁶), only a few studies have attempted to develop claims-based algorithms to identify and measure aspects of poor function for the purposes of risk adjustment or cohort identification. To our knowledge, there are four claims-based models or sets of indicators that have been proposed to identify proxies of poor performance status (Davidoff⁷), poor function (Chrischilles⁸), or frailty (Faurot⁹ and Segal¹⁰). **These models are designed to capture these complex constructs using a mixture of medical and psychiatric diagnoses, medical procedures, service use, disease symptoms, and mobility aids (Table 1).** These tools have never been directly compared in a single cohort **and** it is not known to what extent they capture overlapping constructs.

To understand differences between these four models, we assessed their agreement in a cohort of older adults (aged 66+ years) diagnosed with stage II or III colon cancer. We also evaluated the relationships between each claims-based **proxy** measure and the receipt of cancer treatments and mortality to assess the potential for residual confounding by **poor function** after controlling for comorbidity. Stage II-III colon cancer was chosen as a model clinical setting due to availability of more and less aggressive therapies and clearly defined treatment choices around surgery type, provision of chemotherapy, and chemotherapy

regimen that consider patient functional status. We hypothesized that **proxies of poor function** measured at cancer diagnosis would be positively associated with receipt of less invasive (i.e., laparoscopic) colon cancer surgical procedures and would be negatively associated with receipt of any chemotherapy and of more aggressive combination chemotherapy regimens for colon cancer. As these algorithms were primarily developed to assess associations between poor function and short-term mortality, we also expected that they would be strongly associated with one-year risk of death.

Methods

Data source and study population

Medicare beneficiaries aged 66 and older diagnosed with a first, primary stage II or III colon cancer between 2004–2011 were identified from the linked Surveillance, Epidemiology, and End Results (SEER) program registry and Medicare claims database (SEER-Medicare). We used the SEER cancer registry to identify patients based on tumor site and stage at diagnosis. Medicare enrollment data and Part A and B fee-for-service claims were used to assess periods of continuous enrollment, comorbid conditions, and indicators of poor function in the year prior to cancer diagnosis, as well as to identify cancer treatments following diagnosis.

To be eligible for study inclusion, all patients were required to: (1) undergo guideline-recommended surgical resection of their tumor within 90 days of their diagnosis date (set to the first day of the month of diagnosis), (2) have continuous Medicare Parts A and B coverage for at least 12 months prior to their diagnosis date and through their surgery date, and (3) have at least one claim in the year prior to diagnosis. For analyses evaluating all-cause mortality and receipt of chemotherapy as outcomes, participants were additionally required to survive 90 days from their diagnosis date and 120 days from their surgery date, respectively. This duration of follow up has been shown to ensure claims-based capture of treatment receipt with high sensitivity and specificity.¹¹ For the analysis of chemotherapy type, patients had to initiate adjuvant chemotherapy with either 5-fluorouracil alone (5FU monotherapy) or in combination with oxaliplatin (FOLFOX). The chemotherapy type analysis was restricted to patients who received no other chemotherapy drugs.

Exposure assessment

The four models developed by Davidoff et al., Chrischilles et al., Faurot et al. and Segal et al. were designed to capture **proxy measures of functional status**, Eastern Cooperative Oncology Group (ECOG) performance status, or frailty using International Classification of Diseases, Clinical Modification, 9th Edition (ICD-9) diagnosis and procedure codes and Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) procedure codes from healthcare claims. Indicators included in the four models are presented in Table 1.

Both Davidoff and Faurot developed their models using the Medicare Current Beneficiary Survey (MCBS), a nationally-representative sample of Medicare beneficiaries who completed surveys on self-reported health and functional status. MCBS surveys are linked

with Medicare Parts A and B healthcare claims. After identifying codes for services putatively associated with either poor performance status or frailty, automated model selection strategies were employed to identify indicators that best predicted survey-based measures of disability status (Davidoff) and dependency in activities of daily living (Faurot), such as eating and dressing. The Segal model was developed in the Cardiovascular Health Study using a linkage to Medicare claims **to predict Fried frailty phenotype**.¹² Conditions previously associated with frailty or classified as indicators of frailty were identified from the existing literature in consultation with geriatricians. Penalized logistic regression was used for final model selection. The Davidoff, Faurot, and Segal models generate a continuous measure of the predicted probability of poor function. In contrast, the Chrischilles model was developed using claims for Medicare beneficiaries hospitalized with acute myocardial infarction. After identifying indicators hypothesized to be associated with poor function, function-related indicators were retained in the model if they were negatively associated with cardiac catheterization during the index hospitalization and positively associated with one-year mortality at an alpha level of 0.20. The final validated model included the number of function-related indicators as a discrete variable.

In our study, indicators in the Davidoff, Chrischilles, Faurot, and Segal models were assessed for the entire eligible cohort using inpatient and outpatient healthcare claims for the 12 months prior to cancer diagnosis. For each model, patients were classified based on (1) the number of indicators associated with poor function (Table 1), and (2) tertiles of predicted probability of poor function from the Davidoff, Faurot, and Segal models.

Cut-points for the distribution of the count of Chrischilles indicators were apparent at zero (52%, low), one (28%, intermediate), or more than one (20%, high) indicators of poor function (see figure, Supplemental Digital Content 1, distribution of Chrischilles indicator counts). Thus, all patients were classified as having 0, 1, or 2+ Davidoff; 0, 1, or 2+ Faurot; and <1, 2, or 3+ Segal poor function indicators, which captured similar proportions of individuals in each category as the Chrischilles groups. The distributions of predicted probabilities of poor function were divided into thirds using the first and second tertiles (Davidoff: 0.011, 0.033; Faurot: 0.082, 0.335; Segal: 0.088, 0.188). Lastly, in a sensitivity analysis, we used the 52nd and 80th percentiles as alternative cut-points for the predicted probabilities of poor function in order to match the distribution of Chrischilles **poor** function-related indicator categories.

Outcome assessment

Several outcomes were evaluated with respect to **proxy measures of** poor function (see table, Supplemental Digital Content 2, relevant code lists to identify treatments). Receipt of laparoscopic versus open surgery was assessed for the entire cohort using ICD-9, CPT, or HCPCS procedure codes from inpatient and physician claims. Receipt of chemotherapy was defined using ICD-9 procedure and HCPCS codes and National Drug Codes (NDC) from any inpatient, outpatient, or physician claims among patients that survived 120 days after surgery. Type of chemotherapy (FOLFOX vs. 5FU monotherapy) was also evaluated using HCPCS codes and NDCs for patients who received one of these two chemotherapy

regimens. Lastly, vital status as of December 2013 was assessed using Medicare enrollment information.

Covariate assessment

Several potential confounders were identified and abstracted from the SEER registry. These included sex, race/ethnicity (Non-Hispanic White, Non-Hispanic Black/African-American, Hispanic, Other), cancer stage (II or III), and age at diagnosis (<75, 75–84, 85+). The Gagne combined comorbidity score⁶ and the Charlson comorbidity index (CCI)⁵ were computed using ICD-9 diagnosis codes from inpatient and outpatient healthcare claims in the year prior to cancer diagnosis.

Statistical analysis

Descriptive and graphical analyses were used to examine the distribution and agreement of the four **poor function proxy** measures. The distribution of poor function indicators was plotted for Chrischilles, as was the kernel density of the Davidoff, Faurot, and Segal predicted probabilities (see figure, Supplemental Digital Content 1). Fleiss' kappa¹³ was computed to assess agreement between number of poor function indicators identified in each of the models, and Spearman's rank correlation was used to compare continuous predicted probabilities. The associations between Davidoff, Chrischilles, Faurot, and Segal categories of poor function and receipt of laparoscopic versus open surgery, chemotherapy versus none, FOLFOX versus 5FU, and one-year all-cause mortality were evaluated using log-binomial models, with adjustment for all confounders described above. Log-binomial models were chosen because the outcomes in this study were common (>10%), and odds ratios will overestimate relative risks in this setting.¹⁴ The lowest category of poor function (i.e., the most robust patients) served as the referent group for all analyses. Lastly, Kaplan-Meier curves were used to examine overall survival by **level of function** within age groups (<75, 75–84, 85+) and strata of Gagne combined comorbidity score (0, 1–2, 3+).

Results

A total of 29,687 stage II-III colon cancer patients met the study inclusion criteria (Figure 1). More than half of participants were female (58%), 81% were Non-Hispanic White, and 45% had stage III disease (Table 2). Patients receiving laparoscopic surgeries, chemotherapy, and FOLFOX combination chemotherapy were younger and had lower average Gagne and Charlson comorbidity scores, and patients receiving chemotherapy and FOLFOX had more advanced (e.g., higher stage) disease.

The poor function indicator counts identified by each model had moderate agreement (Fleiss' kappa: 0.64). Rank correlation varied between the three models that generated predicted probabilities of poor function; the Davidoff and Segal models exhibited the lowest correlation (0.25) and the Faurot and Segal models exhibited the highest (0.71). The Davidoff and Faurot models exhibited moderate correlation (0.42).

Individuals with higher predicted probability of poor function (above the first and second tertiles in the Davidoff and Faurot distributions) were significantly less likely to receive laparoscopic procedures for surgical resection of their tumors (Table 3); however, the Segal

model groups were not associated with receipt of laparoscopic surgery. The number of poor function indicators was not consistently associated with receipt of laparoscopic surgery for any model (Table 3).

Receipt of chemotherapy was evaluated in 26,209 individuals who survived at least 120 days from surgery. As expected, having a greater predicted probability of poor function was negatively associated with receipt of any chemotherapy, as was having a greater number of Chrischilles poor function indicators (Table 3). The Faurot and Segal **proxy** measures had the strongest associations with chemotherapy receipt, with those in the top third of the predicted probability of **poor function** distribution having 0.64 or 0.66 times the risk of receiving chemotherapy as individuals with in the lowest third. In contrast, having two or more poor function-related indicators was associated with a modest decrease in the receipt of chemotherapy in the Chrischilles model (adjusted relative risk [aRR]: 0.87, 95% CI: 0.83, 0.91), but this effect estimate was attenuated in the Davidoff, Faurot, and Segal models (Table 3).

The associations between predicted probability of poor function and type of chemotherapy mirrored the “any chemotherapy” results and were of similar magnitude (Table 3). The presence of two or more indicators of poor function was associated with a significant decrease in the receipt of the more aggressive FOLFOX regimen compared to individuals with no Davidoff, Chrischilles, and Faurot indicators of poor function, while an association between having one indicator was only seen in the Chrischilles model. Segal indicators were not associated with receipt of FOLFOX. Again, the strongest association between poor function **proxy measure** and FOLFOX receipt was observed using the predicted probability tertiles from the Faurot and Segal models.

A total of 3,404 individuals died during the first year of follow up (12.5%). Individuals with a higher predicted probability of poor function and those with greater number of Chrischilles function-related indicators had an increased one-year risk of all-cause mortality (Table 3). The Faurot predicted probability model demonstrated the strongest association with mortality, followed by Segal, Davidoff and Chrischilles. The presence of two or more Davidoff or Faurot poor function indicators or of three or more Segal indicators was significantly associated with increased risk of one-year mortality; however, the presence of fewer indicators was not significantly associated with mortality (Table 3).

When alternative cut-points for the Davidoff, Faurot, and Segal models were used to mirror the distribution of Chrischilles indicators, the associations seen between the low, intermediate, and high predicted probability of poor function groups and each of the outcomes were similar to the tertile-based analysis (see table, Supplemental Digital Content 3).

Overall survival was computed for individuals who underwent surgical resection and survived past 90 days from diagnosis (n=27,184) by Davidoff, Faurot, and Segal function tertiles and number of Chrischilles indicators (0, 1, 2+), within age group (<75, 75–84, 85+) and strata defined by Gagne combined comorbidity score (0, 1–2, 3+). Survival curves stratified by age and comorbidity score are presented in Figures 2a and 2b, respectively. For

most models, the functional groups at younger ages were associated with clear differences in survival, whereas poor function was less strongly associated with differences in survival among the oldest individuals (Figure 2a). The separation of survival curves appears most prominent using the Faurot and Segal models and smallest using the Davidoff model across all age groups. Differences in survival due to poor function appeared similar across levels of comorbidity for the Faurot and Segal models (Figure 2b), whereas differences in survival across function groups were more pronounced with increasing levels of comorbidity in the Davidoff and Chrischilles models.

Discussion

Though all four Medicare claims-based models were designed to **measure proxies of poor function** using claims data, agreement between the number of Chrischilles, Davidoff, Faurot, and Segal poor function indicators was moderate, and the correlation between predicted probabilities from the Davidoff, Faurot, and Segal models was variable. **As noted earlier, these measures were developed to capture different components of poor function (i.e., frailty, disability). This may complicate direct comparisons across these measures and efforts to distinguish them from classical comorbidity measures, like Charlson or Gagne score.**

Regardless of the **proxy measure** used, poorer function was negatively associated with receipt of laparoscopic surgery, any chemotherapy, and FOLFOX, and positively associated with one-year mortality, with measures of predicted probability of poor function generally exhibiting stronger associations with treatment and mortality outcomes than number of poor function indicators. In particular, the Faurot and Segal models exhibited the strongest associations with receipt of any chemotherapy and FOLFOX and with one-year mortality.

As supported by the literature,³ this study indicates that poor function acts independently of age and **established measures of comorbidity** to impact both clinical decision-making and health outcomes. This finding is in line with previous studies that demonstrated strong associations between the Davidoff¹⁵ and Chrischilles¹⁶ **proxy** measures with receipt of cancer therapies for multiple cancer sites and clinical outcomes after controlling for comorbidity. Poor function was strongly associated with mortality even after controlling for the Gagne comorbidity score, which itself is highly predictive of one-year mortality.⁶ Based on the results of this analysis, poor function has the potential to act as a strong confounder in CER studies in older adult cancer populations, but there are few examples of implementation of these proxy measures for confounding control.¹⁷

This is the first study to directly compare these four Medicare claims-based **proxy** measures of poor function in terms of agreement and associations with treatment and mortality in a single cohort. Of note, even with the ability to directly assess performance-based components of frailty (e.g., weight loss, cognitive impairment), validated clinical measures of frailty have also been shown to have low levels of concordance in cancer cohorts.¹⁸ Despite being developed in different populations and to predict different gold standard measures, the correlations between the claims-based proxy measures suggests that they

capture overlapping constructs. The ability to predict poor function from healthcare claims might be improved by combining elements of these models.

Given that poor function is an important confounder in many observational studies of cancer treatments among older adults, we recommend that at least one Medicare claims-based **proxy** measure of poor function be used to control for confounding in these settings. Predicted probabilities of poor function from the Davidoff, Faurot, and Segal models exhibited stronger and more consistent associations with treatment and mortality outcomes than indicator count-based measures. This is in part due to the weighting of specific indicators more strongly associated with poor function in the Davidoff, Faurot, and Segal measures, and inclusion of “protective” factors that can lower an individual’s probability of poor function (e.g., immunization) in the Davidoff and Faurot models. In addition, the Davidoff and Segal models propose cut-points that identify individuals likely to have a low performance status or to be frail, respectively. However, poor function indicator counts, and particularly two or more indicators, were still associated with chemotherapy receipt, chemotherapy type, and mortality for most models, and could be used for less precise adjustment in settings where ease of implementation is a concern. In terms of selecting a model, factors that could be considered shortcomings in one setting (e.g., inclusion of procedure codes and not diagnosis codes, shorter look-back periods for variable assessment) may in fact be strengths in others (e.g., limited data access). Depending on study context and data availability, investigators should balance usability and performance when determining which measure to use and how to best operationalize these measures for confounding control or cohort identification.

This study had several important limitations. First, all the poor function measures assessed here are imperfect proxies for complex **and distinct** clinical constructs. Even with relatively detailed data, claims cannot capture a complete picture of age-related functional decline, and residual confounding will likely remain. We restricted the study population to those who underwent surgery within 90 days of their diagnosis date; this excluded individuals with longer time to treatment initiation and those who forewent treatment entirely (i.e., the most frail individuals). This may have led to selection into our study based on care access, socioeconomic status, and **level of function**. This approach was chosen to restrict to a population initiating cure-directed therapy (i.e., with an indication for treatment) to mirror the CER setting; our ability to demonstrate associations after excluding the most frail individuals (thereby restricting the amount of potential confounding by **poor function**) therefore represents a conservative approach. We also did not directly estimate associations between treatments and mortality with and without adjustment for poor function, which would help quantify the potential for residual bias attributable to unmeasured **poor function** in this population. Additionally, this study did not examine whether poor function modified any effects of treatment on outcomes. Lastly, while these models were all developed and evaluated using fee-for-service Medicare claims, it is possible that the relative performance of these algorithms would vary depending on patient population and care setting. These topics are important areas for future research.

In conclusion, Medicare claims-based proxy measures of poor function were found to be strongly associated with both treatment receipt and mortality in a cohort of older colon

cancer patients independent of **established** comorbidity **scores**. In studies of older adults relying upon administrative claims data, use of these proxy measures should be considered for confounding control in addition to traditional measures of comorbidity. Efforts to combine these indicators into a single model and incorporate data on medication use could improve the predictive power of future claims-based **proxy measures** of poor function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study used the linked SEER-Medicare database. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement # U58DP003862-01 awarded to the California Department of Public Health.

The ideas and opinions expressed herein are those of the author(s) and endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred. The authors acknowledge the efforts of the National Cancer Institute; 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.

The database infrastructure for this work was supported by the University of North Carolina Clinical and Translational Science Award (UL1TR001111), UNC Lineberger Comprehensive Cancer Center, and the University Cancer Research Fund via the State of North Carolina.

Funding received for this work:

This work was supported by the National Cancer Institute K12CA120780 (JLL) and the National Institute on Aging R01AG056479 (TS). The database infrastructure used for this project was funded by the CER Strategic Initiative of University of North Carolina at Chapel Hill's (UNC) Clinical & Translational Science Award (UL1TR001111) and the UNC School of Medicine and the University Cancer Research Fund.

References:

1. Brookhart MA, Stürmer T, Glynn RJ, et al. Confounding control in healthcare database research: challenges and potential approaches. *Med Care* 2010;48(60):S114–20. [PubMed: 20473199]
2. Clegg A, Young J, Iliffe S, et al. Frailty in elderly people. *Lancet* 2013;381(9868):752–62. [PubMed: 23395245]
3. Fried LP, Ferrucci L, Darer J, et al. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci*. 2004;59(3):M255–63.
4. Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5:649–655. [PubMed: 7165009]
5. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–83. [PubMed: 3558716]
6. Gagne JJ, Glynn RJ, Avorn J, et al. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011;64(7):749–59. [PubMed: 21208778]
7. Davidoff AJ, Zuckerman IH, Pandya N, et al. A novel approach to improve health status measurement in observational claims-based studies of cancer treatment and outcomes. *J Geriatr Oncol*. 2013;4(2):157–65. [PubMed: 23795223]

8. Chrischilles E, Schneider K, Wilwert J, et al. Beyond comorbidity: expanding the definition and measurement of complexity among older adults using administrative claims data. *Med Care* 2014;52(Suppl 3):S75–84. [PubMed: 24561763]
9. Faurot KR, Jonsson Funk M, Pate V, et al. Using claims data to predict dependency in activities of daily living as a proxy for frailty. *Pharmacoepidemiol Drug Saf.* 2015;24(1):59–66. [PubMed: 25335470]
10. Segal JB, Chang HY, Du Y, et al. Development of a claims-based frailty indicator anchored to a well-established frailty phenotype. *Med Care* 2017;55(7):716–722. [PubMed: 28437320]
11. Lund JL, Stürmer T, Harlan LC, et al. Identifying specific chemotherapeutic agents in Medicare data: a validation study. *Medical Care.* 2013;51(5):e27–e34. [PubMed: 22080337]
12. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* 2001;56(3):M146–56. [PubMed: 11253156]
13. Fleiss JL Measuring nominal scale agreement among many raters. *Psychological Bulletin* 1971;76(5):378–82.
14. McNutt LA, Wu C, Xue X, Hafner JP; Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *Am J Epidemiol.* 2003;157(10):940–3. [PubMed: 12746247]
15. Davidoff AJ, Gardner LD, Zuckerman IH, et al. Validation of disability status, a claims-based measure of functional status for cancer treatment and outcomes studies. *Med Care* 2014;52(6):500–10. [PubMed: 24638118]
16. Tan HJ, Chamie K, Daskivich TJ, et al. Patient function, long-term survival, and use of surgery in patients with kidney cancer. *Cancer* 2016;122(24):3776–84. [PubMed: 27518165]
17. Lund JL, Sturmer T, Sanoff HK. Comparative effectiveness of postoperative chemotherapy among older patients with non-metastatic rectal cancer treated with preoperative chemoradiotherapy. *J Geriatr Oncol.* 2016;7(3):176–186. [PubMed: 26926829]
18. Ferrat E, Paillaud E, Caillet P, et al. Performance of four frailty classifications in older patients with cancer: prospective elderly cancer patients cohort study. *J Clin Oncol.* 2017;35(7):766–77. [PubMed: 28095145]

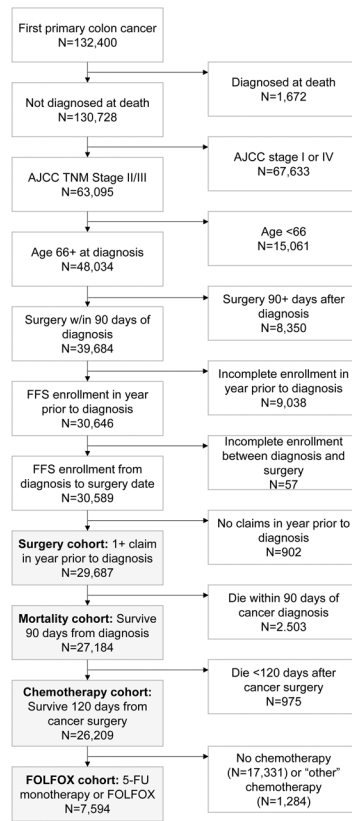


Figure 1.
CONSORT diagram for inclusion and exclusion criteria for each analytic cohort.

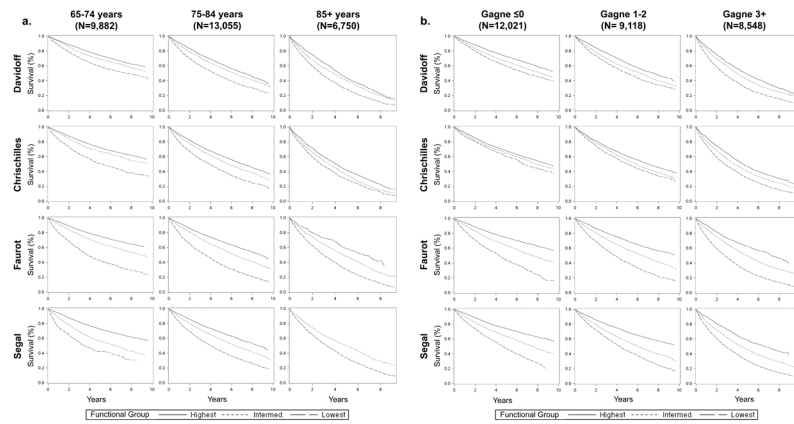


Figure 2. Overall survival from 90 days post-diagnosis stratified by level of function (lowest, intermediate, and highest) and (a) age and (b) Gagne score. Chrischilles stratified by 0, 1, or 2+ indicators of poor function; Davidoff/Faurot/Segal stratified by tertiles of predicted probability of poor function.

Table 1.

Claims-based measures of poor function, performance status, and frailty.

Lead author	Davidoff	Chrischilles	Faurot	Segal
Construct	ECOG performance status	Functional status	Frailty	Frailty
Population	Medicare Current Beneficiary Survey	Medicare beneficiaries hospitalized for AMI	Medicare Current Beneficiary Survey	Cardiovascular Health Study
Proxy measure	Disability status	Function-related indicators	Dependency in ADL	Fried frailty phenotype
Original lookback	Calendar year prior to survey	12 months	8 months	6 months
Claims files	NCH, DME	MEDPAR, OUTPAT, NCH, DME		MEDPAR, NCH OUTPAT
Metric	Predicted probability	Indicator count	Predicted probability	Predicted probability
C-statistic	0.92	0.74–0.79	0.84	0.75
Indicators	<p><u>Positively associated:</u></p> <ul style="list-style-type: none"> Nursing home visit Neurology E&M visit Rheumatology E&M visit Home visit Hospice visit Minor skin procedures Colectomy Major orthopedic procedures Wheelchairs Hospital beds Enteral and parenteral Medical/surgical supplies Other durable medical equipment Imaging/procedure – heart Ambulance Endoscopy - upper gastrointestinal <p><u>Negatively associated:</u></p> <ul style="list-style-type: none"> Dermatology E&M visit Chiropractic E&M visit Musculoskeletal ambulatory procedures Screenings Immunizations/vaccinations 	<ul style="list-style-type: none"> Mobility limitations: Canes; Walkers; Commodes; Bedpans; Urinals; Shower/tub/toilet assistive devices; Lifts, transfer boards, safety belt/harness/vest; Hospital beds; Wheelchairs Blood transfusions Oxygen Supplemental nutrition Hip/pelvic fractures Urinary catheter Chronic ulcer of skin Pneumonia Delirium, dementia, and Alzheimer disease Bone marrow failure/agranulocytosis Depression Respiratory failure, insufficiency, arrest (adult) Sepsis Malnutrition and unintended weight loss 	<p><u>Positively associated:</u></p> <ul style="list-style-type: none"> Arthritis Bladder dysfunction Podiatric care Heart failure Psychiatric illness Home oxygen Hypotension or shock Ambulance transport Stroke/brain injury Dementia Parkinson’s disease Weakness Skin ulcer Paralysis Wheelchair Hospital bed <p><u>Negatively associated:</u></p> <ul style="list-style-type: none"> Cancer screening Lipid abnormality Vertigo Rehabilitation care 	<ul style="list-style-type: none"> Impaired mobility Depression Congestive heart failure Parkinson’s diseases Arthritis (any) Cognitive impairment Stroke Paranoia Chronic skin ulcer Pneumonia Skin and soft tissue infection Mycoses Inpatient hospitalization in past 6 months Gout, crystal-induced arthropathy Falls Musculoskeletal problems Urinary tract infection

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Lead author	Davidoff	Chrischilles	Faurot	Segal
	<ul style="list-style-type: none"> • Major cardiovascular procedures • Bath and toilet aids • Standard imaging - nuclear medicine • Electrocardiography monitoring & cardiovascular stress tests • Endoscopy - sigmoidoscopy, colonoscopy 	<ul style="list-style-type: none"> • Fall-related injury • Syncope 		

Abbreviations: ADL: activities of daily living; AMI: acute myocardial infarction; DME: durable medical equipment; E&M: evaluation and management; ECOG: Eastern Cooperative Oncology Group; MEDPAR: inpatient; NCH: physician services; OUTPAT: outpatient.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Demographic characteristics and comorbidities of full study population by treatment comparison.

	Surgery cohort (N=29,687)			Chemotherapy cohort (N=26,209)			FOLFOX or 5-FU cohort (N=7,594)					
	Laparoscopy (N=8,536)	Open surgery (N=21,151)		Chemotherapy (N=8,878)	No chemotherapy (N=17,331)		FOLFOX (N=3,998)	5-FU (N=3,596)				
Age	N	%	N	%	N	%	N	%	N	%		
<75	3063	35.9	6819	32.2	4798	50.0	4419	25.5	2581	64.6	1535	42.7
75-<85	3714	43.5	9341	44.2	3628	40.9	8031	46.3	1362	34.1	1761	49.0
85+	1759	20.6	4991	23.6	452	5.1	4881	28.2	55	1.4	300	8.3
Sex												
Male	3725	43.6	8636	40.8	4068	45.8	6795	39.2	1893	47.4	1577	43.9
Female	4811	56.4	12515	59.2	4810	54.2	10536	60.8	2105	52.7	2019	56.2
Race												
White, non-Hispanic	6969	81.6	17211	81.4	7061	79.5	14262	82.3	3195	79.9	2868	79.8
Black, non-Hispanic	609	7.1	1806	8.5	690	7.8	1400	8.1	288	7.2	287	8.0
Hispanic	451	5.3	1076	5.1	561	6.3	800	4.6	258	6.5	216	6.0
Other	507	5.9	1058	5.0	566	6.4	869	5.0	257	6.4	225	6.3
Stage												
2	4733	55.5	11598	54.8	2133	24.0	12509	72.2	632	15.8	1138	31.7
3	3803	44.6	9553	45.2	6745	76.0	4822	27.8	3366	84.2	2458	68.4
Comorbidities*												
Mean Gagne score (IQR)	1.52 (0-3)	1.73 (0-3)	1.00 (0-2)	1.78 (0-3)	0.76 (0-1)	1.23 (0-2)						
Mean Charlson score (IQR)	0.87 (0-1)	0.94 (0-1)	0.68 (0-1)	0.94 (0-1)	0.57 (0-1)	0.79 (0-1)						
Hypertension	6479	75.9	15650	74.0	6405	72.1	13064	75.4	2846	71.2	2632	73.2
Dementia	392	4.6	1255	5.9	140	1.6	1093	6.3	36	0.9	75	2.1
Psychosis	378	4.4	1129	5.3	238	2.7	970	5.6	83	2.1	110	3.1
Deficiency anemia	4071	47.7	9944	47.0	3545	39.9	8622	49.8	1442	36.1	1570	43.7
Congestive heart failure	1831	21.5	5221	24.7	1421	16.0	4356	25.1	510	12.8	674	18.7
Fluid/electrolyte disorders	1244	14.6	3706	17.5	1006	11.3	3012	17.4	388	9.7	466	13.0
Chronic pulmonary disease	1746	20.5	4936	23.3	1661	18.7	3911	22.6	646	16.2	744	20.7

	Surgery cohort (N=29,687)		Chemotherapy cohort (N=26,209)		FOLFOX or 5-FU cohort (N=7,594)							
	Laparoscopy (N=8,536)	Open surgery (N=21,151)	Chemotherapy (N=8,878)	No chemotherapy (N=17,331)	FOLFOX (N=3,998)	5-FU (N=3,596)						
Peripheral vascular disease	1633	19.1	4553	21.5	1320	14.9	3866	22.3	493	12.3	619	17.2
Weight loss	118	1.4	576	2.7	63	0.7	392	2.3	19	0.5	31	0.9
Cardiac arrhythmia	2181	25.6	5561	26.3	1693	19.1	4771	27.5	678	17.0	767	21.3
Hemiplegia	103	1.2	373	1.8	57	0.6	304	1.8	16	0.4	33	0.9
Complicated diabetes	916	10.7	2122	10.0	818	9.2	1789	10.3	305	7.6	381	10.6
Renal failure	807	9.5	1882	8.9	515	5.8	1630	9.4	198	5.0	243	6.8
Coagulopathy	343	4.0	921	4.4	293	3.3	739	4.3	119	3.0	129	3.6
Alcohol abuse	87	1.0	222	1.1	61	0.7	182	1.1	27	0.7	26	0.7
Pulmonary circulatory disorders	226	2.7	553	2.6	158	1.8	457	2.6	57	1.4	79	2.2
Liver disease	196	2.3	418	2.0	200	2.3	325	1.9	95	2.4	77	2.1

* Numbers of patients with comorbid HIV/AIDS were suppressed due to cell sizes <11.

Table 3.

Adjusted relative risks and 95% confidence intervals for the association between indicators and predicted probabilities of poor function with treatment or mortality.

	N	Treatment contrast ^a			One-year all-cause mortality (N=27,184) ^b
		Laparoscopic surg. (N=29,687)	Chemotherapy (N=26,209)	FOLFOX (N=7,594)	
Chrischilles					
0 indicators	15,425	1.00	1.00	1.00	1.00
1 indicator	8,275	1.04 (0.99, 1.08)	0.97 (0.94, 0.99)	0.99 (0.95, 1.04)	1.18 (1.09, 1.28)
2+ indicators	5,987	0.89 (0.84, 0.94)	0.87 (0.83, 0.91)	0.92 (0.85, 0.99)	1.45 (1.34, 1.58)
INDICATORS OF POOR FUNCTION					
Faurot					
0	9,238	1.00	1.00	1.00	1.00
1	8,321	1.10 (1.05, 1.15)	1.01 (0.99, 1.04)	0.96 (0.92, 1.00)	1.03 (0.94, 1.14)
2+	12,128	0.99 (0.94, 1.04)	0.90 (0.86, 0.93)	0.93 (0.87, 0.98)	1.42 (1.29, 1.57)
Davidoff					
0	13,504	1.00	1.00	1.00	1.00
1	8,893	1.08 (1.04, 1.13)	1.00 (0.98, 1.02)	1.02 (0.97, 1.06)	1.05 (0.97, 1.14)
2+	7,290	1.03 (0.98, 1.09)	0.95 (0.92, 0.99)	0.86 (0.80, 0.92)	1.25 (1.15, 1.36)
Segal					
0-1	12,915	1.00	1.00	1.00	1.00
2	5,960	1.11 (1.06, 1.17)	1.00 (0.97, 1.03)	0.98 (0.93, 1.03)	1.05 (0.95, 1.15)
3+	10,812	1.01 (0.97, 1.06)	0.97 (0.94, 1.00)	0.96 (0.90, 1.01)	1.21 (1.12, 1.32)
PREDICTED PROBABILITY OF POOR FUNCTION					
Faurot					
Lowest third	9,859	1.00	1.00	1.00	1.00
Middle third	9,932	0.89 (0.85, 0.93)	0.92 (0.90, 0.94)	0.89 (0.85, 0.93)	1.53 (1.38, 1.69)
Highest third	9,896	0.75 (0.71, 0.80)	0.64 (0.60, 0.67)	0.73 (0.67, 0.80)	2.36 (2.12, 2.63)
Davidoff					
Lowest third	9,896	1.00	1.00	1.00	1.00
Middle third	9,895	0.84 (0.80, 0.87)	0.93 (0.91, 0.95)	0.92 (0.88, 0.96)	1.31 (1.20, 1.43)
Highest third	9,896	0.71 (0.68, 0.75)	0.81 (0.79, 0.84)	0.85 (0.81, 0.90)	1.73 (1.60, 1.88)
Segal					
Lowest third	9,899	1.00	1.00	1.00	
Middle third	9,822	1.01 (0.96, 1.07)	0.87 (0.84, 0.90)	0.90 (0.85, 0.96)	1.57 (1.41, 1.76)
Highest third	9,966	0.95 (0.88, 1.02)	0.66 (0.62, 0.71)	0.61 (0.54, 0.68)	2.04 (1.79, 2.33)

^aAdjusted for stage, sex, age (<75, 75-84, 85+), race (White non-Hispanic, Black non-Hispanic, Hispanic, Other), and Gagne comorbidity score (continuous).

^bAdjusted for stage, sex, age (<75, 75-84, 85+), race (White non-Hispanic, Black non-Hispanic, Hispanic, Other), and Gagne comorbidity score (<0, 0, 1, 2, 3, 4+).