

RESEARCH ARTICLE

Open Access

# Patterns of health care utilization preceding a colorectal cancer diagnosis are strong predictors of dying quickly following diagnosis

Robin Urquhart<sup>1,2\*</sup>, Grace Johnston<sup>3</sup>, Mohamed Abdollell<sup>4</sup> and Geoff A Porter<sup>1,2</sup>

## Abstract

**Background:** Understanding the predictors of a quick death following diagnosis may improve timely access to palliative care. The objective of this study was to explore whether factors in the 24 months prior to a colorectal cancer (CRC) diagnosis predict a quick death post-diagnosis.

**Methods:** Data were from a longitudinal study of all adult persons diagnosed with CRC in Nova Scotia, Canada, from 01Jan2001-31Dec2005. This study included all persons who died of any cause by 31Dec2010, except those who died within 30 days of CRC surgery (n = 1885 decedents). Classification and regression tree models were used to explore predictors of time from diagnosis to death for the following time intervals: 2, 4, 6, 8, 12, and 26 weeks from diagnosis to death. All models were performed with and without stage at diagnosis as a predictor variable. Clinico-demographic and health service utilization data in the 24 months pre-diagnosis were provided via linked administrative databases.

**Results:** The strongest, most consistent predictors of dying within 2, 4, 6, and 8 weeks of CRC diagnosis were related to health services utilization in the 24 months prior to diagnosis: i.e., number of specialist visits, number of days spent in hospital, and number of family physician visits. Stage at diagnosis was the strongest predictor of dying within 12 and 26 weeks of diagnosis.

**Conclusions:** Identifying potential predictors of a short timeframe between cancer diagnosis and death may aid in the development of strategies to facilitate timely and appropriate referral to palliative care upon a cancer diagnosis.

**Keywords:** Digestive system neoplasms, Palliative care, Diagnosis, Delivery of health care

## Background

In Canada, colorectal cancer (CRC) has the second and third highest cancer mortality rates for men and women, respectively, with a five-year relative survival of 65% [1]. Stage at diagnosis is the most important prognostic factor for persons diagnosed with CRC, with higher stage at diagnosis associated with decreased probability of survival [2]. In Nova Scotia (NS), ~20% of patients are diagnosed at stage IV, when treatment with curative intent is rarely a viable option [3]. Furthermore, stage-specific five-year relative survival ratios reveal that many persons

diagnosed with less advanced disease, wherein the treatment intent is typically cure, will die *of* or *with* CRC.

Palliative care program (PCP) enrolment is a widely accepted quality indicator of end-of-life care [4,5]. Persons diagnosed with CRC and at risk of dying should benefit from timely enrolment in a PCP. Our research in NS [6-9], as well as that of others' [10], has shown that persons who die quickly following a cancer diagnosis are less likely to access PCPs. Several authors have highlighted the importance of studying populations with lower rates of PCP enrolment in order to enhance access to these important services [11-13].

Understanding the predictors of a quick death following diagnosis may improve timely access to PCPs. Specifically, identifying factors prior to diagnosis that are predictive of dying quickly after a cancer diagnosis may

\* Correspondence: [urquhartr@cdha.nshealth.ca](mailto:urquhartr@cdha.nshealth.ca)

<sup>1</sup>Department of Surgery, Dalhousie University, Halifax, NS, Canada

<sup>2</sup>Cancer Outcomes Research Program, Dalhousie University/Capital District Health Authority, Halifax, NS, Canada

Full list of author information is available at the end of the article

help with early identification of persons who would benefit from early formal palliative care services and perhaps should be targeted for a PCP consultation upon diagnosis. Patient demographics, co-morbidities, and health care utilization prior to diagnosis may be helpful in this regard as they represent potential indicators or “flags” that can be assessed in an efficient manner and at a population-level due to their availability in large administrative databases [14]. These factors also align with the ‘groups’ of predictors classified by Walshe et al. [13] in relation to access to palliative care services.

The objective of this study was to explore whether factors in the 24 months prior to a CRC diagnosis predict a quick death post-diagnosis using classification and regression tree (CART) analysis. CART analysis has been used in health research to explore health service utilization [8,15,16] and to predict health outcomes [17-19]. We previously used CART analysis to identify combinations of factors that, together, identified sub-populations with low rates of PCP enrolment and presented specific areas of potential intervention [8]. In this study, we continue to explore the utility of CART analysis for identifying predictors, and combinations of predictors, that can further our capacity to understand and improve access to palliative services.

## Methods

### Study setting and population

Data for this population-based study are from a large longitudinal, linked administrative database study of all adult persons ( $\geq 20$  years of age) diagnosed with CRC in NS from 01Jan2001 to 31Dec2005. Personnel from the NS Cancer Registry (NSCR) identified all CRC cases diagnosed in NS in the time period, and undertook a comprehensive chart review to stage this cohort using the Collaborative Stage (CS) Data Collection System [20]. Histological classification was determined by International Classification of Diseases (ICD) for Oncology codes [21]: ICD-O3 C18 (excluding C18.1), C19, and C20. Cases diagnosed based on death certificate or autopsy were excluded. If persons had  $>1$  CRC diagnosis in the time period, only one case per patient was retained (the specific rules for retaining cases are reported elsewhere [22]). This cohort ( $n = 3510$ ) was anonymously linked at the patient-level to 14 administrative health databases, providing a comprehensive data source that included clinico-demographic and health service utilization data on all persons with a CRC diagnosis in the province. Further details related to this larger study, including descriptions of the various linked databases and the data linkage process, are reported elsewhere [22]. Ethical approval was obtained from the Capital District Health Authority’s Research Ethics Board.

For the present study, we included all persons in the larger cohort who died of any cause by 31Dec2010. We excluded those who died within 30 days of CRC surgery to remove unexpected deaths arising due to post-operative complications. This resulted in a study cohort of 1885 decedents.

### Variables

Multiple linked databases provided patient and health service utilization data for the 24 months (2 years) prior to CRC diagnosis. The dependent variable was the length of time from CRC diagnosis to death. This variable was analysed as a dichotomous variable (death within various discrete time intervals [yes/no]) and as a continuous variable (in days). Diagnosis and death data were obtained from the NSCR.

The predictor (or independent) variables were all from the 24 months prior to CRC diagnosis, plus stage at diagnosis. Demographic variables included: sex, age at diagnosis (continuous variable, in years), health region, location of residence (urban/rural, using the metropolitan influence zone classification [23]), material and social deprivation indices [24], frailty (yes/no), organ failure (yes/no), and residence in a long-term care facility (yes/no). The presence of frailty and organ failure was computed using algorithms created by Fassbender and colleagues [25], who used a literature review, expert opinion, and cluster analyses to define various causes of death categories (sudden death, terminal illness, organ failure, frailty, and other). The categories ‘organ failure’ and ‘frailty’ are comprised of groups of specific comorbid conditions. Examples of illnesses included in frailty were dementia, Alzheimer’s disease, and Parkinson’s disease; examples of illnesses included in organ failure were chronic obstructive pulmonary disease and congestive heart failure. A person was considered a long-term care facility resident if he/she had at least one physician visit within the last six months of life at a long-term care facility.

Additional socio-demographic variables, obtained from the 2001 census data and linked to patients’ enumeration areas (the smallest geographic unit for which census data are available) were: median household income, unemployment (proportion unemployed), education (proportion without a high school diploma), living arrangement (proportion living alone), marital status (proportion separated, divorced, widowed), single parent (proportion of single parent families), elderly (proportion age 80+; proportion age 85+), immigrants (proportion who are immigrants; proportion who are recent immigrants; proportion who are first generation immigrants), Aboriginal (proportion who are Aboriginal), Black (proportion who are Black), and Francophone (proportion who are Francophone). All census-derived variables were analysed with the proportion represented as a continuous variable.

Health service utilization variables were: number of hospital admissions, total number of days in hospital, number of Emergency Room (ER) visits, number of family physician (FP) visits, number of specialist visits, number of physician visits in a long-term care facility (any physician type), continuity of FP care, and continuity of specialist care. All health service utilization variables were analysed as continuous variables. Cancer site (colon, rectal) was also included as a predictor variable.

Stage at diagnosis was categorized as CS-derived TNM stage I, II, III, IV, or unknown. All variables were extracted from, or computed using variables extracted from, the following administrative databases: NSCR, physicians' billings, hospital discharge abstracts, and national census data.

### Statistical analysis

Binary recursive partitioning is implemented using the CART algorithm. When the outcome is categorical, classification trees are generated that predict the class of a categorical outcome measure (e.g., death within a certain timeframe). When the outcome is continuous, regression trees are generated that predict the value of the continuous outcome measure (e.g., length of time from CRC diagnosis to death). Tree models are generated by repeated binary partitioning of a population, based on a set of predictor variables, such that the cases in the descendant subsets are increasingly more similar within those subsets, so that the overall tree deviance becomes smaller with each successive split. Splitting stops when the cases in a subset are either entirely, or almost entirely, of the same class or the same value, or when further splitting does not improve discrimination between cases (i.e. does not substantially reduce deviance). When a descendant subset of cases cannot be split any further, that subset is labeled as a 'leaf' in the tree. The final set of leaves comprises disjoint sets of cases, with each set being characterized by a series of binary decision rules based on a set of predictor variables [26,27]. As part of this methodology, the algorithm considers all possible splits related to the data, but only selects the optimal splits. Each split is determined in a data driven way, with lower splits all conditional on the prior splits. Researchers do not select or identify *a priori* split-points for each predictor variable; rather, the algorithm selects the split-point that is optimal using a goodness of fit criterion. Strengths of CART methodology include its ability to process large numbers of predictor variables simultaneously, even with considerable interaction amongst variables [28], and the final model is simply interpreted as a set of binary decision rules that together comprise a decision tree and can be implemented intuitively and simply in the clinical setting.

CART models were used to examine pre-CRC diagnosis predictors of time from diagnosis to death using the following discrete time intervals: 2, 4, 6, 8, 12, and 26 weeks from diagnosis to death. Various time intervals were used since there is no agreement in the literature on what constitutes a "quick" death following a cancer diagnosis. All models were run with and without stage at diagnosis as a predictor variable. These models optimally selected split-points identified from across the set of predictor variables, thereby partitioning the data into increasingly more homogeneous, mutually exclusive subsets. All analyses were performed using the R system for statistical computing [29].

### Results

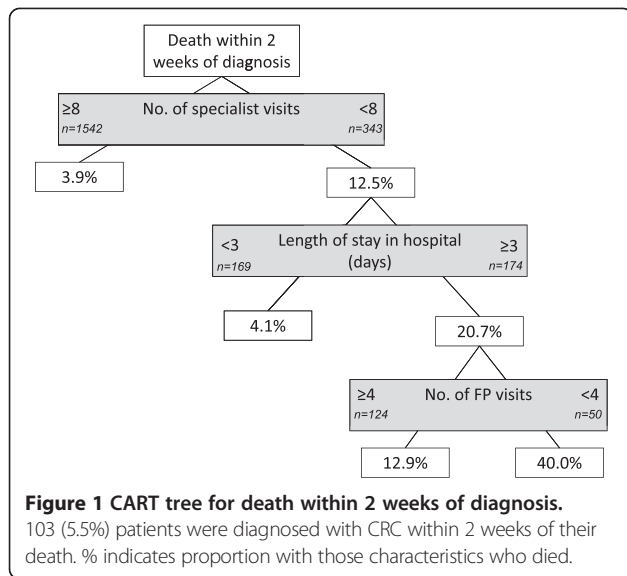
Of this decedent population, 865 (45.9%) were female; 1150 (61.0%) were urban residents; 192 (10.2%) were residents of a long-term care facility; and 183 (9.7%), 435 (23.1%), 467 (24.8%), 668 (35.4%), and 132 (7.0%) were diagnosed with stage I, II, III, IV, and unknown cancer, respectively. The median (range) age at diagnosis was 75.0 (21–101) years. Table 1 depicts the number and percentage of decedents who died within each time interval examined.

Stage at diagnosis was not a strong predictor of dying within 2, 4, 6, and 8 weeks of diagnosis. Stage at diagnosis was the strongest predictor of dying within 12 and 26 weeks (3 and 6 months, respectively) of CRC diagnosis. The strongest, most consistent predictors of dying within 2, 4, 6, and 8 weeks of CRC diagnosis were related to health services utilization in the two years prior to diagnosis: i.e., number of specialist visits, total number of days spent in hospital, and number of FP visits. For all four of these time intervals, having <8 specialist visits in the two years prior to diagnosis was the strongest predictor of dying within the time period. In addition, frailty was a strong predictor of dying within 4 and 6 weeks of CRC diagnosis.

Figures 1, 2, 3, 4 and 5 present the CART trees for the time intervals 2, 4, 6, 8, and 26 (6 months) weeks, respectively. All models presented included stage at diagnosis as a predictor variable, though this factor was *not* a predictor of dying within 2 and 4 weeks of diagnosis

**Table 1 Number and percentage of decedents who died within each time interval**

Time interval	n	%
2 weeks	103	5.5
4 weeks	188	10.0
6 weeks	256	13.6
8 weeks	308	16.3
12 weeks	402	21.3
26 weeks	578	30.7



(and only a moderate predictor of dying within 6 and 8 weeks of diagnosis). Figure 1 shows that among those decedents who had <8 specialist visits, spent  $\geq 3$  days in hospital, and had <4 family physician visits in the 2 years prior to diagnosis, 40.0% died within 2 weeks of diagnosis. Figure 2 shows that among those who had <8 specialist visits, were frail, and had <4 family physician visits in the 2 years prior to diagnosis, 60.0% died within 4 weeks of diagnosis. From Figure 5, it can be seen that among those diagnosed with stage IV or unknown stage with <10 specialist visits and  $\geq 1$  day in hospital in the

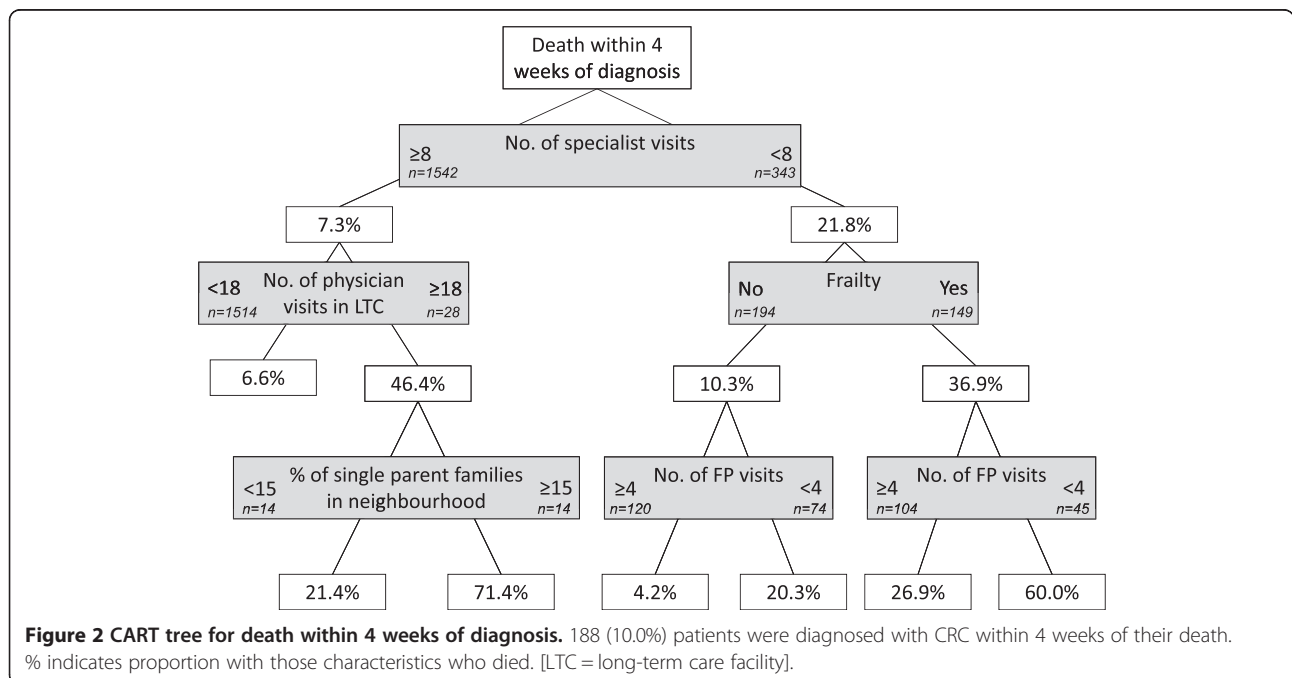
2 years prior to diagnosis, 75.3% died within 26 weeks (6 months) of diagnosis.

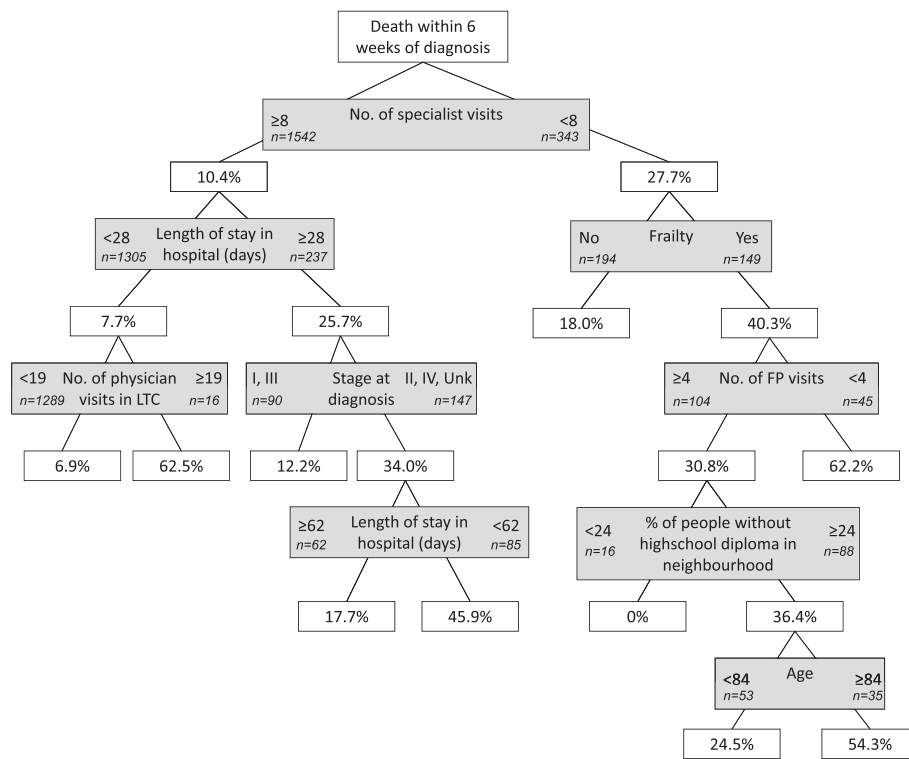
The CART trees also revealed sub-populations with low rates of dying quickly post-CRC diagnosis. Figure 1 reveals that study decedents with high rates of specialist visits and little time spent in hospital preceding diagnosis had low rates of dying within 2 weeks of a CRC diagnosis. From Figure 2, decedents with the lowest rate (4.2%) of dying within 4 weeks of CRC diagnosis were those who had few specialist visits (<8), but were not frail and had more family physician visits ( $\geq 4$ ) in the 2 years preceding diagnosis.

### Discussion

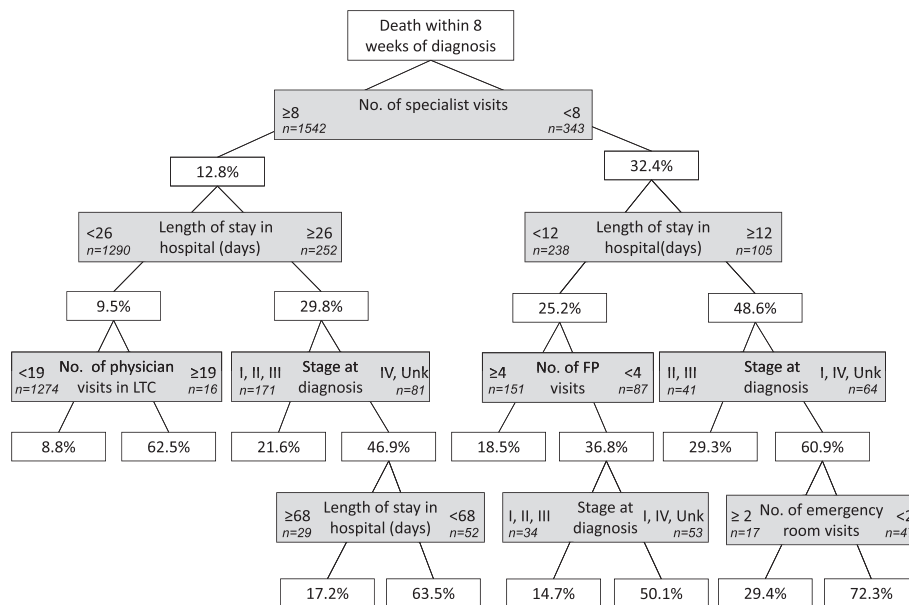
This study used CART methods to explore whether factors in the 24 months prior to a CRC diagnosis predict a quick death post-diagnosis. Identifying pre-diagnosis factors that are predictive of dying within a short time interval following a cancer diagnosis should help us understand whether certain individuals (or groups of individuals) should be targeted for consideration of prompt PCP referral upon diagnosis. We found that fewer physician visits and greater time spent in hospital were strong predictors of dying quickly (i.e., within 8 weeks) after CRC diagnosis whereas stage at diagnosis was not a strong predictor of dying quickly. These findings represent a novel contribution to the palliative care literature; we are unaware of any prior study that has investigated pre-diagnosis predictors of dying quickly following a cancer diagnosis.

The strongest and most consistent predictors of dying within a short timeframe following CRC diagnosis were



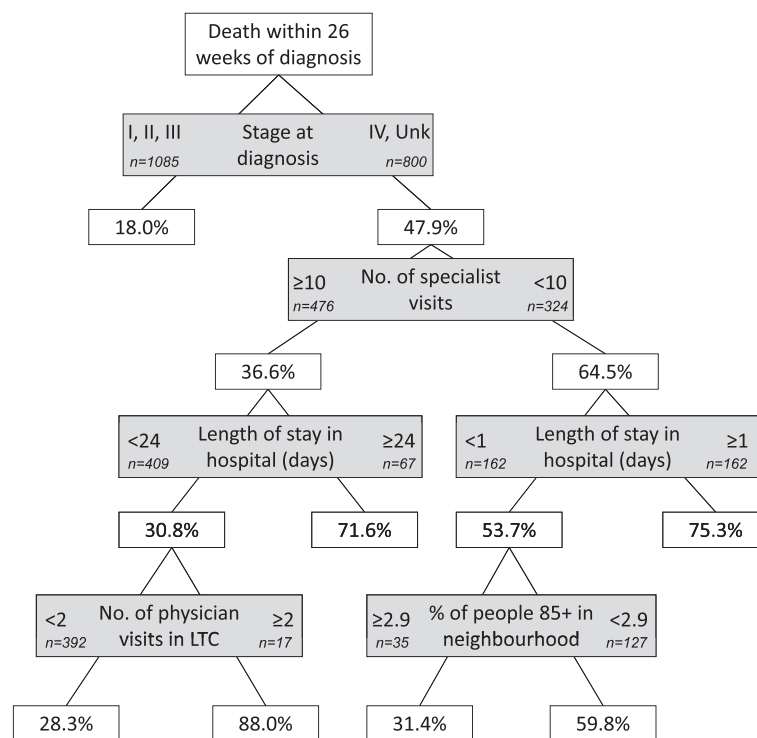


**Figure 3** CART tree for death within 6 weeks of diagnosis. 256 (13.6%) patients were diagnosed with CRC within 6 weeks of their death. % indicates proportion with those characteristics who died. [LTC = long-term care facility].



**Figure 4** CART tree for death within 8 weeks of diagnosis. 308 (16.3%) were diagnosed with CRC within 8 weeks of their death. % indicates proportion with those characteristics who died. [LTC = long-term care facility].





**Figure 5** CART tree for death within 26 weeks (6 months) of diagnosis. 578 (30.7%) were diagnosed with CRC within 26 weeks of their death. % indicates proportion with those characteristics who died. [LTC = long-term care facility].

related to health service utilization in the two years prior to diagnosis. Specifically, fewer visits to specialists and FPs, and longer cumulative time spent in hospital, prior to diagnosis were relatively consistent in terms of identifying sub-populations at risk of dying quickly following CRC diagnosis. The optimum split-points for the factors sometimes differed depending on the time interval studied. Together, these factors may be indicative of poor disease management, with lower use of physician services and higher use of hospital services when conditions or illnesses are not adequately diagnosed and managed out of hospital, leading to deterioration in health and requiring hospital admission. Alternatively, the burden of illness or severity of symptoms experienced by these decedents might have made community-based disease management not feasible, or simply quite difficult. In the latter case, accessing outpatient clinics may have been problematic for these individuals due to high disease burden, resulting in them eventually being admitted to hospital when they could no longer stay at home (or at a long-term care facility). While the exact reasons for this pattern of health care use are unknown, identifying such predictive patterns of health services may provide an opportunity to develop and target interventions to improve patient care [14]. The limited contact with healthcare providers in the sub-populations who die relatively quickly following CRC

diagnosis highlights the importance of ‘upstream’ care in identifying and managing health issues that might increase an individual’s vulnerability for adverse outcomes.

One health issue may be frailty. In this study, frailty was a predictor of dying within 4 and 6 weeks of CRC diagnosis. Numerous studies have shown that frailty is an independent risk factor for major morbidity and mortality in various patient populations [30-33]. Notwithstanding the conceptual, definitional, and measurement issues associated with assessing frailty [34], most researchers agree that frailty increases a person’s vulnerability to adverse outcomes [35], even in the presence of relatively minor stressors [36]. Although frailty is likely highly correlated with age [35], in this study, age was not a strong predictor (nor a weak consistent predictor) of dying quickly after CRC diagnosis. However, a greater number of physician visits in a long-term care facility was a predictor of dying within 4 weeks of CRC diagnosis (Figure 2). It is possible that the CRC diagnosis was an incidental or unanticipated diagnosis for persons with a complex set of symptoms or multi-morbidity.

As demonstrated in the CART trees, our analyses also defined subpopulations, based on interactions amongst predictor variables, with *low* rates of dying quickly after a CRC diagnosis. The decedents with the lowest rate of dying within four weeks of diagnosis were those with few specialist visits, but were not frail and had four or more

family physician visits in the two years before diagnosis (Figure 2). This is interesting given that fewer specialist visits (<8) was often the strongest predictor of dying within a short timeframe. While we can only speculate on possible explanations for this finding, one is that the interaction of these particular factors is indicative of well health. That is, decedents in this subgroup may have been relatively healthy prior to their CRC diagnosis with few comorbid conditions or complex medical needs, at least not severe enough to require high specialist care, and were receiving regular primary care (wherein any comorbid conditions may have been well-managed). Thus, their CRC diagnoses and management may not have been complicated by other health/illness factors.

A number of socio-demographic characteristics that are frequently associated with inequities in cancer care and outcomes, such as income, education, sex, and geographic location, were not key predictors of time from cancer diagnosis to death in this study. In a review of Canadian literature on inequities in cancer care [37], income, geography, sex, and age were frequently associated with both access to palliative care services as well as mortality outcomes. For instance, when logistic regression methods are used, persons diagnosed with cancer from rural areas have been shown to have poorer access to palliative care services [7], while persons with lower socioeconomic status have been shown to have an increased likelihood of dying within six months of diagnosis [38]. We employed a different analytic methodology to identify and group potential predictors of an outcome, which may account for the differences with prior studies.

Certainly, early and complete assessment of a person's health is important in ensuring timely palliative support. Even those who die quickly following a cancer diagnosis should benefit from care that focuses on pain and symptom management, psychosocial well-being, and bereavement support, which are integral features of palliative care. However, care providers are challenged to consistently and accurately anticipate death at the time of diagnosis [39-41], thus potentially compromising timely access to PCPs. This may be especially pertinent if the condition(s) leading to a quick death are outside the scope of a care provider's own medical specialty. That stage at diagnosis was not a strong predictor of a quick death following diagnosis highlights the complexity of anticipating impending death. This study represents an important step toward identifying and elucidating potential predictors of a short timeframe between a cancer diagnosis and death that can aid in the development of flags or algorithms that ultimately facilitate care providers' decision-making with respect to palliative care referral. The utility of pre-diagnosis factors, especially patterns of health care utilization, for predicting a quick death following a cancer diagnosis requires further study.

This study has a number of strengths. First, we examined all persons who were diagnosed with and died of or with CRC in one province. Thus, we identified predictors of a quick death at a population level (versus in a clinical subset of patients). Second, unlike regression analyses, CART can handle large numbers of predictor variables [42] and identifies optimal split-points, eliminating the need to categorize variables *a priori*. As we demonstrated in a prior study [8], by partitioning the data through recursive splitting, CART analysis defines target subpopulations for further investigation and intervention in a more straightforward manner than multiple logistic regression. That is, CART methodology allows one to identify subpopulations that are defined by interactions amongst predictor variables (as shown in the figures). In addition, the final CART model is presented as a decision tree that lays out a set of intuitive decision rules easily used in the field. There are also limitations. The predictor variables were limited to data available in large administrative databases; thus, clinical factors (e.g., functional status, severity of co-morbid illness) were not included in the analyses. Nonetheless, the identification of predictors from administrative databases allows assessment on a population level and provides a means for developing a monitoring system that facilitates the early identification of persons who may benefit from timely referral to palliative care services.

## Conclusion

The strongest and most consistent predictors of dying quickly following a CRC diagnosis were related to health service utilization in 24 months prior to diagnosis whereas stage at diagnosis became the strongest predictor of dying within 12 and 26 weeks after a CRC diagnosis. The pattern of health care utilization predictive of dying quickly following diagnosis may highlight a situation of poor disease management, frailty, or multi-morbidity. Further research is required to corroborate these findings and assist in the development of algorithms that may facilitate timely and appropriate referral to PCPs upon a cancer diagnosis.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

RU participated in the design, coordination, and interpretation of the study, and drafted the manuscript. GJ conceived of the study, and participated in its design and interpretation. MA participated in the design of the study and performed the statistical analysis. GAP participated in the design of the study and its interpretation. All authors read and approved the final manuscript.

## Acknowledgements

We gratefully acknowledge Martha Cox (programmer/analyst) for her work on this project and Cynthia Kendell for her helpful review of the manuscript. This study was funded by a Team Grant from the Canadian Institutes of Health Research (CIHR) and local funding partners (Cancer Care Nova Scotia, Nova Scotia Department of Health and Wellness, Capital District Health Authority, Dalhousie Medical Research Foundation, and Dalhousie University's Faculty of Medicine). The CIHR Grant Number is AQC-83513.

#### Author details

<sup>1</sup>Department of Surgery, Dalhousie University, Halifax, NS, Canada. <sup>2</sup>Cancer Outcomes Research Program, Dalhousie University/Capital District Health Authority, Halifax, NS, Canada. <sup>3</sup>School of Health Administration, Dalhousie University, Halifax, NS, Canada. <sup>4</sup>Department of Diagnostic Radiology, Dalhousie University, Halifax, NS, Canada.

Received: 21 January 2014 Accepted: 14 January 2015

Published: 20 January 2015

#### References

- Canadian Cancer Society's Advisory Committee on Cancer Statistics. Canadian Cancer Statistics 2013. Toronto, ON: Canadian Cancer Society; 2013.
- Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A. AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer; 2010.
- Bu J, Urquhart R, Dewar R, Grunfeld E. Factors Associated With the Late-Stage Diagnosis of Colorectal Cancer In Nova Scotia (NS) [abstract], Health Statistics Data Users Conference: September 21–22, 2009; Ottawa, ON. 2009.
- Earle CC, Park ER, Lai B, Weeks JC, Ayanian JZ, Block S. Identifying potential indicators of the quality of end-of-life cancer care from administrative data. *J Clin Oncol*. 2003;21(6):1133–8.
- Grunfeld E, Urquhart R, Mykhailovskiy E, Folkes A, Johnston G, Burge FI, et al. Toward population-based indicators of quality end-of-life care: testing stakeholder agreement. *Cancer*. 2008;112(10):2301–8.
- Johnston GM, Gibbons L, Burge FI, Dewar RA, Cummings I, Levy IG. Identifying potential need for cancer palliation in Nova Scotia. *CMAJ*. 1998;158(13):1691–8.
- Burge FI, Lawson BJ, Johnston GM, Grunfeld E. A population-based study of age inequalities in access to palliative care among cancer patients. *Med Care*. 2008;46(12):1203–11.
- Gao J, Johnston G, Lavergne MR, McIntyre P. Identifying population groups with low palliative care program enrollment using classification and regression tree analysis. *J Palliat Care*. 2011;27(2):98–106.
- Maddison AR, Asada Y, Burge F, Johnston GW, Urquhart R. Inequalities in end-of-life care for colorectal cancer patients in Nova Scotia. *Canada J Palliat Care*. 2012;28(2):90–6.
- Gray JD, Forster DP. Factors associated with utilization of specialist palliative care services: a population based study. *J Public Health Med*. 1997;19(4):464–9.
- Addington-Hall J, Altmann D. Which terminally ill cancer patients in the United Kingdom receive care from community specialist palliative care nurses? *J Adv Nurs*. 2000;32(4):799–806.
- Burt J, Raine R. The effect of age on referral to and use of specialist palliative care services in adult cancer patients: a systematic review. *Age Ageing*. 2006;35(5):469–76.
- Walshe C, Todd C, Caress A, Chew-Graham C. Patterns of access to community palliative care services: a literature review. *J Pain Symptom Manage*. 2009;37(5):884–912.
- Corkum M, Urquhart R, Kendall C, Burge F, Porter G, Johnston G. Impact of comorbidity and healthcare utilization on colorectal cancer stage at diagnosis: literature review. *Cancer Causes Control*. 2012;23(2):213–20.
- Gregory KD, Korst LM, Gornbein JA, Platt LD. Using administrative data to identify indications for elective primary cesarean delivery. *Health Serv Res*. 2002;37(5):1387–401.
- Peters D, Chen C, Markson LE, Allen-Ramey FC, Vollmer WM. Using an asthma control questionnaire and administrative data to predict health-care utilization. *Chest*. 2006;129(4):918–24.
- Anderson RT, Balkrishnan R, Camacho F. Risk classification of Medicare HMO enrollee cost levels using a decision-tree approach. *Am J Manag Care*. 2004;10(2 Pt 1):89–98.
- Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med*. 2003;26(3):172–81.
- Rovlias A, Kotsou S. Classification and regression tree for prediction of outcome after severe head injury using simple clinical and laboratory variables. *J Neurotrauma*. 2004;21(7):886–93.
- National Cancer Institute. Collaborative Stage. <http://seer.cancer.gov/tools/collabstaging/>.
- Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, et al., editors. International classification of diseases for oncology - ICD-O-3. Geneva, Switzerland: World Health Organization; 2000.
- Porter G, Urquhart R, Bu J, Kendall C, Macintyre M, Dewar R, et al. A team approach to improving colorectal cancer services using administrative health data. *Health Res Policy Syst*. 2012;10(1):4.
- Janes D, McNiven C, Puderer H. Census Metropolitan Area and Census Agglomeration Influenced Zones (MIZ): A Description of the Methodology. Ottawa, ON: Geography Division, Statistics Canada; 2000.
- Pampalon R, Raymond G. A deprivation index for health and welfare planning in Quebec. *Chron Dis Can*. 2000;21(3):104–13.
- Fassbender K, Fainsinger RL, Carson M, Finegan BA. Cost trajectories at the end of life: the Canadian experience. *J Pain Symptom Manage*. 2009;38(1):75–80.
- Breiman L. Classification and Regression Trees. New York, NY: Chapman & Hall; 1984.
- Zhang H, Singer BH. Recursive Partitioning and Applications. 2nd ed. New York, NY: Springer; 2010.
- Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods*. 2009;14(4):323–48.
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; ISBN 3-900051-07-0; 2012. [<http://www.R-project.org/>]
- Dasgupta M, Rolfson DB, Stolee P, Borrie MJ, Speechley M. Frailty is associated with postoperative complications in older adults with medical problems. *Arch Gerontol Geriatr*. 2009;48(1):78–83.
- Garcia-Gonzalez JJ, Garcia-Pena C, Franco-Marina F, Gutierrez-Robledo LM. A frailty index to predict the mortality risk in a population of senior Mexican adults. *BMC Geriatr*. 2009;9:47.
- Lee DH, Buth KJ, Martin BJ, Yip AM, Hirsch GM. Frail patients are at increased risk for mortality and prolonged institutional care after cardiac surgery. *Circulation*. 2010;121(8):973–8.
- Makary MA, Segev DL, Pronovost PJ, Syin D, Bandeen-Roche K, Patel P, et al. Frailty as a predictor of surgical outcomes in older patients. *J Am Coll Surg*. 2010;210(6):901–8.
- Partridge JS, Harari D, Dhesi JK. Frailty in the older surgical patient: a review. *Age Ageing*. 2012;41(2):142–7.
- Rockwood K. What would make a definition of frailty successful? *Age Ageing*. 2005;34(5):432–4.
- Campbell AJ, Buchner DM. Unstable disability and the fluctuations of frailty. *Age Ageing*. 1997;26(4):315–8.
- Maddison AR, Asada Y, Urquhart R. Inequity in access to cancer care: a review of the Canadian literature. *Cancer Causes Control*. 2011;22(3):359–66.
- Groome PA, Schulze KM, Keller S, Mackillop WJ. Demographic differences between cancer survivors and those who die quickly of their disease. *Clin Oncol (R Coll Radiol)*. 2008;20(8):647–56.
- Chow E, Harth G, Hruby G, Findelstein JW CD. How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review. *Clin Oncol (R Coll Radiol)*. 2001;13(3):209–18.
- Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ*. 2000;320(7233):469–72.
- Glare P, Virik K, Jones M, Hudson M, Eychmuller S, Simes J, et al. A systematic review of physician's survival predictions in terminally ill cancer patients. *BMJ*. 2003;327(7408):195.
- Berger AM, Berger CR. Data mining as a tool for research and knowledge development in nursing. *Comput Inform Nurs*. 2004;22(3):123–31.

doi:10.1186/1472-684X-14-2

Cite this article as: Urquhart et al.: Patterns of health care utilization preceding a colorectal cancer diagnosis are strong predictors of dying quickly following diagnosis. *BMC Palliative Care* 2015 :2.