Joint Modelling of Colorectal Cancer Recurrence and Death after resection using Multi-State Model with Cured Fraction

Behnaz Alafchi

The University of Manchester Faculty of Biology Medicine and Health

Ghodratollah Roshanaei

Hamadan University of Medical Sciences School of Public Health

Leili Tapak

Hamadan University of Medical Sciences School of Public Health

Mohammad Abasi

Hamadan University of Medical Sciences Medical School

Hossein Mahjub (✉ mahjub@umsha.ac.ir)

Hamadan University of Medical Sciences School of Public Health <https://orcid.org/0000-0003-3274-2650>

Research article

Keywords: Colorectal Neoplasms , Recurrence, Cure model, Multi-state model, Markov, Survival

Posted Date: October 9th, 2019

DOI: <https://doi.org/10.21203/rs.2.15882/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on January 13th, 2021.
See the published version at <https://doi.org/10.1038/s41598-020-79969-6>.

Abstract

Background Curing of colorectal cancer (CRC) occurs at the time of resection, but it is not immediately observable. If the cancer is not completely eliminated, the patient will not be cured of cancer and will experience a recurrence as the tumor has regrown to a detectable size. The main proposes of the present study was to assess the effects of different covariates on the probability of being cured as well as the time to recurrence, time to death, and time to death after recurrence in CRC patients by using multi-state cure model. **Methods** In the present study, the information of 283 patients with adenocarcinoma CRC, who underwent resection, from 1992 to 2015 in Imam Khomeini Hospital of Hamadan, Iran, were analyzed. A multi-state cure model is used to joint modeling the recurrence and death in patients with CRC when a fraction of patients was apparently cured after resection. **Results** The results revealed that females, patients diagnosed at stage IV and whom underwent radiotherapy were less likely to be apparently cured. Also, more than 50% of not cured patients recurred later. Moreover, the survival time of patients was affected by the stage of disease, age at diagnosis and receiving radiation therapy. In addition, sex had a significant effect on the time-to-recurrence. **Conclusions** The multi-state cure model provided a flexible framework to study and compare the effects of prognostic factors simultaneously on the transition between different health states and the probability of being apparently cured of CRC.

Background

Colorectal cancer (CRC) is a major cause of mortality and morbidity. It is one of the most commonly diagnosed cancer (1) and the fourth common cause of cancer-related death worldwide (2, 3). The CRC survival is highly dependent upon the stage of disease at diagnosis as well as the possibility of resection of the tumor (7, 8).

The only curative treatment for CRC patients is complete surgical resection. About 70% to 80% of patients are eligible for curative resection (9, 10). Moreover, almost two-thirds of CRC patients underwent resection but 30% to 50% of these patients will experience recurrence and will die of CRC (9, 14). However, curing happens at the time of treatment, and it is not immediately observable. If cancer cannot be completely eliminated, the patient will not be cured and will experience a recurrence as the tumor has regrown to a detectable size (15). Patients who experience a recurrence are substantially at higher risk of mortality, and it is essential to find out who prognostic factors predispose a patient to recurrence (16-20). The time of occurrence of recurrence also influences the overall survival, such that worse survival is associated more with the early recurrence than the late one (21).

Recurrence can modify the cancer progression, and the effects of prognostic factors may have be changed by recurrence. So, the effects of variables on the time of recurrence, the survival time and the recurrence-free survival time may be different. Therefore, especial statistical methods are required to obtain such effects (22, 23). Various analytic methods can be used to analyze the time of recurrence or time-to-death in such settings (17, 18, 24-26). Usually, separate survival analyses are carried out for these clinical events. Nevertheless, these separate analyses are not completely satisfying because they may fail

to reveal the relationships between recurrence and death event (27, 28). On the other words, a patient may experience a clinical progression (e.g. a local recurrence, followed by a metastatic recurrence and then death), so instead of the occurrence of a single event, the progression of disease should be modeled jointly.

A common way to joint modeling of different types of events is to use multi-state models, which describe the progression of the disease and transitions between different states over time. In this model, each event or each transition between events is considered as a disease state (29). Multi-state models are usually specified by using transition intensities and can be based on two-time scales including: the calendar time and the duration time in the current state, called Markov and semi-Markov model, respectively. Markov models assume that the future evolution of the system depends on the history through current state. However, if the sojourn/duration time in the current state has been indicated as a significant covariate in the model, semi-Markov model should be used. This type of Markovian model is called the clock-reset model, setting the clock back to zero at the time of entry in a new state (30, 31). On the other hand, as some patients may be apparently cured after treatment and will never experience recurrence of CRC, cure models should be used.

Cure models are used to model many different types of diseases when a substantial proportion of patients are completely cured by the treatment and will never experience the clinical recurrence (32). A multi-state cure model is a multi-state model which incorporates a latent cured state and combines the aspects of both multi-state and cure model to investigate the effects of the covariates on both curing of the disease and the disease processes as well as dealing with the association between different events of interest (recurrence/death), simultaneously (33). Although the progression of CRC disease includes different states and a fraction of patients apparently cured after the resection, there has not yet been conducted any study that analyzes CRC data using this model. So, in this study, a multi-state cure model is used to joint modeling of the recurrence and death in patients who developed CRC considering the probability of being apparently cured after resection.

Methods

2.1. Data

The information of 283 patients with adenocarcinoma CRC who underwent curative resection and admitted to Imam Khomeini Hospital in Hamadan in the west of Iran, between 1992 and 2015 were analyzed in this study. Patients were followed to August 2017. The information of vital status and date of death was obtained through medical and administrative recorded sources. Here, all deaths were considered as CRC-related deaths. The information of baseline demographic and clinical variables including sex (male/female), age at diagnoses (year), body mass index (BMI), metastasis (yes/no), stages of the disease based on AJCC (36) classification (stage I/stage II/stage III/stage IV), receiving chemotherapy and/or radiotherapy (yes/no), were collected from medical records. The outcomes of

interest were time of entry in each state including time to recurrence, time to death and time to death after recurrence. All patients who were alive at the end of study were censored for death and who did not experience recurrence of CRC were also censored for the recurrent event.

2.2. Multi-state Cure Model

According to the available information, the patients who recurred during their follow-up were assumed as not cured patients with observed recurrence time, while the patients who did not recur during their follow-up were assumed to be censored for recurrence. Apparently, the cured patients would have never recurred even if they had been followed longer. Each patient can transient to death state either with or without a prior recurrence. So, the multi-state model (shown in figure 1) consisted of four states (apparently cured, not cured, recurrence and death) and then there were four transitions between these states including: (1) the transition from not cured group to death (2) the transition from apparently cured group to death, (3) the transition from not cured group to recurrence and (4) the transition from recurrence to death. Patients with unknown exact time of death were considered as censored. As shown in figure 1, there were two latent states that each patient was assigned to one of them based on her/his information. It should be noted that the censored times for death and recurrence were not necessarily equal as the recurrent time needs an active follow-up while (death obtaining) death/live status can be obtained at a later time.

In the standard mixture formulation of the cure model, when there is one event of interest, the marginal survival function, $S(t)$, is given by $S(t) = P + (1 - P)S_u(t)$, where $S_u(t)$ is the conditional survival function for the uncured group and the survival for the cured group is equal to 1. P shows the proportion of the population who has never experienced the event of interest (in this case CRC recurrence). Here P provides information about the tumor and the effect of the treatment on cell killing.

A logistic model, $P = \text{logit}^{-1}(X\beta)$, is used to describe cured fraction, where X is the vector of the covariates (32). The area under the ROC curve (AUC) was used as evaluation criteria for the logistic regression model as a classifier.

The distributions of event times given cured status are described by proportional hazards models. The hazard for transition from state k to state j for the i th subject based on proportional hazards model is defined as,

$$\lambda_{jk}(t_i; X_i) = \lambda_{jk}^0(t_i) \exp(\beta_{jk}^T X_i) \quad (1)$$

Where X_i is the vector of covariates that their effects (their associated coefficients β_{jk} for transition from state j to state k) can change over states and $\lambda_{jk}^0(t_i)$ is the baseline hazard for subject i for transition from state j to state k at time t.

Then the survival distribution for each transient state is constructed by $\lambda_{jk}(t_i)$ as follows;

$$\begin{aligned} S_1(t) &= \exp\left(-\int_0^t \lambda_{14}(u) du\right) \\ S_2(t) &= \exp\left(-\int_0^t \lambda_{23}(u) du - \int_0^t \lambda_{24}(u) du\right) \\ S_3(t | t_r) &= \exp\left(-\int_0^{t-t_r} \lambda_{34}(u) du\right) \end{aligned} \quad (2)$$

Where t_r is the entry time into state 3.

Here, a Markov multi-state model was used, with the clock set back to 0 at the time of entry into a new state.

So for transition 1→4 and 2→4, t_i is the death or censored time, for transition 2→3, t_i is the recurrence or censored time, and for transition 3→4, t_i is the gap time between entry into state 3 and state 4 (i.e. $t - t_r$) and a proportional hazard regression was used to model the residual time in state 3 before entering state 4 as follow:

$$\lambda_{34}(t - t_{ir}) = I(t > t_{ir}) \lambda_{34}^0(t - t_{ir}) \exp(\beta_{34}^T X_i). \quad (3)$$

In addition, for the transition from state 3 to state 4, the time-to-recurrence (sojourn time in the recurrence state) is considered as a covariate in addition to other covariates. All baseline hazards are assumed unrestricted (i.e. the equality of baseline hazards for transition 1→4 and 2→4 was not compulsory).

The likelihood function for the multi-state cure model can be found in Conlon et al (34). An EM algorithm proposed by Lauren et al (35) was used to estimate the parameters of the model. For variable selection, separate proportional hazard models were fitted for all transitions. If a variable was statistically significant in a model, it was added to the multi-state cure model. However, for the transition from recurrence to death state, the model was faced with the problem of non-convergence due to the small sample size. To solve this issue, we included only the clinically important variables in the multi-state cure model for this transition. The proportional hazards assumption was checked for each variable in each transition using Schoenfeld residuals. All statistical analyses were performed at a significance level of 0.05 using the MultiCure package of R software, version 3.5.3 (The R Foundation for Statistical Computing, Vienna, Austria, RC Team. URL <http://www.R-project.org>).

Results

Of 283 patients underwent curative resection for CRC, 99 (35%) patients had rectal and 184 (65%) patients had colon adenocarcinomas. The frequency of CRC in both male and female sexes was almost the same (52.7% female and 47.3% male) with mean age of 55.58 ± 13.127 . For more than 90% of patients, surgery was the first treatment that they received. About 67 (23.7%) patients were diagnosed with metastasis and 44 patients developed metastasis during their follow up. Overall, 40 percentages of the patients had metastatic CRC (45.5%, 9.1%, 8.2%, and 37.2% of metastasis were in liver, lung, lung and liver, and other tissues, respectively). All of the patients were diagnosed being at more advanced stages such that none of them were diagnosed at stage I, 132 (46.6%) at the stage II, 84 (29.7%) at the stage III, and 67 (23.7%) at the stage IV). The number of patients who have received chemotherapy and radiotherapy after resection were 242 (85.5%) and 89 (31.4%), respectively. Duration, frequency, type/dose of drug and the number of cycles of chemotherapy (ranged from one to 39 by mean of 6.41 ± 4.6 sessions) were different among patients. During the study, 44 (15.5%) patients experienced the recurrence after resection.

Figure 2 shows the Kaplan-Meier survival curve for recurrence and death times. According to the figure, there is no clear level plateau, especially for recurrence. The lines in Figure 3 depict the follow-up times of overall survival for each subject. The events and censoring times for both recurrence and death (dots) were also shown in Figure 3 (Figure 3 (a), (b), (c), and (d)), indicating that there was not unequal censorship. So, we assumed the censoring times for both events (recurrence and death) were equal. It could be seen that the majority of the observed recurrences occurred early in follow-up and were slowed down substantially by about 50 months (Figure 3 (a)). So, as there was a very small probability to experience the recurrence after 50 months, it seems reasonable to assume that the subjects who were still at risk for recurrence and death were apparently cured of CRC. Markov assumption was satisfied in our CRC data, as the coefficient of the sojourn time was not significant ($p=0.655$).

Figure 1 (b) represents the number of subjects in each transition where the number of apparently cured and not cured patients were estimated by the model. Most of the patients, about 70%, were apparently cured of CRC. The summary statistics including the number of patients, the number of events, and the number of censored patients in each transition as well as the estimation of mean (SE), median and other quartiles of the length of stay in each transition were provided in Table 1. It should be noted that the 25th quartile (Q25) and 50th quartile (Q50) did not exist for the transition apparently cured to death (see its survival plot in the appendix).

We assessed the accuracy of the logistic regression model as a classifier using AUC. We obtained $AUC = 0.784$ which indicated a promising predictive accuracy for this model. The results of multi-state cure model were provided in Table 2 which shows the effects of the variables including age at diagnosis, sex, BMI, stage of the disease, and treatments (chemotherapy and/or radiotherapy) on the hazards of different transitions between various states and also on the probability of being not cured of CRC after resection. As there was a significant association between metastasis and the stage of disease ($p<0.001$), metastasis was not included in the model. Moreover, it should be noted, as all of the patients who were not cured of CRC and recurred during the follow-up received chemotherapy, it was not possible to assess

the effects of chemotherapy on the survival time of these patients.

As the results showed (Table2 (A)), sex, stage of disease at the time of diagnosed and receiving radiotherapy had a significant effect on the probability of not being cured, such that females ($p=0.027$), patients at advanced stage IV ($p=0.045$), and who received radiotherapy ($p=0.041$) were less likely to be cured of CRC. The results also shown that the hazard of death among apparently cured patients increased as the age of the patients increased ($p=0.011$) (Table2 (B)).

Table2 part C showed that the risk of death for not cured patients increased substantially by diagnosing the disease at more advanced stages, as patients at stage III ($p<0.001$) or stage IV ($p=0.003$) compared with patients at the stage II were at higher risk of death. Moreover, receiving radiotherapy significantly decreased the risk of death among these patients ($p=0.046$). Our results also revealed that the risk of recurrence in not cured patients was significantly associated with sex ($p=0.015$) such that the risk was about 2.6 times higher in men (Table2 (D)). Table2 part E showed the results for the transition from recurrence to death state. This table showed that after recurrence of CRC, patients who diagnosed at stage IV were at higher risk of death, compared with patients at stage II ($p=0.042$).

Figure 4 (a) shows the state occupancy probabilities (constructed by the probability that an individual be in a specific state at any given time) as a method of visualizing the results of the multi-state model. In this figure, the horizontal axis represents time in months and the vertical axis shows the cumulative probability of being in a particular state. It can be seen that the risk of death has increased over time either in apparently cured or not cured patients. It also exhibited that most of the patients did not experience the recurrence and were alive at the end of the follow-up period. The overall survival (the time from treatment until death or censoring at the last follow up.) and progression free (event-free) survival (the time from treatment to recurrence or death whichever occurs first) were two interested endpoints. Figure 4 (b) and (c) show these two survival probabilities at different times. The cumulative hazards for all four transitions were also represented in Figure 4 part (d).

Discussion

In the present study, a multi-state Markov model was used to joint modeling of recurrence and death in colorectal cancer (CRC) with an incorporated cured fraction, in order to study the factors that influence the transition intensities between different states. The structure of this multi-state cure model was motivated by the disease process of CRC. This model was first introduced by Conlon et al in 2014 to analyze colon cancer data (34) and Lauren et al in 2018 extended an EM algorithm to estimate the parameters of this model and applied their model on head and neck cancer data (35). As we were awarded, there were no studies that have applied this model on CRC data and assessed the effects of variables on recurrence and death events jointly in the presence of the cured fraction of patients.

It has been reported that 30% to 50% of CRC patients who underwent resection will experience the recurrence (9, 14). Although, our results revealed that the tumor of a significant proportion of patients was eliminated by the treatment so that they will never experience a recurrence of CRC. Moreover, as there were a sufficient follow-up period and a number of patients who were censored for recurrence after the

last observed time (the Kaplan–Meier survival plot for recurrence event in Figure2 shows a clear level plateau), it was justifiable to use a mixture cure model for the recurrence event (32). On the other hand, as the recurrence and death events were correlated, joint modeling of recurrence and death events could diminish the bias which might occur in separated model.

The joint modeling of the recurrence and survival time could also aid in identifiability of the cure part of the model because subjects with survival greater than the last observed recurrence time (which was approximately 50 months) were considered to be apparently cured of the recurrence. The probability of being cured for subjects with survival times lower than 50 months did not experience the recurrence have been estimated by logistic model and according to this model the subjects were appointed to apparently cured/not cured latent states. About 70% of our CRC patients were classified in cured group based on the model.

According to the findings, there was a significant association between sex, stage and receiving radiotherapy with the probability of being cured. The result suggested that females were less likely to be cured of CRC (the odds of being cured in males was 2.747 times of females). Based on our results, patients who diagnosed at stage IV were at about twofold higher odds of not being cured of CRC compared to patients at the stage II. In addition, patients who underwent radiotherapy were about 2.841 times more likely to be cured. However, the effect of the variables on the recurrence of CRC have been assessed in other studies (37-39), we have not found any study that have investigated its effect on the probability of being cured.

Generally, in patients with CRC, death can occur with or without a prior recurrence. The deaths following a recurrence may be due to the cancer, whereas the deaths without a prior recurrence are known not to be directly due to the regrowth of the tumor (34). However, the cause of death was not considered in this study, and we have not followed this issue as a competing risk event.

The results showed that the survival time of patients after resection was affected by age at diagnosis (among patients who their cancer was completely eliminated by resection), and stage of the disease and receiving radiotherapy (among patients who may experience the recurrence lately). It can be concluded from the findings, as the age of patients increased, the hazard of death in apparently cured patients increased. The effect of age on the survival of CRC patients had been assessed in different studies which are controversy. Some of them indicated that older patients are at higher risk of death (25, 43, 44), while the results of some others did not show a significant association between age and the risk of death in CRC patients (45-48). However, none of these studies assessed the effect of age separately on the survival rate of apparently cured and not cured CRC patients. Moreover, it has been shown that the risk of death was substantially higher in patients diagnosed with more advanced stages (patients at stage IV and III were at higher risk of death compared those at stage II). It should be noted that stage of the disease at diagnosis were just significantly effective on the survival time of patients whose disease was not cured by resection and the tumor of more than 50% of these patients, was regrowth. Other studies

also showed a significant association between the stage of the CRC at diagnosis and the survival time (40, 46, 49, 50).

Based on our results, the risk of death and recurrence of CRC was lower among patients who underwent radiation therapy. Different clinical trials also have shown that the adjuvant radiotherapy can decrease the rate of death and recurrence among patients with CRC after surgery and improve their survival time. However, its effect was dependent on dose regimen (53, 54).

On the other hand, two competing risks (recurrence/death) were encountered by the patients after resection. Among 30% (85 of 283) of patients who were not cured by resection, 51.8% (44 of 85) experienced the recurrence. In this study we have founded that males were at higher risk of recurrence after resection by $HR=2.614$. Tartter (38) and Kobayashi et al (51) have showed that the risk of recurrence is significantly different in both males and females patients with colon and rectal cancer while based on the survival analysis there were no association between sex and recurrence time or disease-free survival. Dancourt et al (25) by joint modeling of recurrence and death in CRC data using a multi-state model showed that the time of recurrence is affected by sex and males were more likely to be recurred, which was aligned with our finding.

The results also showed that the risk of death after recurrence among patients who were diagnosed at stages III and IV, were 2.965 and 4.169 times of the patients who were at the stage II, respectively. Other clinical study also had showed that the patients who underwent resection and diagnosed at the stage III had a greater probability of death after experiencing the recurrence than the patients at the stage I&II (25). The coefficient of sojourn time in recurrent state, which had been included as a covariate in transition from state 3 to state 4, indicated that the Markov assumption was satisfied for our CRC data.

The results also revealed that BMI did not have a substantial effect on any transition intensities. Vrieling and Kampman conducted a review to investigate the role of BMI, diet, and physical activity in CRC recurrence and survival rate. They revealed that there was a high association between BMI before or at the time of diagnosis and mortality or recurrence of CRC. In this review, they just found one study that assessed the effect of BMI on CRC survival rate, so there was a need to more studies to examine the effect of this variable on progression or survival time of the CRC (52).

As most of our patients underwent chemotherapy, assessing the effect of this variable were not possible in two transitions. Chemotherapy schedule was different. Most of the studies have assessed the effect of chemotherapy on the rate and time of (local) recurrence after resection (9, 55-58). Collaborative Group showed that the relative risk of recurrence and death were higher in patients underwent chemotherapy. However, according to their findings there was no significant difference in efficacy of treatment by chemotherapy schedule (57).

This study has some limitations. First, for survival analysis, reliable data based on prospective cohort studies are required. However, our data were based on a retrospective study, and information was based on the data recorded by registry centers. Therefore, we were unable to assess the accuracy of the data.

This issue may introduce information bias. Moreover, according to this limitation some important variables such as period /exposition to chemotherapy /radiotherapy and clinical state of the patients were not included in the collected data. Second, although patients were followed about 25 years, the number of all available patients who underwent resection was limited. On the other hand, as in the multi-state cure model, there were many parameters to be estimated and their number increased by the number of variables in each transition, our sample size was relatively small. Due to this limitation, the confidence interval of some HR was relatively wide. It is clear that bigger sample sizes will provide much more precise estimates.

Despite these limitations, the main purpose of the present study was to use powerful statistical methods (here multi-state cure model) which take all aspects of the data into account. In the future clinical studies, it is suggested that if there were different states of disease, multi-state or multi-state cure model would be used instead of separated models to analyze the data. Moreover, in this study, the used model was based on the semi-parametric Cox model. Although, based on the Schoenfeld residuals, the PH assumption was reasonable, it is suggested to use log-linear models and to develop R codes for this extended model in future studies.

Conclusions

The multi-state cure model provided a flexible framework to study and compare the effects of prognostic factors simultaneously on transition between different health states and the probability of being cured of CRC.

Declarations

Ethics approval and consent to participate

This study has been approved by the Research Council of Hamadan University (No. 9504151863). Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Availability of data and material

The datasets analyzed during the current study are available from the corresponding author on reasonable request with permission of MA.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was a part of Ph.D. thesis of the first author and it was supported by the Hamadan University of Medical Sciences.

Authors' contributions

MA and BA designed and performed the research. Data analysis, interpretation and manuscript preparation was performed by BA, HM, GhR and LT. All authors contributed to the final version of the manuscript; revised the article critically for important intellectual content and approved the final manuscript.

Acknowledgement

We would like to thank the Vice-Chancellor of Research and Technology, Hamadan University of Medical Sciences for the approval and support of the study.

References

1. Cancer IAFRo. GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012. 2012.
2. Fund WCR, Research AlfC. Food, nutrition, physical activity, and the prevention of cancer: a global perspective: Amer Inst for Cancer Research; 2007.
3. Janout V, Kollárová H. Epidemiology of colorectal cancer. Acta-Universitatis Palackianae Olomucensis Facultatis Medicae. 2001:5-10.
4. Boyle P, Langman J. Epidemiology. Bmj. 2000;321(7264):805-8.

5. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA: a cancer journal for clinicians*. 2005;55(2):74-108.
6. Jackson-Thompson J, Ahmed F, German RR, Lai SM, Friedman C. Descriptive epidemiology of colorectal cancer in the United States, 1998–2001. *Cancer*. 2006;107(S5):1103-11.
7. Jemal A, Clegg LX, Ward E, Ries LA, Wu X, Jamison PM, et al. Annual report to the nation on the status of cancer, 1975–2001, with a special feature regarding survival. *cancer*. 2004;101(1):3-27.
8. Ries L, Eisner M, Kosary C, Hankey B, Miller B, Clegg L, et al. *Edwards. SEER cancer statistics review*. 1975;2000.
9. Abulafi A, Williams N. Local recurrence of colorectal cancer: the problem, mechanisms, management and adjuvant therapy. *British journal of surgery*. 1994;81(1):7-19.
10. Faivre-Finn C, Bouvier AM, Mitry E, Rassiati E, Clinard F, Faivre J. Chemotherapy for colon cancer in a well-defined French population: is it under-or over-prescribed? *Alimentary pharmacology & therapeutics*. 2002;16(3):353-9.
11. Kleihues P. *World Cancer Report*. International Agency for Research on Cancer IARC Press, Lyon, France. 2003.
12. Gatta G, Faivre J, Capocaccia R, De Leon MP, Group EW. Survival of colorectal cancer patients in Europe during the period 1978–1989. *European journal of cancer*. 1998;34(14):2176-83.
13. Berrino F, Gatta G, Chessa E, Valente F, Capocaccia R. The Eurocare II study. *European Journal of Cancer*. 1998;34(14):2139-53.
14. Gill S, Loprinzi CL, Sargent DJ, Thomé SD, Alberts SR, Haller DG, et al. Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: who benefits and by how much? *Journal of clinical oncology*. 2004;22(10):1797-806.
15. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*. 1982;1041-6.
16. BORIE F, DAURES J-P, MILLAT B, FOLSCHVEILLER-BRUGGEMAN M, TRETARRE B. Surveillance des malades opérés à visée curative d'un cancer colorectal dans le département de l'Hérault: Etude médico-économique. *Gastroentérologie clinique et biologique*. 2001;25(10):881-4.
17. Schwandner O, Schiedeck T, Bruch H-P, Duchrow M, Windhoevel U, Broll R. p53 and Bcl-2 as significant predictors of recurrence and survival in rectal cancer. *European Journal of Cancer*. 2000;36(3):348-56.
18. Morales-Gutiérrez C, Vegh I, Colina F, Gómez-Cámara A, García-Carranza A, Landa JI, et al. Survival of patients with colorectal carcinoma. *Cancer*. 1999;86(9):1675-81.
19. Obrand DI, Gordon PH. Incidence and patterns of recurrence following curative resection for colorectal carcinoma. *Diseases of the colon & rectum*. 1997;40(1):15-24.
20. Gaines S, Shao C, Hyman N, Alverdy J. Gut microbiome influences on anastomotic leak and recurrence rates following colorectal cancer surgery. *BJS*. 2018;105(2).

21. Fredriksson I, Liljegren G, Arnesson L-G, Emdin S, Palm-Sjövall M, Fornander T, et al. Local recurrence in the breast after conservative surgery—a study of prognosis and prognostic factors in 391 women. *European journal of cancer*. 2002;38(14):1860-70.
22. Lê MG, Arriagada R, Spielmann M, Guinebretière JM, Rochard F. Prognostic factors for death after an isolated local recurrence in patients with early-stage breast carcinoma. *Cancer*. 2002;94(11):2813-20.
23. Ueno H, Mochizuki H, Hashiguchi Y, Hase K. Prognostic determinants of patients with lateral nodal involvement by rectal cancer. *Annals of surgery*. 2001;234(2):190.
24. Cheah P, Cao X, Eu K, Seow-Choen F. NM23-H1 immunostaining is inversely associated with tumour staging but not overall survival or disease recurrence in colorectal carcinomas. *British journal of cancer*. 1998;77(7):1164.
25. Dancourt V, Quantin C, Abrahamowicz M, Biquet C, Alioum A, Faivre J. Modeling recurrence in colorectal cancer. *Journal of clinical epidemiology*. 2004;57(3):243-51.
26. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *International journal of epidemiology*. 2015;44(1):324-33.
27. Conlon AS, Taylor JM, Sargent DJ, Yothers G. Using cure models and multiple imputation to utilize recurrence as an auxiliary variable for overall survival. *Clinical trials*. 2011;8(5):581-90.
28. Andersen PK, Keiding N. Multi-state models for event history analysis. *Statistical methods in medical research*. 2002;11(2):91-115.
29. Ferrer L, Rondeau V, Dignam J, Pickles T, Jacqmin-Gadda H, Proust-Lima C. Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. *Statistics in medicine*. 2016;35(22):3933-48.
30. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in medicine*. 2007;26(11):2389-430.
31. Spitoni C, Verduijn M, Putter H. Estimation and asymptotic theory for transition probabilities in markov renewal multi-state models. *International Journal of Biostatistics*. 2014;8(1).
32. Berkson J, Gage RP. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*. 1952;47(259):501-15.
33. Conlon A, Taylor J, Sargent DJ. Multi-state models for colon cancer recurrence and death with a cured fraction. *Statistics in medicine*. 2014;33(10):1750-66.
34. Conlon A, Taylor J, Sargent D. Multi-state models for colon cancer recurrence and death with a cured fraction. *Statistics in medicine*. 2014;33(10):1750-66.
35. Beesley LJ, Taylor JM. EM algorithms for fitting multistate cure models. *Biostatistics*. 2018.
36. Edge SB, Cancer AJCo. *AJCC cancer staging handbook: from the AJCC cancer staging manual*: Springer New York; 2010.
37. Schatzkin A, Lanza E, Corle D, Lance P, Iber F, Caan B, et al. Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. *New England Journal of Medicine*. 2000;342(16):1149-55.

38. Tartter PI. The association of perioperative blood transfusion with colorectal cancer recurrence. *Annals of surgery*. 1992;216(6):633.
39. Guyot F, Faivre J, Manfredi S, Meny B, Bonithon-Kopp C, Bouvier A. Time trends in the treatment and survival of recurrences from colorectal cancer. *Annals of oncology*. 2005;16(5):756-61.
40. Pagès F, Berger A, Camus M, Sanchez-Cabo F, Costes A, Molitor R, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *New England journal of medicine*. 2005;353(25):2654-66.
41. Liang H, Wang X-N, Wang B-G, Pan Y, Liu N, Wang D-C, et al. Prognostic factors of young patients with colon cancer after surgery. *World Journal of Gastroenterology: WJG*. 2006;12(9):1458.
42. Park YJ, Park KJ, Park J-G, Lee KU, Choe KJ, Kim J-P. Prognostic factors in 2230 Korean colorectal cancer patients: analysis of consecutively operated cases. *World journal of surgery*. 1999;23(7):721-6.
43. Ansari R, Amjadi H, Norozbeigi N, Zamani F, Mir-Nasseri S, Khaleghnejad A, et al. Survival analysis of colorectal cancer in patients underwent surgical operation in Shariati and Mehr Hospital-Tehran, in a retrospective study. *Govaresh*. 2007;12(1):7-15.
44. Vakili M, Aghakoochak A, Pirdehghan A, Shiryazdy M, Saffarmoghadam A. The Survival Rate of Patients with Colorectal Cancer in Yazd during 2001-2011. *Journal of Shahid Sadoughi University of Medical Sciences*. 2014;22(3):1187-95.
45. Chong DQ, Banbury BL, Phipps AI, Hua X, Kocarnik J, Peters U, et al. Association of family history and survival in patients with colorectal cancer: a pooled analysis of eight epidemiologic studies. *Cancer medicine*. 2018;7(5):2192-9.
46. Jayasekara H, English DR, Haydon A, Hodge AM, Lynch BM, Rosty C, et al. Associations of alcohol intake, smoking, physical activity and obesity with survival following colorectal cancer diagnosis by stage, anatomic site and tumor molecular subtype. *International journal of cancer*. 2018;142(2):238-50.
47. Meyerhardt JA, Giovannucci EL, Holmes MD, Chan AT, Chan JA, Colditz GA, et al. Physical activity and survival after colorectal cancer diagnosis. *Journal of clinical oncology*. 2006;24(22):3527-34.
48. Fernandez FG, Drebin JA, Linehan DC, Dehdashti F, Siegel BA, Strasberg SM. Five-year survival after resection of hepatic metastases from colorectal cancer in patients screened by positron emission tomography with F-18 fluorodeoxyglucose (FDG-PET). *Annals of surgery*. 2004;240(3):438.
49. Parsaee R, Fekri N, SHAHID SS, AFZAL AM, SHAARBAF EE, Esmaeily H. Prognostic factors in the survival rate of colorectal cancer patients. 2015.
50. O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. *Journal of the National Cancer Institute*. 2004;96(19):1420-5.
51. Kobayashi H, Mochizuki H, Sugihara K, Morita T, Kotake K, Teramoto T, et al. Characteristics of recurrence and surveillance tools after curative resection for colorectal cancer: a multicenter study. *Surgery*. 2007;141(1):67-75.

52. Vrieling A, Kampman E. The role of body mass index, physical activity, and diet in colorectal cancer recurrence and survival: a review of the literature–. The American journal of clinical nutrition. 2010;92(3):471-90.
53. Colorectal CCG. Adjuvant radiotherapy for rectal cancer: a systematic overview of 8,507 patients from 22 randomised trials. 2001.
54. O'Neill BD, Tait DM. Radiotherapy in colorectal cancer. Colorectal Cancer. 2007:110.
55. Elsaleh H, Joseph D, Grieu F, Zeps N, Spry N, Iacopetta B. Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer. The Lancet. 2000;355(9217):1745-50.
56. Mitry E, Fields AL, Bleiberg H, Labianca R, Portier G, Tu D, et al. Adjuvant chemotherapy after potentially curative resection of metastases from colorectal cancer: a pooled analysis of two randomized trials. Journal of clinical oncology. 2008;26(30):4906-11.
57. Group QC. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. The Lancet. 2007;370(9604):2020-9.
58. Fong Y, Fortner J, Sun RL, Brennan MF, Blumgart LH. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. Annals of surgery. 1999;230(3):309.

Tables

n	N	No. death/ recurrence	No. censor	Mean (month)		Q ₂₅ (SE)* (month)	Q ₅₀ (SE) (month)	Q ₇₅ (SE) (month)
				Estimate	SE			
ntly cured 1	198	49	149	126.35	6.675	Na	Na	53(9.5)
ed to nce	85	44	41	36.82	5.16	49(6.69)	27(6.49)	13(1.71)
ed to death	41	39	2	18.84	2.126	25(3.06)	15(1.16)	7(0.63)
nce to	44	32	12	30.58	7.731	42(14.68)	10(1.47)	7.11(12.9)

,

'able 2.

(A) The probability of not being cured (logistic Model)					
Variable	Coefficient(B)	OR (e ^B)	SE	0.95% CI(B)	P-value
BMI	0.049	1.050	0.063	(-0.07,0.17)	0.440
Age	0.441	1.554	0.343	(-0.23,1.11)	0.198
Sex (male)	-1.010	0.364	0.458	(-1.91, -0.11)	0.027*
Stage III (ref:stageII)	0.611	1.842	0.347	(-0.07,1.29)	0.078
Stage IV (ref:stageII)	0.744	2.104	0.371	(0.01,1.47)	0.045*
Radiotherapy (yes)	-1.043	0.352	0.511	(-2.04, -0.04)	0.041*
Chemotherapy (yes)	-0.494	0.610	0.917	(-2.29,1.30)	0.590
(B) Apparently cured to death (Transition 1 4)					
Variable	Coefficient(B)	HR (e ^B)	SE	0.95% CI(B)	P-value
BMI	-0.075	0.927	0.069	(-0.21,0.06)	0.278
Age	0.151	1.163	0.060	(0.03,0.27)	0.011*
Sex (male)	0.091	1.095	0.349	(-0.59,0.77)	0.795
Stage III (ref:stageII)	0.825	2.281	0.638	(-0.42,2.07)	0.195
Stage IV (ref:stageII)	0.958	2.606	0.685	(-0.38,2.30)	0.162
Radiotherapy (yes)	-0.402	0.669	0.627	(-1.63,0.83)	0.521
Chemotherapy (yes)	0.477	1.611	0.345	(-0.20,1.15)	0.167
(C) Not cured to death (Transition 2 4)					
Variable	Coefficient(B)	HR (e ^B)	SE	0.95% CI(B)	P-value
BMI	0.197	1.217	0.332	(-0.45,0.84)	0.553
Age	0.583	1.79	1.65	(-2.65,3.82)	0.724
Sex (male)	0.269	1.309	0.624	(-0.95,1.49)	0.666
Stage III (ref:stageII)	1.863	6.443	0.566	(0.75,2.97)	<0.001*
Stage IV (ref:stageII)	2.505	12.251	0.833	(0.87,4.14)	0.003*
Radiotherapy (yes)	-0.998	0.369	0.501	(-1.98, -0.02)	0.046*
(D) Not cured to recurrence (Transition 2 3)					
Variable	Coefficient(B)	HR (e ^B)	SE	0.95% CI(B)	P-value
BMI	-0.132	0.876	0.082	(-0.29,0.03)	0.106
Age	0.102	1.107	0.062	(-0.02,0.22)	0.099
Sex (male)	0.961	2.614	0.395	(0.18,1.74)	0.015*
Stage III (ref:stageII)	0.557	1.746	0.787	(-0.99,2.11)	0.479
Stage IV (ref:stageII)	0.587	1.798	0.778	(-0.94,2.11)	0.451
Radiotherapy (yes)	-0.874	0.417	0.455	(-1.77,0.02)	0.548
Chemotherapy (yes)	-0.337	0.714	1.269	(-2.82,2.15)	0.790
(E) Recurrence to death (Transition 3 4)					
Variable	Coefficient(B)	HR (e ^B)	SE	0.95% CI(B)	P-value
Stage III (ref:stageII)	1.087	2.965	0.614	(-0.11,2.29)	0.077
Stage IV (ref:stageII)	1.428	4.169	0.701	(0.05,2.80)	0.042*
Radiotherapy (yes)	-0.487	0.614	0.387	(-1.25,0.27)	0.208
Time-to-recurrence	-0.005	0.994	0.012	(-0.03,0.02)	0.676

* Significant at 0.05; ref=reference level;

Figures

Figure 1

(a) Multi-state cure model structure. The arrows show the directions of possible transitions (N=283). (b) The number of subjects in each transition in colorectal cancer data.

Figure 2

Kaplan-Meier (KM) survival curves for recurrence and death times.

Figure 3

Length of follow-up for recurrence and death events. The lines represent follow-up time for death events. a) the overall follow-up time for all patients, b) follow-up time for recurrence event, c) follow-up time for recurrence censoring, d) follow-up time for death event, and e) follow-up time for death censoring.

Figure 4

State occupancy probability (a), overall survival (b), progression free survival (c), and cumulative hazards for all transitions (d) according to Multi-state cure model.