JAMA Network | Open

ESTIMATES OF OVERALL SURVIVAL IN PATIENTS WITH CANCER RECEIVING DIFFERENT TREATMENT REGIMENS EMULATING HYPOTHETICAL TARGET TRIALS IN THE SURVEILLANCE, EPIDEMIOLOGY, AND END RESULTS (SEER)–MEDICARE LINKED DATABASE

JAMA Network Open. 2020;3(3):e200452. doi:10.1001/jamanetworkopen.2020.0452
Lucia C. Petito, PhD; Xabier Garcia-Albeniz, MD, PhD; RogerW. Logan, PhD; Nadia Howlader, PhD; Angela B. Mariotto, PhD; Issa J. Dahabreh, MD, ScD; Miguel A. Hernan, MD, DrPH

Chuenkamon Charakorn, M.D.
June 10, 2022

- When randomized clinical trials are not available at the time a clinical decision needs to be made, clinicians often must choose between making decisions based on either observational data or no human data at all.
- A cautious use of observational data has been recently endorsed by the American Society of Clinical Oncology, which states that observational studies can "inform questions that either have not been or cannot be answered by randomized trials" and "complement the evidence collected in randomized trials."
- Observational studies may be particularly important for patients who are underrepresented in trials, such as elderly patients

- Although health care databases make it possible to study cancer treatments as they are used in clinical practice, observational analyses are subject to several sources of bias.
- Some of these biases, however, can be eliminated by using observational data to emulate a (hypothetical) pragmatic target trial.
- Explicitly emulating a target trial is important because some well-known failures of observational research were the result of deviating from basic principles of study design rather than with the shortcomings inherent to observational data.

- The author demonstrate how to use the Surveillance, Epidemiology, and End Results (SEER)–Medicare database to emulate target trials that compare cancer treatments for elderly individuals.
- Their goal is to describe procedures to increase the validity of comparative effectiveness analyses using the SEER-Medicare database.
- They emulated 2 target trials for estimating the effect of the addition of a cancer drug to an existing treatment regimen on overall survival:

  (1) adjuvant fluorouracil after curative surgery in stage II colorectal cancer

  (2) addition of erlotinib to a regimen of gemcitabine for advanced pancreatic adenocarcinoma.
- They also compared the emulated randomized clinical trial estimates with those from naive observational analyses that did not attempt to emulate a target trial.

METHODS

- Study Data: SEER-Medicare Linked Database
- Emulating a Target Trial in the SEER-Medicare Linked Database
- Statistical Analysis

Study Data: SEER-Medicare Linked Database

- This is a linkage of patient demographic and cancer-related variables collected by 17 SEER cancer registries across 12 states with Medicare claim files from the Centers for Medicare & Medicaid Services.
- The SEER data are summarized in the Patient Entitlement and Diagnosis Summary File, which is linked to Medicare claims.
- This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) reporting guideline for good practices for comparative effectiveness research.
- The institutional review board at the Harvard T.H. Chan School of Public Health deemed this study exempt from review because the research involves the study of publicly available deidentified data.

Study Data: SEER-Medicare Linked Database

- To identify eligible individuals, construct a combined comorbidity score and a performance status proxy score, and also identify all other variables in this study, they used Medicare claims from the Inpatient, Outpatient, Home Health Agency, Durable Medical Equipment, Medpar, National Claims History, and Patient Entitlement and Diagnosis Summary Files.
- To identify treatment, they used the Healthcare Common Procedure Coding System and *Common Procedural Terminology* codes.
- To identify erlotinib, they determined its use by records of filled prescriptions in Medicare Part D, which started in 2007.

eAppendix 1. Codes Used to Identify Variables Used in the Analyses

Description	Code source	Codes	Analysis[a]
Cancer codes	PDESF file		
Pancreatic cancer	ICD-O-3 recode	21100	E
Non-melanoma skin cancer	ICD-O-3 recode	25020	B
Colorectal cancer	ICD-O-3	C18.0, C18.2, C18.3, C18.4, C18.5, C18.6, C18.7, C18.8, C18.9, C19.9, C20.9, C21.8	F
Adenocarcinoma	ICD-O-3	8140, 8500, 8010, 8560, 8490, 8000, 8260, 8255, 8261, 8263, 8020, 8050, 8141, 8144, 8210, 8211, 8262	E
Treatment codes			
Erlotinib	*PDE file only* NDC Brand name Generic	 50242006201, 50242006301, 50242006401 Tarceva Erlotinib HCL	E
Gemcitabine	HCPCS/CPT NDC Brand name Generic	J9201 00002750101, 00002750201, 00409018101, 00409018201, 00409018501, 00409018601, 00409018701, 00781328275, 00781328379, 16729009203, 16729011711, 25021020810, 47335015340, 47335015440, 55111068607, 55111068725, 63323010213, 63323012550 Gemzar Gemcitabine HCL	E
Fluorouracil	HCPCS/CPT NDC	J9190 00703301513, 00703301812, 00703301912, 25021021598, 25021021599, 16729027667, 16729027668, 16729027611, 16729027638, 00069016902,	F
		66758005402	
Radiotherapy	ICD-9 Revenue center HCPCS/CPT	V58.0, V66.1, V67.1, 92.2x, 92.3x, 92.4, 92.41 0330, 0333, 0339 G0174, G0251, G0339, G0340, 77401-77499, 77750 - 77899	B
Surgery (pancreatic)			E
Surgery (colorectal)	ICD-9-CM	17.3x, 45.00, 45.03, 45.4x, 45.7x, 45.8x, 46.04, 48.4xx, 48.5xx, 48.6xx	F
Staging tests			
Colonoscopy	HCPCS ICD-9	45378, 45380, 45381, 45383, 45384, 45385, G0105, G0121 45.23, 45.25	F
Abdominal CT scan	HCPCS ICD-9	74150, 74160, 74170 88.01, 88.02	F
Pelvic CT scan	HCPCS	72192, 72193, 72194	F
Emergency room visit	HCPCS	99281, 99282, 99283, 99284, 99285, 99291, 99292	B
Sentinel symptoms[c]			
Intestinal obstruction or perforation	ICD-9	560, 560.8, 560.89, 560.90, 569.83	F
Anemia	ICD-9	280, 280.0, 280.9, 281.9, 285.1, 285.2, 285.22, 285.29, 285.9	F
Abdominal distention	ICD-9	787.3	F
Change in bowel habit	ICD-9	787.99	F
Constipation	ICD-9	564.0, 564.00, 564.01, 564.02, 564.09	F

- Study Data: SEER-Medicare Linked Database
- **Emulating a Target Trial in the SEER-Medicare Linked Database**
- Statistical Analysis

Emulating a Target Trial in the SEER-Medicare Linked Database

- The first step was to fully articulate the clinical question by specifying the protocol of the (hypothetical) target trial.
- The second step was to emulate the target trial using the observational data.

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Eligibility Criteria	– Histologic diagnosis of stage II colorectal cancer (node negative) between 2008 and 2012	- Histologic diagnosis of colorectal cancer with no evidence of distant metastases between May 1994 and December 2003
	- Medicare beneficiaries ages 66 years or older To satisfy insurance and entitlement requirements, individuals must have aged into Medicare and been continuously enrolled in Parts A & B and not enrolled in an HMO for 12 months before diagnosis.	
	-Evidence of complete resection of colon or rectal cancer with "uncertain indication for chemotherapy"	-Evidence of complete resection of colorectal cancer with "uncertain indication for chemotherapy"
	- No history of prior cancer	- No definite contraindications to any of the chemotherapy regimens
	- No prior chemotherapy	- Resection margins and peritoneal washings negative for malignant cells

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Treatment Strategies	A. Initiate any dose of fluorouracil as first line treatment up to 3 months after post-surgery hospital discharge.	A. 30 doses of fluorouracil (370mg/m2 IV), given either as six 5-day courses with 4 weeks between the start of the courses or as 30 once-weekly doses. Ideally this treatment begins within 6 weeks of surgery. Patients can take high-dose folinic acid (175 mg IV), low-dose folinic acid (25 mg IV), or levamisole (50 mg) at their discretion.
	B. Do not initiate any chemotherapy within 3 months of post-surgery hospital discharge	B. Observation – do not initiate any chemotherapy

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Assignment Procedures	Participants are randomized to either treatment strategy at baseline, and are aware of the strategy they are assigned to.	Participants are randomized to a strategy by phone call to a central office. A “minimized” randomization procedure was used, ensuring balance with respect to age-group, site of cancer, stage, portal-vein infusion, preoperative radiotherapy, planned postoperative radiotherapy, and chemotherapy schedule (weekly versus not). Treatments were balanced within participating centers

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Follow-up Period	Time zero of follow-up is the first time an individual meets all eligibility criteria (when the person is assigned to one of the treatment strategies), here assumed to be the date of postsurgery discharge from the hospital. Follow-up ends at • death, • loss of insurance eligibility (loss of enrollment in Medicare Parts A or B; enrollment in an HMO), or • administrative end of follow-up (December 31, 2013 or 60 months after time zero)	Follow-up begins at randomization. Follow-up ends at the earliest of death, loss to follow-up, or administrative end of follow-up (January 2005 or 10 years after time zero).

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Outcome	All-cause mortality. Death certified by a physician, reported to Medicare and confirmed by the National Death Index within 5 years of time zero.	All-cause mortality within 10 years of time zero.
Causal contrasts of interest	Intention-to-treat effect: effect of being assigned to the strategies at baseline, regardless of whether individuals adhere to them during follow-up Per-protocol effect: effect of adhering to the strategies (as defined in the protocol) during follow-up	Intention-to-treat effect only

Protocol of the Target Trial to Study Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer and Protocol of the Existing QUASAR Trial (2007)

Protocol Component	Description of target trial	Description of existing trial
Analysis Plan	• Intention-to-treat effect estimated via comparison of 5-year risk of all-cause mortality among individuals assigned to each treatment strategy from a pooled logistic regression model adjusted for baseline covariates. • Per-protocol effect estimates are calculated from an inverse probability weighted pooled logistic regression model, adjusted for baseline : year, sex, race/ethnicity, marital status, region of the United States, metropolitan county, median household income in census tract, percentage of households under poverty line in census tract, time between diagnosis and surgery, prolonged hospitalization after surgery, preoperative radiotherapy, cancer type (colon, rectum, or both), tumor grade, and in the year before surgery, and postbaseline covariates: anemia, abdominal distention, abnormal weight loss, asthenia, change in bowel movements, constipation, diarrhea, irritable bowel syndrome, # of emergency department visits, colonoscopy, and abdominal or pelvic CT scan.	• Intention-to-treat effect estimated via comparison of 10-year risk of all-cause mortality among individuals assigned to each treatment strategy using the Kaplan-Meier method.

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
Eligibility Criteria	-Histologic diagnosis of adenocarcinoma of the pancreas between April 2007 and July 2013	- Histologic or cytologic evidence of locally advanced or metastatic adenocarcinoma of the pancreas between October 2001 and January 2003
	- Medicare beneficiaries ages 66 years or older o To satisfy insurance and entitlement requirements, individuals must have aged into Medicare and been continuously enrolled in: ▪ Parts A & B for 12 months before diagnosis ▪ Part D for 3 months before diagnosis and not enrolled in an HMO for 12 months before diagnosis.	– ECOG performance status 0, 1, or 2

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
	– No history of prior cancer	– Adequate hematologic, renal, and hepatic function
	– If diagnosis at late stage (stage IV or stage III with no surgery): o Initiation of gemcitabine (any dose) within 12 weeks of cancer diagnosis o Treatment naïve	– No prior chemotherapy except fluorouracil or gemcitabine given concurrently as a radiosensitizer.
	– If diagnosis at early stage (stage I, II, or III) with record of surgery (recurrence): o Initiation of gemcitabine (any dose) after 12 weeks post-surgery o No chemotherapy or radiation post-surgery	

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
Treatment Strategies	A. Initiate gemcitabine as first line treatment. Initiate erlotinib (any dose) within the grace period: up to 12 weeks after gemcitabine initiation.	A. Gemcitabine (1,000 mg/m2 intravenously) plus erlotinib (100 or 150 mg/d orally)
	B. Initiate gemcitabine as first line treatment within the grace period. Do not initiate erlotinib.	B. Gemcitabine (1,000 mg/m2 intravenously) plus placebo
	Under both strategies, the decision to discontinue gemcitabine or erlotinib, as well as to initiate any additional therapies, is left to the patient and physician's discretion.	Under both strategies, gemcitabine was administered on days 1, 8, 15, 22, 29, 36, and 43, followed by a 1-week rest, and on days 1, 8, and 15 in subsequent 4-week cycles. Erlotinib was taken once daily.

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
Assignment Procedures	Participants are randomized to either treatment strategy at baseline, and are aware of the strategy they are assigned to.	Patients are randomized to either treatment strategy at baseline, stratified by center, performance status (ECOG 0 versus 1-2), and stage (locally advanced versus metastatic). Patients and physicians are blinded to treatment assignment.
Follow-up Period	Time zero of follow-up is the first time an individual meets all eligibility criteria (when the person is assigned to one of the treatment strategies). Follow-up ends at the earliest of death, loss to follow-up (loss of enrollment in Medicare Parts A, B, or D; enrollment in an HMO), or administrative end of follow-up (December 31, 2013 or 18 months after time zero)	Follow-up begins at randomization. Follow-up ends at the earliest of death, loss to follow-up, or administrative end of follow-up (September 2004 or 24 months after time zero).

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
Outcome	All-cause mortality. Death certified by a physician, reported to Medicare and confirmed by the National Death Index within 18 months of time zero.	All-cause mortality within 24 months of baseline.
Causal contrasts of interest	Intention-to-treat effect: effect of being assigned to the strategies at baseline, regardless of whether individuals adhere to them during follow-up	Intention-to-treat effect only.
	Per-protocol effect: effect of adhering to the strategies (as defined in the protocol) during follow-up	

Protocol of the Target Trial to Study the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer and Protocol of the Existing Trial (Moore et al. 2007)

Protocol Component	Target trial	Description of existing trial
Analysis Plan	Intention-to-treat effect estimated via comparison of 18-month risk of all-cause mortality among individuals assigned to each treatment strategy from a pooled logistic regression model adjusted for baseline covariates.	Intention-to-treat effect estimated via comparison of 24-month risk of all-cause mortality among individuals assigned to each treatment strategy using the Kaplan-Meier method.
	Per-protocol effect estimates are calculated from an inverse probability weighted pooled logistic regression model, adjusted for baseline and postbaseline covariates: number of emergency department visits, Charlson Comorbidity Index, cholangitis, and pneumonia (each defined using claims from the previous week).	

Eligibility Criteria

- In addition to the eligibility criteria of the existing trials,
- + 1 year of continuous enrollment in Medicare Parts A and B, without being enrolled in a health maintenance organization,
- and having less than 3 months between diagnosis and initiation of first-line treatment.

Treatment Strategies and Assignment

- All components of the treatment strategy needed to be in place within a prespecified period (the grace period) after baseline.
- Because the existing trials did not explicitly specify the length of the grace period, they chose a 3-month period based on the clinical management.

Treatment Strategies and Assignment

- Adjusted for factors and demographic characteristics, geographic characteristics, and comorbidities (as identified by previously validated procedure and diagnosis codes).
- Target trials must be pragmatic trials in which treatment assignment is not blinded because observational data typically cannot be used to emulate double-blind trials.

Start and End of Follow-up

- The start of follow-up (baseline or time zero) for each individual was the time of first eligibility:
  - the date of the first gemcitabine claim in the emulation of the erlotinib trial, and
  - the date of hospital discharge after surgery in the emulation of the fluorouracil trial.
- The end of follow-up was death, loss of insurance eligibility follow-up (loss of enrollment in Parts A, B, or D; enrollment in a health maintenance organization), or administrative end of follow-up (December 31, 2013; 60 months for the fluorouracil trial and 72 weeks for the erlotinib trial), whichever was earliest.

Outcome

- All-cause mortality by using the Medicare date of death, which is more up to date and is validated by the National Death Index.
- The SEER-Medicare data do not include enough information to reliably study progression-free survival.

Causal Contrast

- The causal contrast of interest was the per-protocol effect, that is, the effect that would have been observed if all trial participants had adhered to the treatment strategies specified in the protocol.
- Because treatment strategies simply required the initiation of either fluorouracil or erlotinib within the grace period (regardless of future continuation), the causal contrast was similar to the intention-to-treat effect in a trial in which all individuals initiated their assigned treatment during the grace period, even if some of them later discontinued treatment.

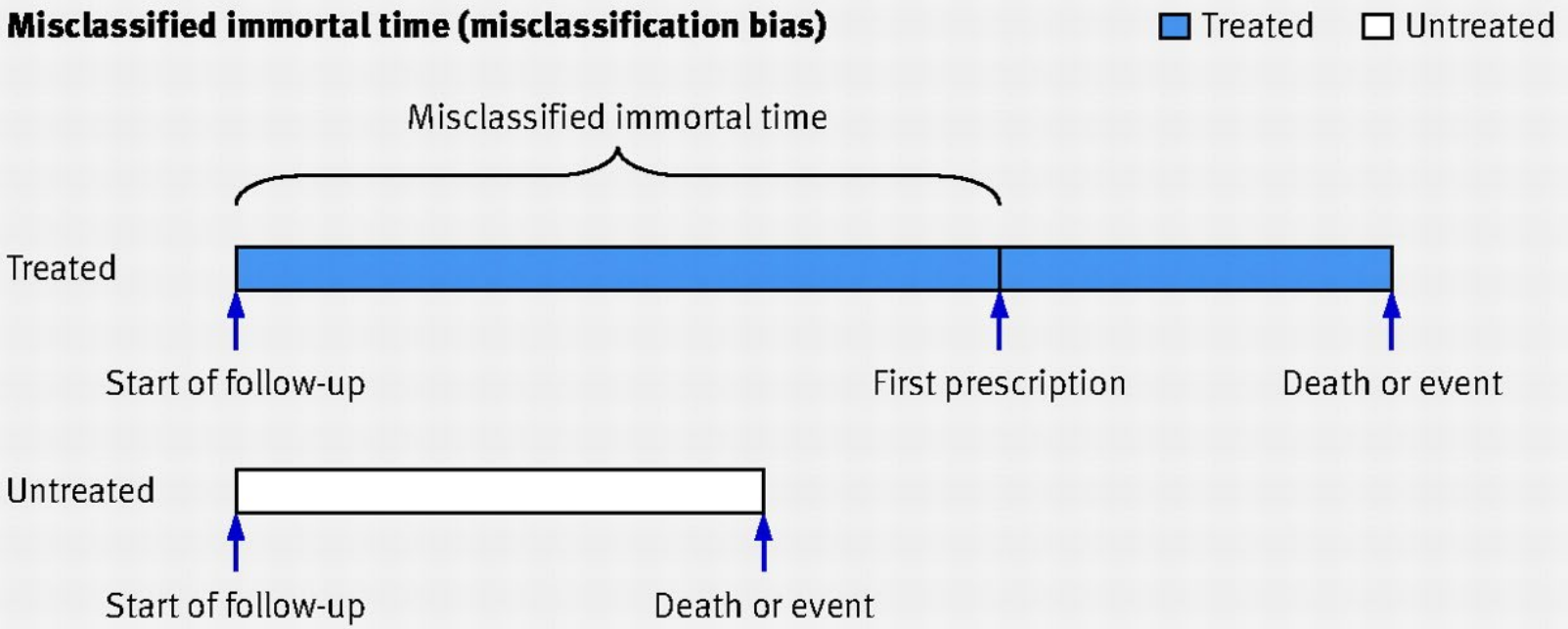
- Study Data: SEER-Medicare Linked Database
- Emulating a Target Trial in the SEER-Medicare Linked Database
- Statistical Analysis

Statistical Analysis

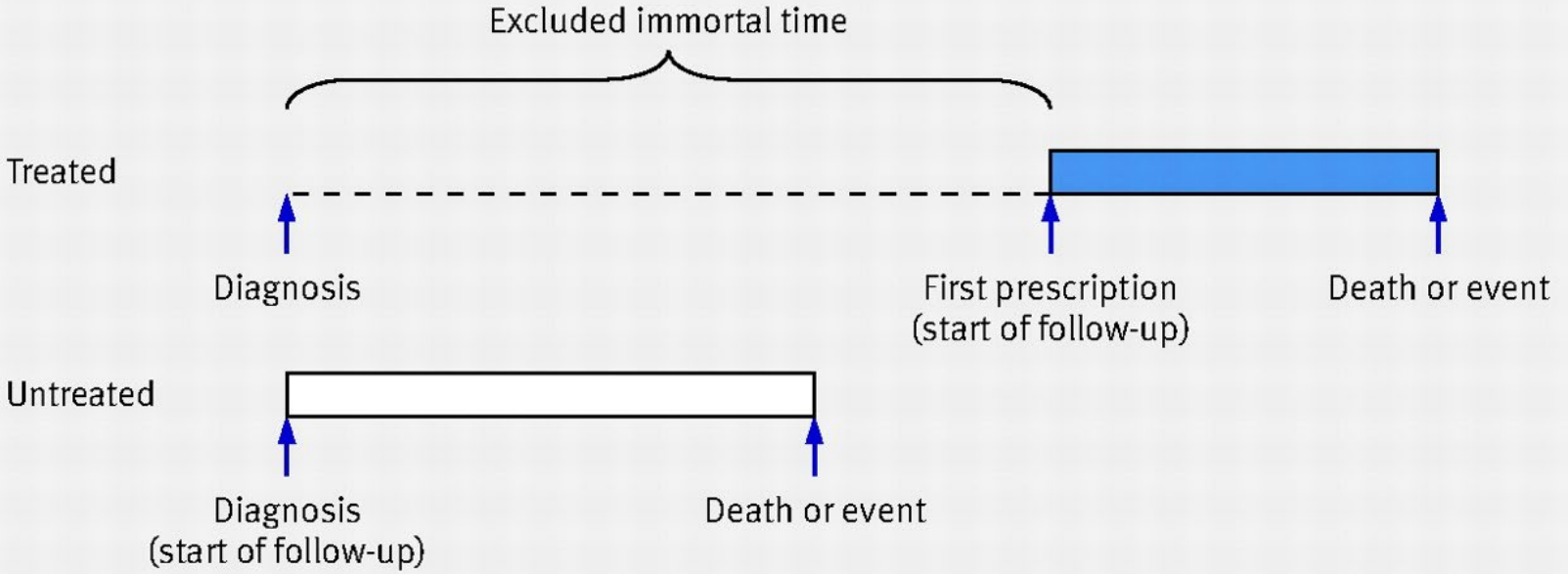
- There are 3 steps of the per-protocol analysis that they implemented in the emulation of the target trials

1. cloning (to avoid immortal time bias),
2. censoring at deviation from protocol (to ensure adherence), and
3. inverse probability weighting (to adjust for selection bias).

Misclassified immortal time (misclassification bias)



Excluded immortal time (selection bias)



Cloning and censoring

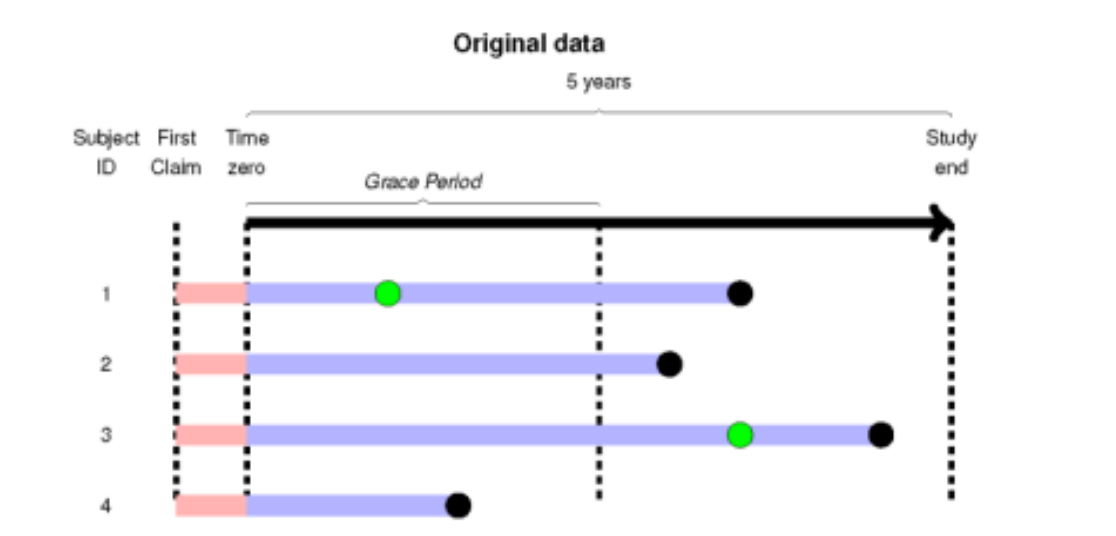
- Duplicating the original data for each individual and assigning each clone or replicate (two per individual) to either strategy A or B.

Protocol Component	Description of target trial
Treatment Strategies	A. Initiate any dose of fluorouracil as first line treatment up to 3 months after post-surgery hospital discharge.
	B. Do not initiate any chemotherapy within 3 months of post-surgery hospital discharge

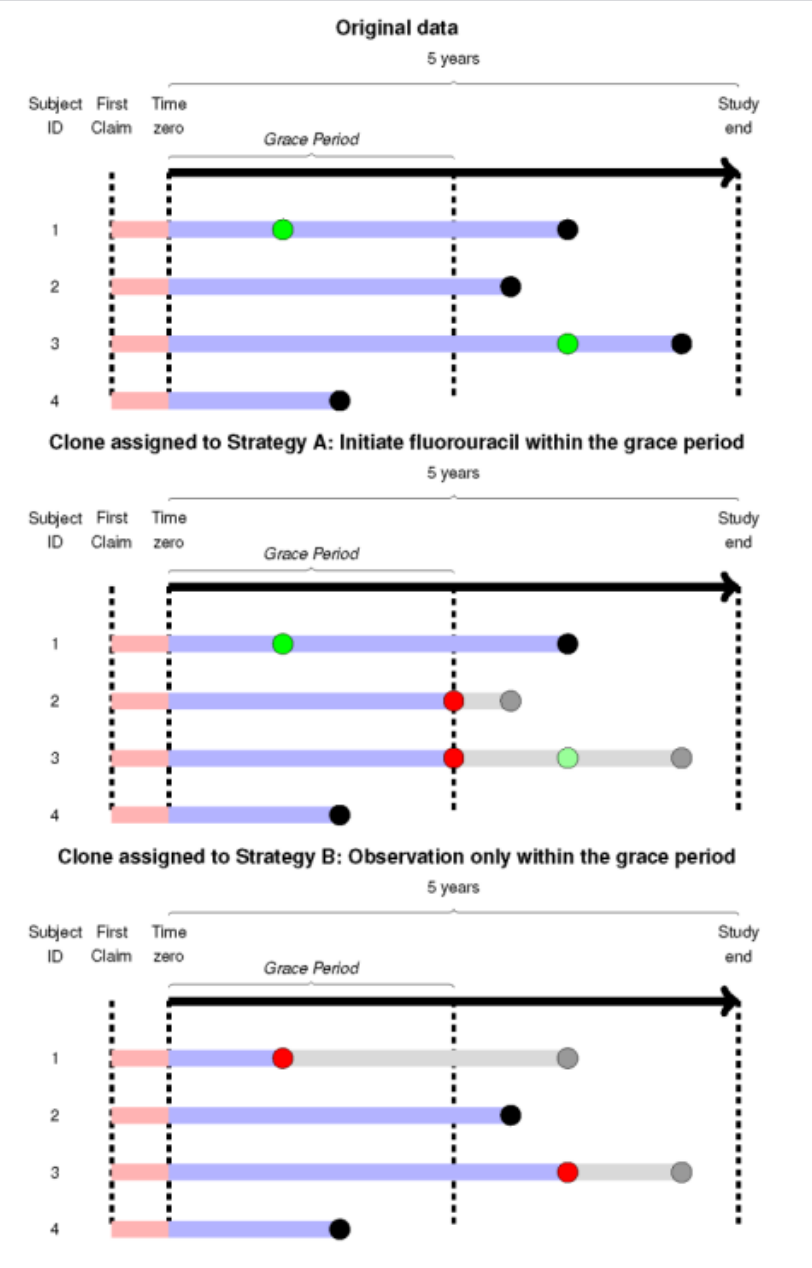
- Created a new variable to indicate which treatment strategy the replicate was assigned to – this variable is "treated" in the model summaries.
- Replicates were then censored when they deviated from the protocol of the treatment strategy.

Cloning and censoring

- Each individual's treatment strategy was completely determined by the end of the grace period, so at most only one replicate from each individual still contributes person-time to the analysis by the end of the grace period.

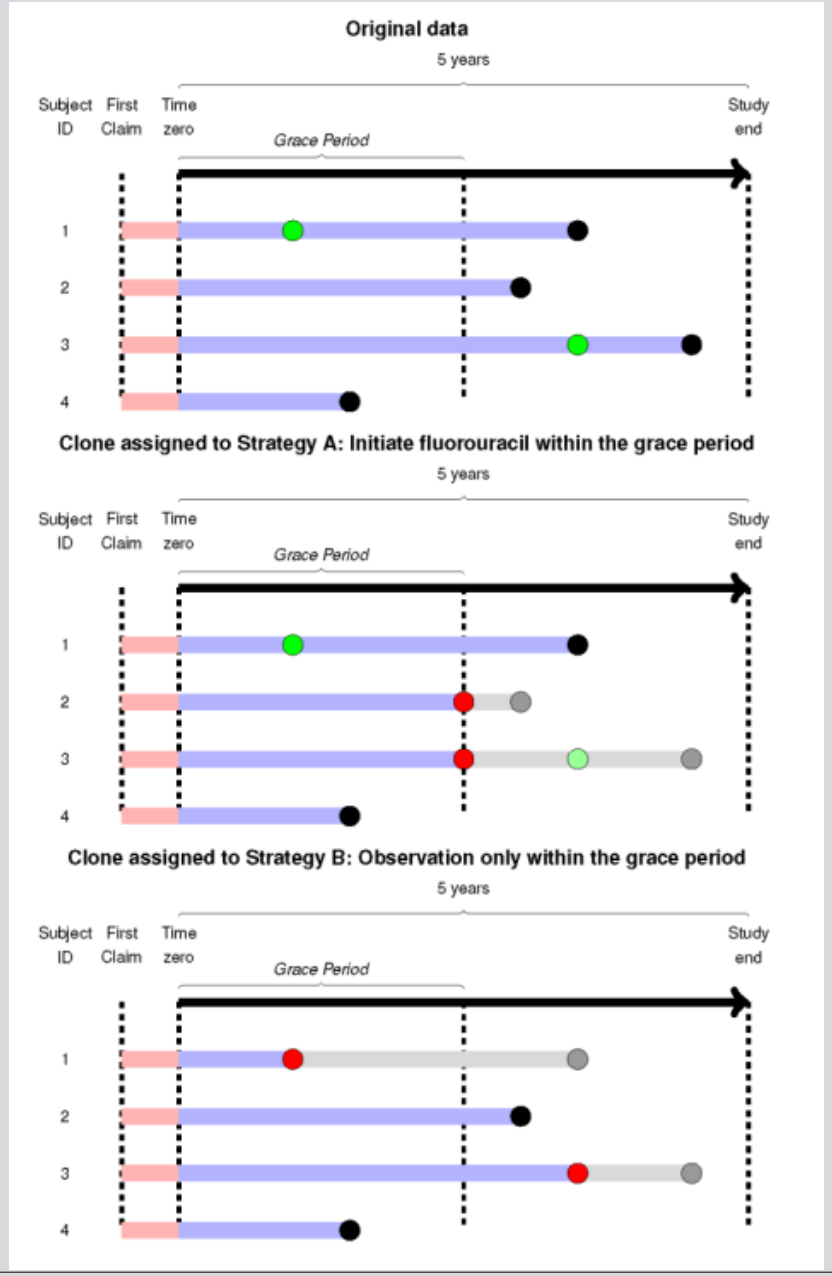


4 types of individuals would be treated in this setting



- Subject 1 initiates fluorouracil (green circle) during the grace period, and then dies or is censored after the grace period ends (black circle). Their complete person-time contributes to the Strategy A clone. However, only their person-time before initiating fluorouracil contributes to the Strategy B clone, resulting in "artificial" censoring (red circle) at the time of fluorouracil initiation.

- Subject 2 does not initiate fluorouracil during the grace period, and then dies or is censored after the grace period ends. Their complete person-time contributes to the Strategy B clone. However, only their person-time during the grace period contributes to the Strategy A clone, as they have not initiated fluorouracil by the end of the grace period. The Strategy A clone here is "artificially" censored at the end of the grace period.



- **Subject 3** initiates fluorouracil after the end of the grace period, and then dies or is censored. Only their person-time during the grace period contributes to the Strategy A clone, as they have not initiated fluorouracil. Their Strategy B clone only includes the person-time contributed before they initiate fluorouracil – they are "artificially" censored at that time.

- **Subject 4** dies or is censored during the grace period. This censoring includes initiating other chemotherapy for Strategy B. For Strategy A, it also includes initiating other chemotherapy before. Individuals like Subject 4 contribute their complete person-time to both clones.

Section C.1. Model coefficients for hazard ratio estimates

Note: In all reported models, t represents the linear term for time, and t*, t**, and t*** represent the estimates for the 1st, 2nd, and 3rd spline basis terms (knots prespecified at 3, 16, 30, 44, and 57 months).

From the unadjusted model (without product term)

	Estimate	Std. Error
Intercept	-4.339399	0.048883
t	-0.051215	0.006478
t*	0.217710	0.046279
t**	-0.511636	0.142585
t***	0.459383	0.209772
treated	-0.000395	0.058004

From the adjusted model (without product term)

	Estimate	Std. Error
Intercept	-4.5559	0.1451
t	-0.0475	0.0066
t*	0.2223	0.0467
t**	-0.5233	0.1436
t***	0.4680	0.2110
treated	0.0183	0.0587
Diagnosed in 2010-2011	0.0095	0.0456
Male	-0.1071	0.0436
Non-Hispanic Black	-0.2090	0.0825
Grade: poor	-0.1545	0.0489
Anemia (b)	0.1177	0.0499
Abdominal Distension (b)	0.0835	0.0800
Abnormal weight loss (b)	0.2528	0.0509
Asthenia (b)	0.0998	0.0424
Change in bowel movement (b)	0.0152	0.0674
Constipation (b)	0.0623	0.0514
Diarrhea (b)	0.0415	0.0546
Irritable bowel syndrome (b)	-0.1601	0.1286
At least 1 ED visit (b)	0.4024	0.0466
Colonoscopy (b)	-0.4659	0.0458
Abdominal or pelvic CT scan (b)	0.0028	0.0497
Charlson (b)	0.1530	0.0094

Section C.2. Model coefficients for risk estimates

From the adjusted weighted model with product terms between time and treatment.

	Estimate	Std. Error
Intercept	-4.5598	0.1464
t	-0.0435	0.0070
t*	0.2072	0.0487
t**	-0.4985	0.1487
t***	0.4772	0.2163
treated	0.2100	0.0812
treated x t	-0.1022	0.0232
treated x t*	0.5457	0.1830
treated x t**	-1.1769	0.5660
treated x t***	0.5105	0.8337
Diagnosed in 2010-2011	0.0176	0.0457
Male	-0.1172	0.0434
NH Black	-0.1486	0.0801
Hispanic/Other	-0.1679	0.0729
Married	-0.3370	0.0435
Region: NE	-0.0199	0.0472
Region: S	0.0745	0.0810
Region: MW	0.0027	0.0626
Urban center	0.0310	0.0569
Median HHI	0.0000	0.0000

Section C.3. Model coefficients for numerator and denominator of weights

From the model for the numerator of the weights (adjusted for baseline covariates only).

	Estimate	Std. Error
Intercept	-2.8139	0.0920
t	0.0782	0.0042
t*	-0.3552	0.0266
t**	0.8955	0.0791
t***	-0.9103	0.1117
Diagnosed in 2010-2011	0.0633	0.0282
Male	-0.3133	0.0249
NH Black	0.3962	0.0488
Hispanic/Other	0.4919	0.0397
Married	0.4515	0.0259
Region: NE	0.7009	0.0283
Region: S	-0.1204	0.0591
Region: MW	0.5761	0.0351
Urban center	0.3388	0.0306
Median HHI	0.0000	0.0000
% Poverty	-0.0219	0.0019
Time between DX and Fluoro: 1-30D	-0.3379	0.0276
Time between DX and Fluoro: 31-60D	-0.5339	0.0414
Time between DX and Fluoro: 61-90D	-0.5279	0.0751
Prolonged post-surgery hospitalization	-0.2634	0.0483
Pre-operative radiation	-0.8624	0.0599
Rectal cancer	1.3841	0.0392
Both Colon and Rectal cancer	0.5576	0.0438
Grade: poor	-0.3297	0.0283
Anemia (b)	-0.0644	0.0256

From the model for the denominator of the weights (adjusted for baseline and time-varying covariates).

	Estimate	Std. Error
Intercept	-2.9602	0.0930
t	0.0872	0.0042
t*	-0.3844	0.0267
t**	0.9628	0.0796
t***	-0.9690	0.1122
Diagnosed in 2010-2011	0.0686	0.0284
Male	-0.3332	0.0250
NH Black	0.4151	0.0490
Hispanic/Other	0.4907	0.0398
Married	0.4562	0.0260
Region: NE	0.6751	0.0284
Region: S	-0.1308	0.0592
Region: MW	0.5637	0.0352
Urban center	0.3544	0.0307
Median HHI	0.0000	0.0000
% Poverty	-0.0222	0.0019
Time between DX and Fluoro: 1-30D	-0.3246	0.0278
Time between DX and Fluoro: 31-60D	-0.5218	0.0416
Time between DX and Fluoro: 61-90D	-0.4905	0.0754
Prolonged post-surgery hospitalization	-0.3158	0.0486
Pre-operative radiation	-0.8819	0.0601
Rectal cancer	1.3701	0.0394
Both Colon and Rectal cancer	0.5422	0.0441

Statistical Analysis

- Used inverse probability weighting to adjust for the potential selection bias introduced by censoring.
  - Weighting process
    - Estimated subject-specific time-varying stabilized inverse-probability (IP) weights, which create a pseudopopulation where time-varying prognostic factors are independent of future treatment.

Ak is an indicator for use of fluorouracil at time k (1: ever initiated, 0: never initiated), L0 is the vector of baseline prognostic factors, and Lk is the vector of time-varying prognostic factors at time k. The overbar denotes the history of a variable since start of follow-up. The stabilized IP weights can then be written as

$$SW_t = \prod_{k=0}^{t} \frac{f(A_k|\bar{A}_{k-1}, L_0)}{f(A_k|\bar{A}_{k-1}, \bar{L}_k)},$$

  - Weighted outcome model-using a pooled logistic regression model

$$logit(\Pr(Y_{t+1} = 1 \,|Y_t = 0, A, L_0) = \beta_0 + \vec{\beta_1} f(t) + \beta_2 A + \vec{\beta_3} f(t) A + \vec{\beta_4} L_0$$

Statistical Analysis

- Used the inverse probability weighting to adjust for the potential selection bias introduced by censoring.
  - Weighting process
  - Weighted outcome model
- The predicted values form this IP weighted model are used to compute the cumulative incidence of mortality.

Statistical Analysis

- Fit an inverse probability weighted discrete-time hazard model by pooled logistic regression, with death as the response and the following regressors: the indicator for the assigned treatment strategy, a function of time of follow-up (restricted cubic spline with knots at 3, 16, 30, 44, and 57 months for the fluorouracil emulation; quadratic for the erlotinib emulation), product terms for treatment strategy and time, and baseline covariates.

- To calculate a single summary (average) hazard ratio as reported in trials, use the predicted values from the weighted model to simulate the trajectory of each original individual under complete follow-up (10 simulations per individuals were used to reduce simulation uncertainty).

Statistical Analysis

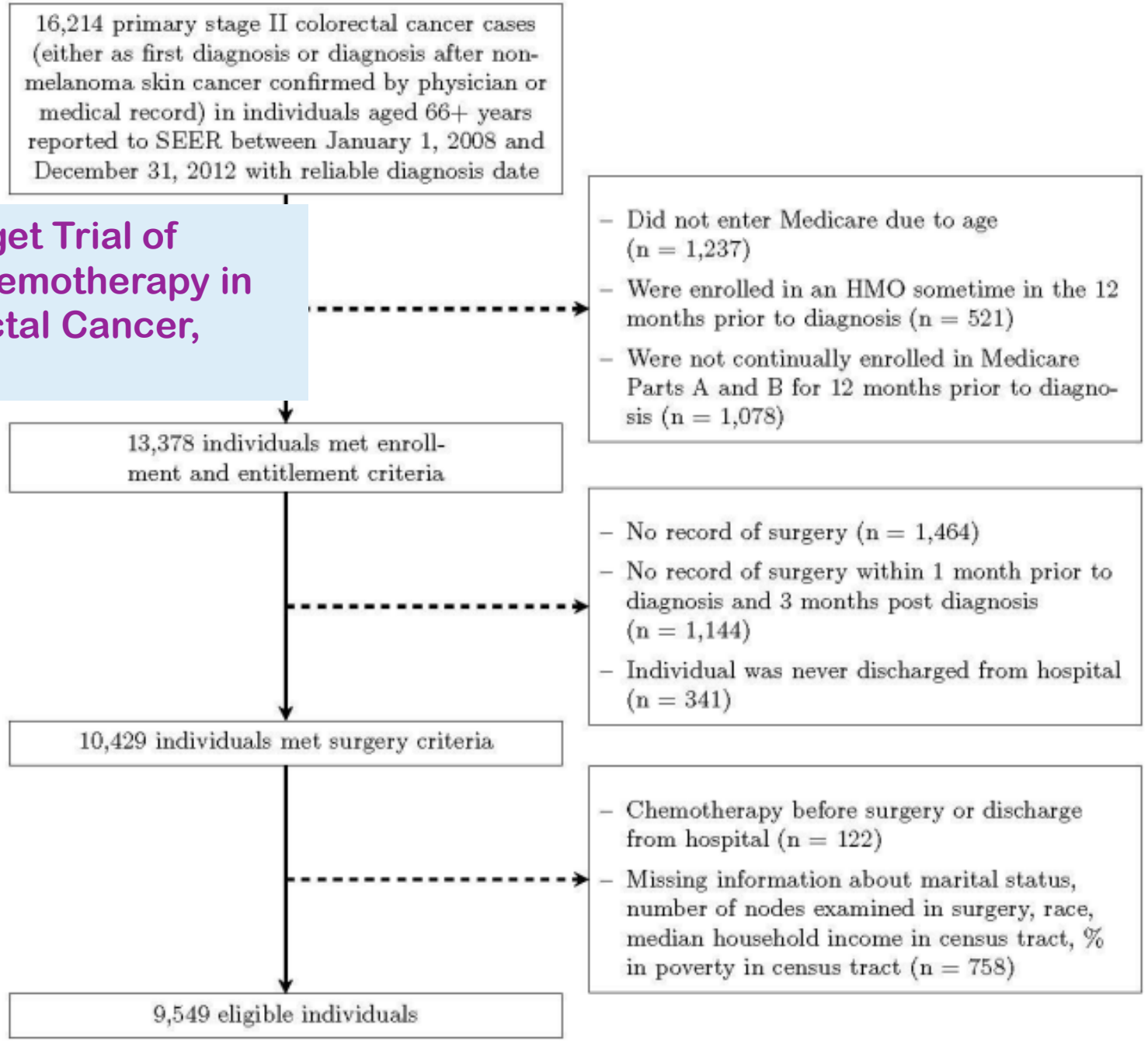
- Used the estimated probability of death for a random Bernoulli flip to determine if an individual was alive at a given time. The first instance of death was deemed to be end of follow-up.
- Then fit an unadjusted Cox proportional hazards model in the simulated data, using the predicted time of end of follow-up as the outcome, and treatment assignment (as determined by the end of the grace period) as the sole predictor.
- To obtain standardized, treatment-specific risks for time points between baseline and the end of follow-up, standardized the model-derived estimated values to the joint distribution of the baseline covariates.
- To estimate the mean HR commonly reported in randomized clinical trials, fit a Cox proportional hazards model with treatment strategy as the only covariate.

Statistical Analysis

- Conducted several sensitivity analyses to test the robustness of the estimates to different choices of functional forms and grace periods (results were similar)
- Computed 95% CIs via a nonparametric bootstrap based on 500 resamplings.
- Used SAS, version 9.4 (SAS Institute Inc), for data processing, and R, version 3.4.4 (The R Foundation), for data analyses.
- The SEER-Medicare data were from those released in 2015 (containing content through December 31, 2013), and their analyses were conducted from January 2018 to March 2019.

RESULTS

Flowchart of Eligibility for a Target Trial of Adjuvant Fluorouracil-Based Chemotherapy in Individuals With Stage II Colorectal Cancer, SEER-Medicare 2008- 2013



eTable 4. Comparison of Individuals in the Existing QUASAR Trial (2007) in the Emulation of the Fluorouracil Target Trial Using SEER-Medicare 2008-2013

	SEER-Medicare Eligible Sample (n = 9,549)		QUASAR Participants (n = 3,239)	
	n	%	n	%
Site				
Colon	8,565	89.7	2291	70.7
Rectum (or both)	984	10.3	948	29.3
Sex				
Male	5,524	57.8	1979	61.1
Female	4,025	42.2	1260	38.9
Age				
<59	---	---	1225	37.8
60-69	1132	11.9	1351	41.7
70+	8417	88.1	663	20.5
Median age (IQR)	79		63	
IQR	73 to 84		56 to 68	
Other adjuvant therapy				
Pre-operative radiotherapy	782	8.2	203	6.3

IQR: Inner quartile range

--- reported when cell size is ≤10, as per the SEER-Medicare Data Use Agreement

Fluorouracil for Stage II Colorectal Cancer

- Of the 9549 eligible individuals included in the present analysis (23 447 person-years of follow-up), 204 initiated fluorouracil-based chemotherapy within 3 months of their hospital discharge.
- Fluorouracil initiation was more likely for people who were younger, married, and had a T4 tumor category at diagnosis and rectum involvement.
- Individuals were less likely to initiate fluorouracil-based chemotherapy if they had a prolonged hospitalization after surgery, preoperative radiotherapy, and, in the year before diagnosis, anemia and asthenia.
- By the end of the grace period, 185 individuals remained in the fluorouracil group and 6150 in the observation group

- 2148 deaths
- adjusted estimated 5-year survival was 66.6% for fluorouracil and 62.7% for no fluorouracil
- the 5-year risk difference was −3.8% (95% CI, −14.8% to 12.6%)
- The mortality HR for fluorouracil vs no fluorouracil was 1.00 (95% CI, 0.89-1.12) without any adjustment, 1.01 (95% CI, 0.97-1.05) after adjustment for baseline covariates, and 0.95 (95% CI, 0.85-1.04) after adjustment for baseline and postbaseline covariates.

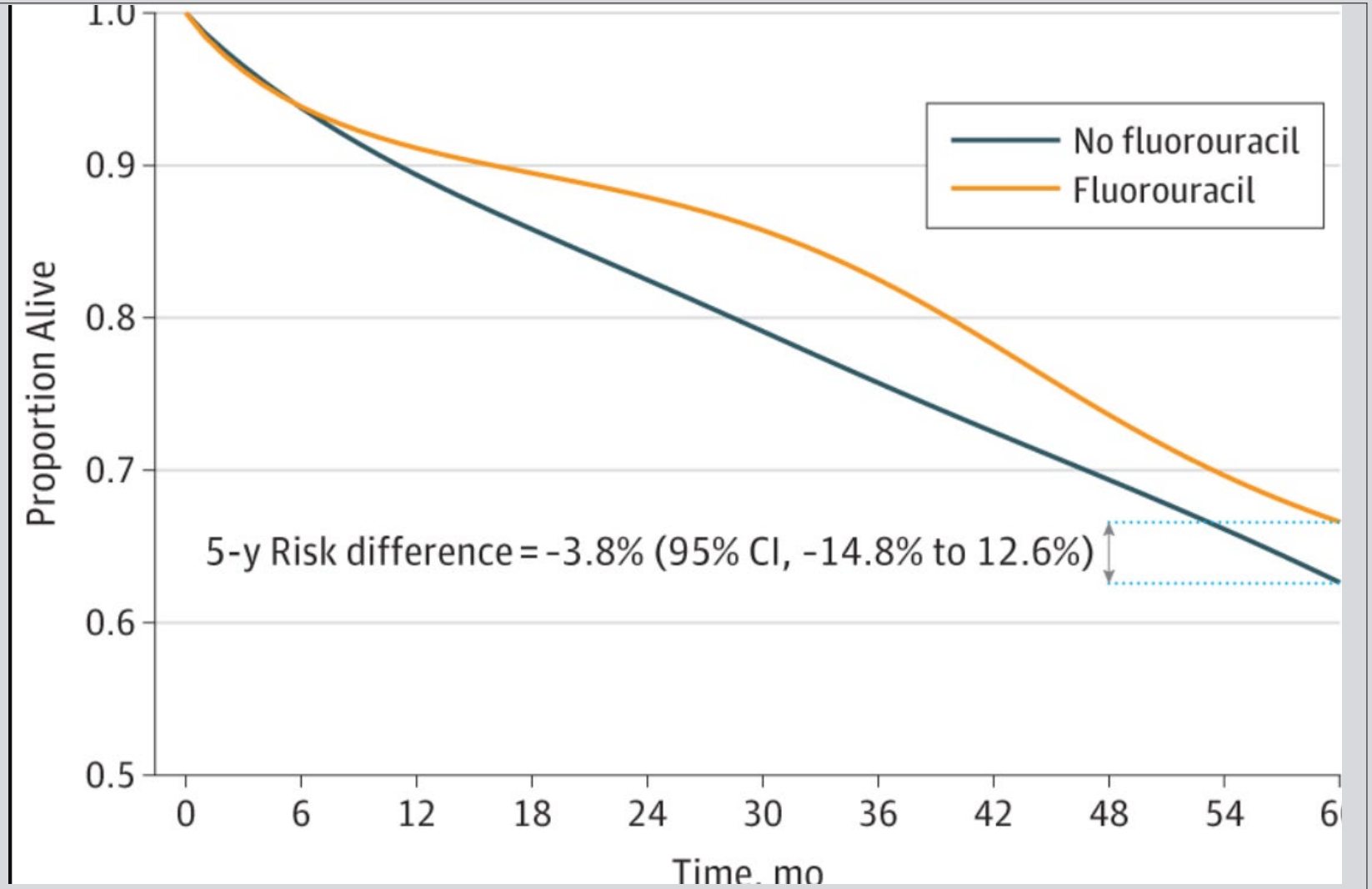
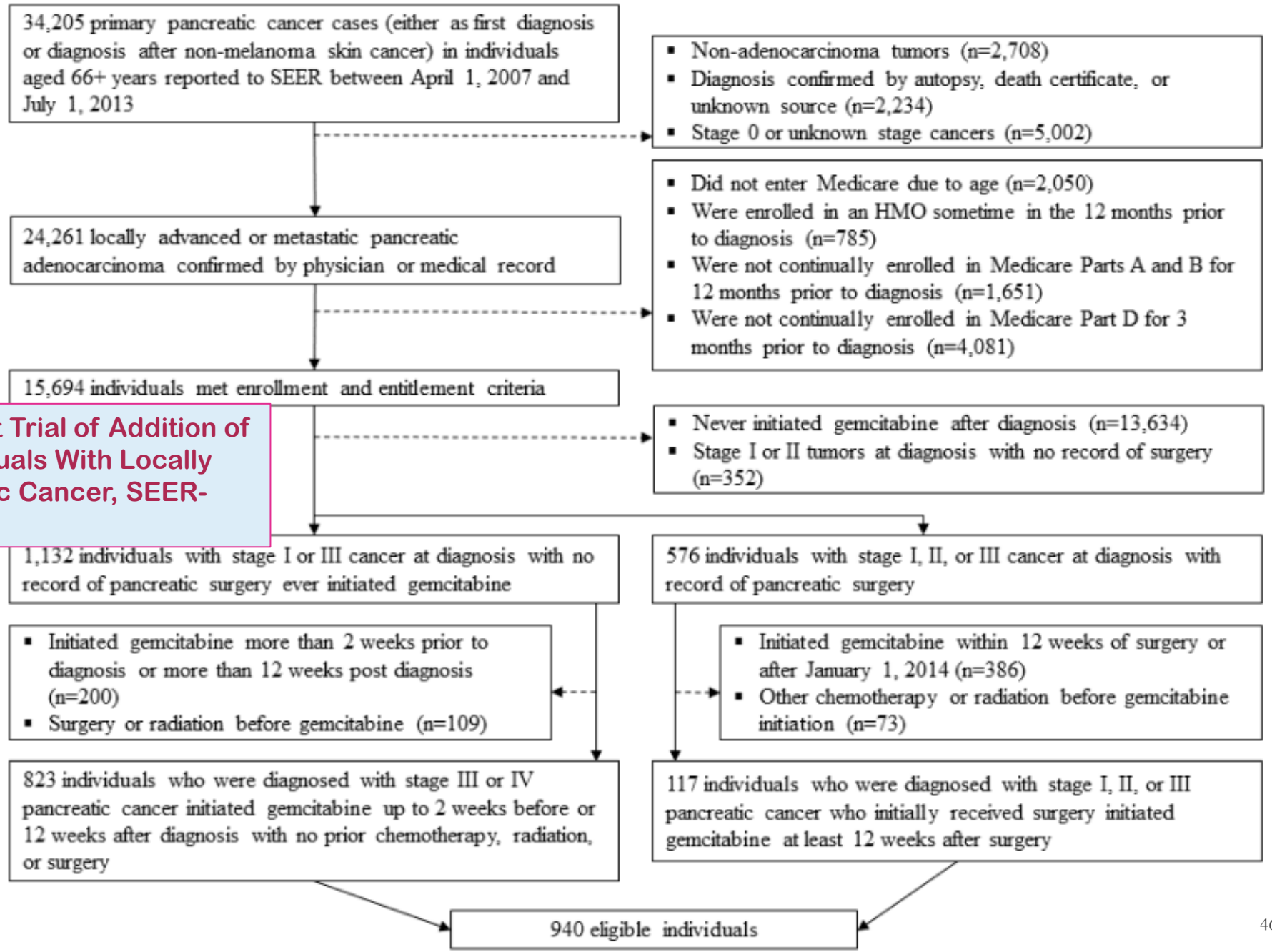


Figure 1. Survival Curves From the Emulated Target Trial of Adjuvant Fluorouracil-Based Chemotherapy in Stage II Colorectal Cancer Using SEER–Medicare Data From 2008 to 2013

Flowchart of Eligibility for a Target Trial of Addition of Erlotinib to Gemcitabine in Individuals With Locally Advanced or Metastatic Pancreatic Cancer, SEER-Medicare 2007-2013



eTable 6. Comparison of Individuals in the Existing Trial (Moore et al. 2007) and in the Emulation of the Erlotinib Target Trial Using SEER-Medicare 2007-2013

	SEER-Medicare Eligible Sample		Moore et al. (2007) Participants	
	N = 940		N = 569	
	n	%	n	%
Sex				
Female	393	41.8	271	47.6
Male	547	58.2	298	52.4
Age, years				
Median	74.0		63.9	
Range	66.0-93.0		36.1-92.4	
ECOG performance status[a]				
0-2	858	91.3	569	100.0
3+	82	8.7	0	0.0
Extent of disease				
Locally advanced	240	25.5	138	24.3
Distant metastases	700	74.5	431	75.7
Prior therapy[b]				
Radiotherapy	---	---	47	8.3
Chemotherapy	53	5.6	45	7.9
Prior surgical resection of primary tumor	117	12.4	48	8.4

[a]ECOG performance status < 3 was an eligibility criteria for Moore et al. (2007)

[b]In SEER-Medicare eligible sample, prior therapy is only possible in individuals with prior surgical resection of primary tumor

--- reported when cell size is ≤10, as per the SEER-Medicare Data Use Agreement

Erlotinib for Advanced Pancreatic Cancer

- Of the 940 eligible individuals (412 person-years of follow-up), 62 initiated erlotinib within 12 weeks of their initial gemcitabine dose.
- Erlotinib initiation was more likely for individuals with stage IV cancer and for those with comorbidities, and less likely for those with a poor performance status and older age.
- By the end of the grace period, 44 individuals remained with the gemcitabine plus erlotinib strategy and 480 remained with the gemcitabine alone strategy.

- 659 deaths
- The adjusted estimated 1-year survival was 15.6% for gemcitabine plus erlotinib and 20.4% in gemcitabine alone; the risk difference was 4.7% (95% CI, −9.4% to 18.0%).
- The all-cause mortality HR for gemcitabine plus erlotinib vs gemcitabine alone was 1.08 (95% CI, 0.93-1.25) without any adjustment, 1.03 (95% CI, 0.96-1.09) after adjustment for baseline covariates, and 1.04 (95% CI, 0.86-1.42) after adjustment for baseline and postbaseline covariates.

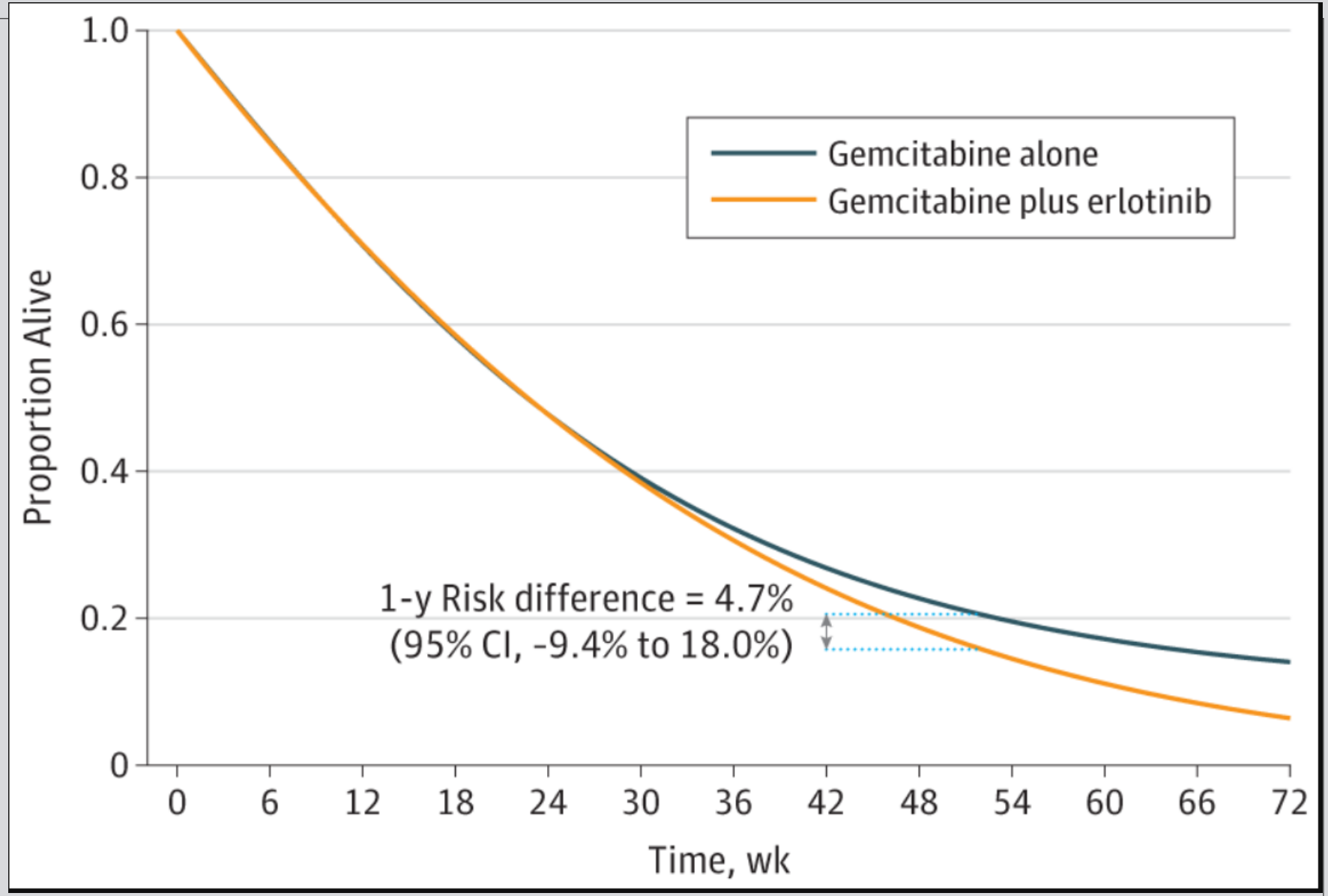


Figure 2. Survival Curves From the Emulated Target Trial of the Addition of Erlotinib to a Regimen of Gemcitabine in Locally Advanced or Metastatic Pancreatic Cancer Using SEER –Medicare Data From 2007 to 2013

Naive Analyses

- For comparison purposes
- The mortality HR estimate was 1.14 (95% CI, 0.95-1.36). In a similar analysis for the erlotinib comparison, the mortality HR estimate was 0.68 (95% CI, 0.54-0.87).

- The mortality HR for fluorouracil vs no fluorouracil was
  - 1.00 (95% CI, 0.89-1.12) without any adjustment,
  - 1.01 (95% CI, 0.97-1.05) after adjustment for baseline covariates, and
  - 0.95 (95% CI, 0.85-1.04) after adjustment for baseline and postbaseline covariates.

- The all-cause mortality HR for gemcitabine plus erlotinib vs gemcitabine alone was
  - 1.08 (95% CI, 0.93-1.25) without any adjustment,
  - 1.03 (95% CI, 0.96-1.09) after adjustment for baseline covariates, and
  - 1.04 (95% CI, 0.86-1.42) after adjustment for baseline and postbaseline covariates.

The age-specific estimates from 2 published randomized clinical trials: hazard ratios (HRs) of 1.02 (95% CI, 0.70-1.48) for fluorouracil and 0.96 (95% CI, 0.74-1.24) for erlotinib.

DISCUSSION

- These target trials were based on 2 existing randomized clinical trials: 1 pragmatic trial (fluorouracil) and 1 placebo-controlled, double-blind trial (erlotinib).
- They emulated these target trials using observational data from individuals 65 years of age or older whose data were in the SEER-Medicare database.
- Their observational estimates were consistent with the null estimates reported by the existing trials when they explicitly emulated the target trials but not when they conducted naive analyses that did not attempt to emulate a target trial.

- They were able to obtain that information from the linkage of detailed cancer information from the SEER registries with administrative data from Medicare claims.
- Detailed timing information for comorbidity onset and treatment administration exists solely in Medicare; thus, using SEER only would preclude them from emulating target trials. The use of Medicare data enabled them to reduce some errors present in SEER alone (eg, in the radiation variable) but does not guarantee perfect coding reliability of all variables (eg, rare histologic results).

- Similar to all observational analyses, theirs assume that the lack of randomized treatment assignment can be approximately replaced by adjustment for the measured confounders.
- After the target trial was specified and the emulation procedures followed, they found no indication of strong confounding by measured variables.
- In fact, adjustment for multiple prognostic factors only modestly changed the estimates in both emulations. Given that they adjusted for some of the most important indications for the studied treatments, it is likely that confounding was not the most important validity threat in the examples.

- Leaving aside a lack of randomization, a major threat to the validity of observational comparative effectiveness analyses is a failure to emulate components of a target trial other than randomization, which may introduce biases such as selection bias and immortal time bias. Explicit emulation of a target trial helps eliminate these problems.
- For the same reasons, prior comparisons of randomized clinical trials and observational studies based on SEER-Medicare or other databases are difficult to interpret because, in the absence of an explicit emulation of a target trial, any discrepancies may be due to incorrect analysis rather than to inherent limitations of the observational data.

Limitations

- The use of large observational databases does not guarantee transportable or precise estimates.
- The requirements of target trial emulation (eg, requiring individuals to participate in Medicare Part D, which was necessary in the erlotinib analysis) may limit the generalizability.
- Small sample sizes despite the apparently large amount of data in the SEER-Medicare database.

Conclusion

- They emulated 2 target trials by using the SEER-Medicare database and found that observational estimates from this database are consistent with those from randomized clinical trials in elderly populations, but only when an appropriate causal inference approach is implemented.
- As the population ages, understanding how cancer treatments affect elderly individuals will have profound implications for clinical decision-making.
- Because elderly individuals tend to be underrepresented in randomized clinical trials, observational health care databases can be used to obtain more precise estimates in that segment of the population or to guide the design of future randomized clinical trials by helping to rule out clearly ineffective or harmful treatment strategies.

THANK YOU VERY MUCH FOR THE ATTENTION

Have a nice weekend!!