IDENTIFYING MORTALITY IN CANCER PATIENTS USING CLAIMS DATA

Wei-Jhih Wang[1,3], MS, Craig S. Meyer[1], PhD, Ning Wu[4], MD, PhD,Nicolas Sommer[2],PharmD, MSc, Daniel Sheinson[1], PhD

[1] Technical Institute, Evidence for Access, USMA, Genentech, Inc., South San Francisco, CA
[2] BioOncology, USMA, Genentech, Inc., South San Francisco, CA
[3] The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, University of Washington, Seattle, WA
[4] Health Economic and Outcome Research, Medical Affairs, Regeneron Pharmaceuticals, Inc., Tarrytown, NY

INTRODUCTION

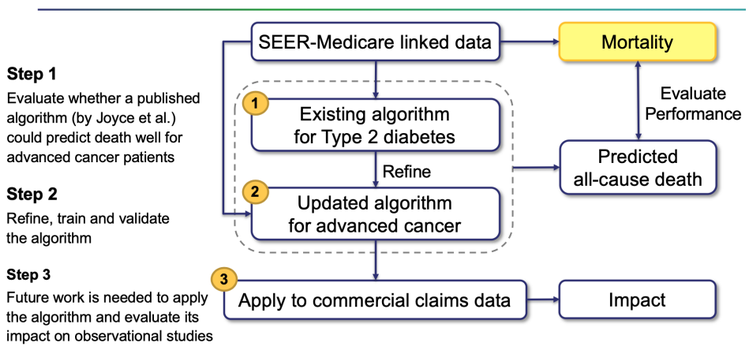
Background / Rationale

- Claims data are commonly used to generate real-world evidence in health economics and outcomes research studies.
- However, most claims databases do not provide death information, which limits available methods for research. For example:
  - Mortality information is needed to investigate survival
  - To apply common statistical methods for calculating costs in the presence of censoring, the reason for censoring (death versus disenrollment) is needed, since patients who died stop accumulating costs while those who are censored continue accumulating.
- An algorithm was developed to predict death in patients with Type 2 diabetes using claims data.[1,2] This algorithm (referred to as the DM algorithm throughout) includes fatal events such as cardiac arrest and organ transplant failure, hospitalization, ER visits, and ambulance services. Yet, it is unclear how the algorithm would perform in other disease areas, such as cancer.

Objective / Impact

- To refine and validate an algorithm to identify mortality from claims data for advanced cancer patients

Develop an algorithm to predict mortality in cancer population



METHODS

Data and Patient Selection

- The SEER-Medicare linked database, which includes mortality information, was used to modify and validate the existing algorithm.
- Patient conditions were identified from multiple Medicare claims files (i.e.Medpar, NCH, OPT, DME Hospice, HHA) using ICD-9 diagnosis/procedure codes, HCPCS / CPT codes, revenue center codes, and place of services codes.
- Patients newly diagnosed in 2012-2013 with advanced colorectal cancer/ advanced breast cancer were followed to the earliest of end of enrollment or December 31, 2014.
- The patient sample was split between a 50% training set and 50% test set

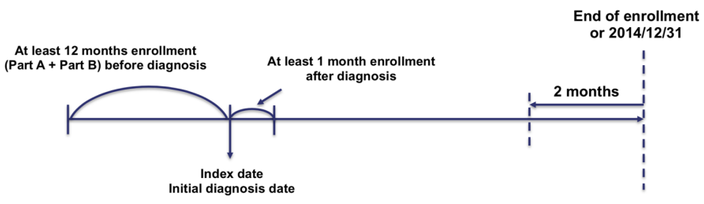
Patient Selection

Selection criteria	Advanced breast cancer (aBC)	Advanced colorectal cancer (aCRC)
Patients had at least 1-year of continuous enrollment in Medicare (Part A and Part B) before their initial diagnosis with advanced breast / colorectal cancer in 2012-2013, and had at least one month of continuous enrollment Medicare after diagnosis*	2,309	5,534
Without HMO enrollment	1,465	3,525
Training set	716	1,746
Test set	749	1,779

*Advanced cancer included stage 3B and stage 4

Development and the evaluation of the algorithm

- Three algorithms were developed and evaluated: one applied to only the aBC cohort, one applied to only the aCRC cohort, and a combined algorithm applied to both cohorts
- The algorithms defined death as the presence of any listed claims during the last 2 months of enrollment or 2014/12/31, whichever is earliest
- Mortality status from the Medicare date-of-death as of 2014/12/31 was used to benchmark the performance of the algorithm



Identification of Predictors

- Codes were identified as important predictors of mortality at end of enrollment by ranking them according to the fold-change difference in the number of deceased patients for which they appear in the last two months of enrollment versus the number of surviving patients for which they appear in the last two months of enrollment

$$\text{Fold-change difference} = \frac{\text{\# of patients who died}}{\text{\# of patients who survived}}$$

- Codes not appearing on claims in the last two months of enrollment for any surviving patients (i.e. infinite fold-change) were ranked according to the percentage of deceased patients in the cohort with the code in the last two months of enrollment
- Fold-change difference and % patients died (among codes with infinite fold-change difference) were varied in order to identify code sets with optimal performance **(Table 1)**
- Final code sets were adapted based on clinical feedback prior to validating the algorithm on the test set cohorts
- The performance of the algorithm was evaluated using sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)

RESULTS

Figure 1 - Mortality rate among the test set cohorts by cancer type

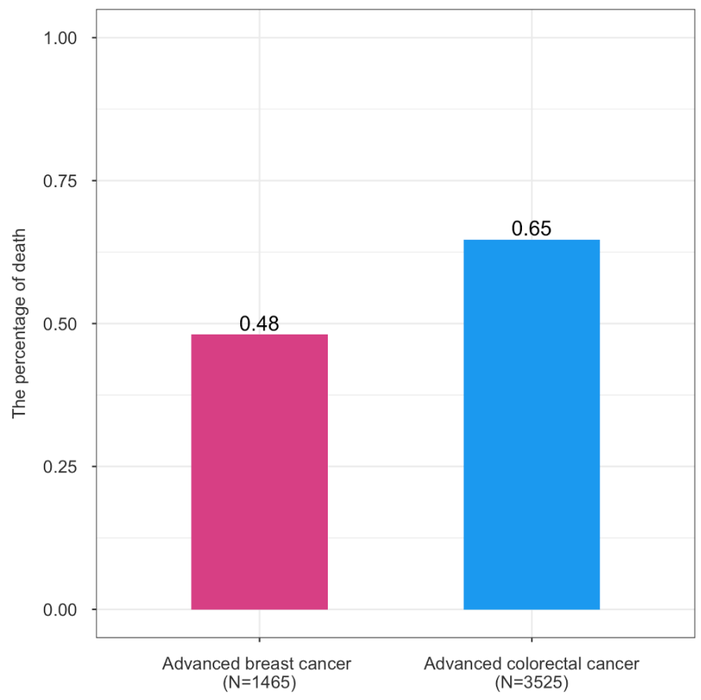


Table 1 - Predictor selection process: an example

Criteria		Number of codes satisfying all criteria in both aBC and aCRC training cohorts	Combined training cohort			
Fold-change difference	% of patients who died among those codes which were not shown in any patient who survived (From >=0.5% to >=5 %)		Sensitivity	Specificity	PPV	NPV
> 5	>= 1%	184	98.27	65.23	80.03	96.37
> 10	>= 1%	73	96.47	85.85	90.63	94.49
> 15	>= 1%	45	95.15	91.16	93.85	92.99
> 20	>= 1%	32	87.4	94.99	96.12	84.16
> 25	>= 1%	26	61.01	97.94	97.67	63.91
> 30	>= 1%	23	56.86	98.72	98.44	61.73
> 5	>= 3%	135	98.13	66.01	80.37	96.14
> 10	>= 3%	43	95.57	86.94	91.21	93.26
> 15	>= 3%	25	92.73	91.94	94.23	89.91
> 20	>= 3%	15	80.54	95.58	96.27	77.59
> 25	>= 3%	11	44.53	98.43	97.57	55.57
> 30	>= 3%	9	40.17	99.02	98.31	53.85

Table 2 - Codes included in the final algorithms; cells marked 'V' indicate inclusion in the algorithm

Category	Type	Codes	Short description	The aBC algorithm (39 codes)	The aCRC algorithm (44 codes)	The combined algorithm (32 codes)
Acute/critical care	HCPCS / CPT	99223	INITIAL ADMISSION E&M, HIGH COMPLEXITY		V	
		99238	HOSPITAL DISCHARGE DAY MANAGEMENT, 30 MIN	V		
		99292	CRITICAL CARE E&M, EACH ADDITIONAL 30 MIN	V	V	V
		99356	PROLONGED INPATIENT FACE-TO-FACE SERVICES		V	V
Advanced directive	ICD-9 diagnosis	V49.86	DO NOT RESUSCITATE STATUS	V	V	
Complications	ICD-9 diagnosis	038.9	SEPTICEMIA NOS			V
		514	PULMONARY CONGESTION AND HYPOSTASIS		V	
		570	ACUTE NECROSIS OF LIVER	V	V	V
		586	RENAL FAILURE NOS		V	
		273.8	OTHER DISORDERS OF PLASMA PROTEIN METABOLISM		V	
		275.3	DISORDERS OF PHOSPHORUS METABOLISM		V	
		572.8	OTHER SEQUELAE OF CHRONIC LIVER DISEASE		V	
		403.90	HYPERTENSIVE CHRONIC KIDNEY DISEASE, NOS	V		
		518.81	ACUTE RESPIRATORY FAILURE	V		V
		569.83	PERFORATION OF INTESTINE		V	
		780.02	TRANSIENT ALTERATION OF AWARENESS		V	
		785.52	SEPTIC SHOCK	V	V	V
		995.92	SEVERE SEPSIS	V	V	V
		V49.84	BED CONFINEMENT STATUS	V	V	V
		427.5	CARDIAC ARREST	V	V	V
		578.9	HEMORRHAGE OF GASTROINTESTINAL TRACT, NOS	V		V
		584.5	ACUTE KIDNEY FAILURE WITH LESION OF TUBULAR NECROSIS	V	V	V
		783.7	ADULT FAILURE TO THRIVE	V	V	V
		799.4	CACHEXIA	V		
Hospice/palliative	HCPCS / CPT	Q5005	HOSPICE, INPATIENT HOSPITAL	V	V	V
		Q5006	HOSPICE, HOSPICE FACILITY	V	V	V
	ICD-9 diagnosis	V66.7	ENCOUNTER FOR PALLIATIVE CARE	V	V	V
	Revenue center code	0651	HOSPICE-ROUTINE HOME CARE	V	V	
		0652	HOSPICE-CONTINUOUS HOME CARE	V	V	V
		0656	HOSPICE-GENERAL INPATIENT CARE	V	V	V
		0657	HOSPICE-PHYSICIAN SERVICES	V	V	V
Lab/evaluation procedure	HCPCS / CPT	88313	SPECIAL STAIN INCLUDING INTERPRETATION AND REPORT		V	
		G8721	TUMOR PATHOLOGY REPORT		V	
		G8784	NOT ELIGIBLE FOR BLOOD PRESSURE DUE TO URGENT SITUATION	V	V	V
	ICD-9 procedure	45.25	BIOPSY OF LARGE INTESTINE		V	
Non-urgent care	HCPCS / CPT	99304	INITIAL NURSING FACILITY CARE		V	
		A0428	AMBULANCE SERVICE, NON-EMERGENCY TRANSPORT		V	
		G0154	SKILLED NURSING SERVICES (HOME HEALTH/HOSPICE, 15 MIN)	V		
		G0155	SERVICES OF CLINICAL SOCIAL WORKER (HOME HEALTH/HOSPICE, 15 M	V	V	
		G0156	SERVICES OF HOME HEALTH/HOSPICE AIDE, 15 MIN	V	V	V
		G8553	ELECTRONIC PRESCRIPTIONS GENERATED AND TRANSMITTED		V	
	Revenue center code	0550	SKILLED NURSING-GENERAL CLASSIFICATION	V	V	V
		0551	SKILLED NURSING-VISIT CHARGE	V		
		0552	SKILLED NURSING-HOURLY CHARGE	V		
		0560	MEDICAL SOCIAL SERVICES-GENERAL CLASSIFICATION		V	
		0561	MEDICAL SOCIAL SERVICES-VISIT CHARGE	V	V	
		0569	MEDICAL SOCIAL SERVICES-OTHER	V		V
		0571	HOME HEALTH AID-VISIT CHARGE	V	V	
		0572	HOME HEALTH AID-HOURLY CHARGE	V		V
Treatment procedure	HCPCS / CPT	31500	ENDOTRACHEAL INTUBATION	V	V	V
		36561	INSERTION OF TUNNELED CENTRALLY INSERTED CENTRAL VENOUS ACC	V		
		92950	CARDIOPULMONARY RESUSCITATION		V	V
		4045F	EMPIRIC ANTIBIOTIC PRESCRIBED	V	V	V
		A0422	AMBULANCE, LIFE SUSTAINING SITUATION	V		
	ICD-9 procedure	38.93	VENOUS CATHETERIZATION, NOT ELSEWHERE CLASSIFIED	V		V
		96.04	INSERTION OF ENDOTRACHEAL TUBE	V	V	V
		96.71	CONTINUOUS INVASIVE MECHANICAL VENTILATION (< 96 HOURS)	V		V
		96.72	CONTINUOUS INVASIVE MECHANICAL VENTILATION (96+ HOURS)		V	
		99.07	TRANSFUSION OF OTHER SERUM		V	V

Figure 2 - Performance of algorithms on the test cohort data

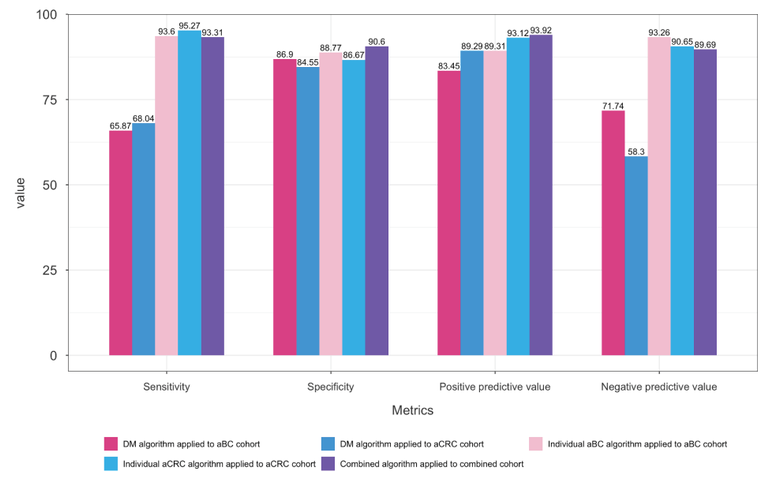


Table 3 - Comparison of the performance between the algorithms with and without codes for metastases

- From a clinical perspective, brain metastases / liver metastases could be a potential predictor for death among patients with breast cancer / colorectal cancer
- The performance of the individual aBC algorithm was consistent regardless of including a code for brain metastases, whereas the specificity of the individual aCRC algorithm dramatically decreased from 86.67% to 69.92% after including a code for liver metastases

	The individual aBC algorithm		The individual aCRC algorithm	
	No code for brain metastases	Add a dignosis code for brain metastases	No code for liver metastases	Add a dignosis code for liver metastases
Sensitivity	93.6	93.87	95.27	96.13
Specificity	88.77	87.43	86.67	69.92
Positive predictive value	89.31	88.22	93.12	85.81
Negative predictive value	93.26	93.43	90.65	90.53

Table 4 - the performance of selected codes

- The top five codes in Table 4 represent high performing predictors included in all three algorithms. The bottom two codes are for liver and brain metastases, which were not included in the final algorithms
- We observed that a subset of colorectal cancer survivors had liver metastases in their last 2 months of claims, which lead to a large decrease in specificity.

Type	Code	Code Description	Advanced Breast Cancer				Advanced Colorectal Cacner			
			Number of deceased patients	Number of surviving patients	Fold-change difference	The percentage of deceased patients	Number of deceased patients	Number of surviving patients	Fold-change difference	The percentage of deceased patients
HCPCS	Q5005	HOSPICE, INPATIENT HOSPITAL	30	0	Inf	4.19%	77	0	Inf	4.41%
HCPCS	Q5006	HOSPICE IN HOSPICE FACILITY	26	0	Inf	3.63%	109	1	109	6.24%
ICD 9 diag	4275	ARREST, CARDIAC	21	0	Inf	2.93%	80	1	80	4.58%
ICD 9 diag	V4984	STAUS, BED CONFINEMENT	21	0	Inf	2.93%	70	2	35	4.01%
ICD 9 diag	78552	SYMPTOM, SHOCK, SEPTIC	25	0	Inf	3.49%	85	1	85	4.87%
ICD 9 diag	1977	NEOPLASM, METASTATIC, LIVER	91	27	3.37	12.71%	558	132	4.23	31.96%
ICD 9 diag	1983	NEOPL, MTSTC, BRAIN AND SPINAL CORD	44	9	4.89	6.15%	36	3	12	2.06%

CONCLUSIONS

- Relative to its performance in Type 2 diabetes patients, the performance of the DM algorithm decreases considerably when applied to the aBC and aCRC patients
- The algorithms developed and shown here for aBC and aCRC patients have at least 90% sensitivity and 85% specificity for identifying death status at the end of enrollment when the information is not available in claims. While other approaches utilizing predictive modeling could potentially improve performance, these algorithms are easy to apply using a short list of codes.

LIMITATIONS

- The algorithms were developed by observing conditions among the elderly with aBC or aCRC who were diagnosed between 2012 and 2013. Future work is needed to validate these algorithms in a younger population and those cancer patients diagnosed over other years.
- The algorithms are not applicable using claims after 2015 when ICD-9 codes were replaced with ICD-10 codes. Additional mapping of codes is required.
- While the performance of the algorithms is similar in aBC and aCRC patients, they may vary when applied to other types of cancers.

REFERENCES

1. Stafkey-Maliley, D., Wang, W., Murty, S., & Shetty, S. Validation of an Algorithm to Identify Death in Administrative Claims Data. AMCP 2017 Nexus, Dallas, TX.
2. Joyce, A. T., Iacoviello, J. M., Nag, S., Sajjan, S., Jilinskaia, E., Throop, D., Alexander, C. M. (2004). End-Stage Renal Disease-Associated Managed Care Costs Among Patients With and Without Diabetes. *Diabetes Care*, 27(12), 2829.