RWD140

Validity of Diagnostic Codes to Identify Metastatic Prostate Cancer in Medicare Claims Database

Bo Zhou[1], **Amit D. Raval**[2], Yifan Zhang[1], Nethra Sambamoorthi[1], Usha Sambamoorthi[1]
[1]University of North Texas Health Science Center, Fort Worth, TX, USA; [2]Bayer HealthCare Pharmaceuticals, Whippany, NJ, USA

BACKGROUND

- With the increasing acceptance of real-world evidence (RWE) derived through real-world data (RWD) from regulatory bodies, one cannot undermine the need of have valid operational definitions to generate robust and reliable RWE
- Prostate cancer (PC) is the most common cancer in men in the United States, with a rise in the incidence of metastases (mPC) at diagnosis and parallel advancement in therapeutic alternatives for mPC in the last decade
- Accurate identification of mPC for RWE generation using claims remains a challenge
  - Prior research on the identification of diagnosis codes for mPC using claims showed poor validity compared to gold-standard registry data--- resulting in a guidance from the National Cancer Institute (NCI) advising against claims codes to identify mPC[1]
  - Prior studies were based on an older database covering the period until 2007, with coding semantics including International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)[2]
- Since 2015, the US has adopted more precise ICD-10 coding systems, and metastasis-directed therapies for PC have also been available with potential prior authorization restriction. These changes, along with RWD field in general, may have resulted in the adoption of coding practices to reflect the mPC diagnosis through claims

OBJECTIVE

To examine the validity of diagnostic codes to identify mPC in Medicare claims (mPC_{claims}) against the gold-standard cancer registry-based mPC_{SEER} using the Surveillance, Epidemiology, and End Results (SEER) Registry-linked Medicare claims database 2016-2019

METHODS

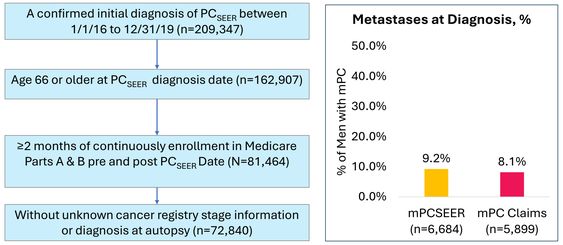
- ***Design:*** Retrospective observational cohort
- ***Data source***: SEER Registry-linked Medicare administrative claims
- ***Study population:*** Men diagnosed with PC_{SEER} between 1/1/2016 to 12/31/2019, continuous continuous enrollment in Medicare fee-for-service Parts A and B for ≥ 2 months pre- and post-PC diagnosis
- ***Study measures:***
  - mPC_{SEER} : Registry documented combined Summary Stage 7: Distant and M1b for the presence of bone metastases
  - mPC_{claims} : ≥ 1 claims with the International Classification of Disease (ICD)-10 diagnosis codes for any metastases (ICD-10: C77.x-C80.0, C7B) and bone metastase)ICD-10: within 2 months pre- and post- (registry-based) mPC_{SEER} date
- ***Statistical analysis:***
  - Diagnostic validity parameters, such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and concordance (accuracy) of claim-based mPC, were estimated using mPC_{SEER} as a gold standard
  - Population characteristics were descriptively summarized for demographic, clinical and survival outcomes for mPC_{SEER} and mPC_{claims}
  - Logistic regression was used to examine factors associated with the discordance (sum of false positives and false negatives) in mPC diagnosis using demographic and socio-economic factors, including age, race, marital status, and county-level socioeconomic YOST index

RESULTS

Study Population

- The study cohort comprised 72,840 men with PC between 2016 to 2019 (**Fig. 1**)

Figure 1. Study Cohort Identification and



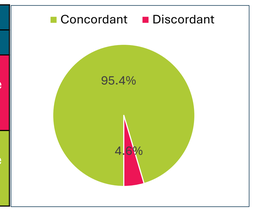
- Among 72,840 men diagnosed with PC, a total of 6,684 (9.2%) had registry-based mPC, and 5,899 (8.1%) had claims-based mPC at diagnosis

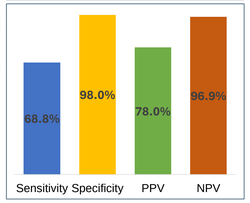
Diagnostic Parameters

- The agreement (concordance) in accurately classifying the presence and absence of true metastases (mPC_{SEER} registry) was 95.4% with mPC_{claims}

Figure 2. mPC Diagnosis among PC using Registry and Claims with Confusion Matrix

Confusion Matrix		mPC_{SEER} Yes	mPC_{SEER} No
mPC_{claims}	Yes	True Positive 4,600	False Positive 1,299
	No	False Negative 2,084	True Negative 64,857





- *Sensitivity/Recall*: mPC_{claims} correctly identified 68.8% of actual positive mPC_{SEER}
- *Specificity*: mPC_{claims} correctly identified 90.0% of actual positive mPC_{SEER}
- *PPV/Precision*: Out of mPC_{claims} based positive cases, 78.0% are actually positive mPC_{SEER}
- *NPV*: Out of mPC_{claims} based negative cases, 96.9% are actually positive mPC_{SEER}
- *F1 score* of 73.1% highlights adequate performance (>70%) of mPC_{claims}
- *False Positive Rate*: Only 1.96% of negative cases are incorrectly classified as positive using mPC_{claims}

Study Population Characteristics

- Study Demographics, socio-economic status and clinical characteristics are detailed in **Table 1**

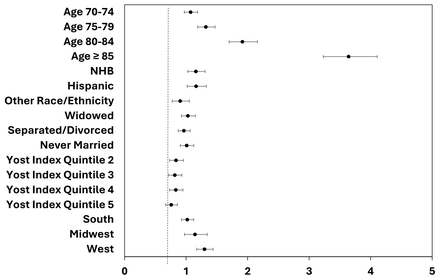
Table 1: Baseline demographic, clinical characteristics and follow-up survival outcomes

	Cohort	mPC_{SEER}	mPC_{claims}
Cohort size	72,840	6,684	5,889
Demographic and Socio-economic Characteristics			
Age group, n (%)			
66–69	23,480 (32.2%)	1,264 (18.9%)	1192 (20.2%)
70–74	24,468 (33.6%)	1,541 (23.1%)	1396 (23.7%)
75–79	14,655 (20.1%)	1,355 (20.3%)	1222 (20.7%)
80–84	6,548 (9.0%)	1,168 (17.5%)	999 (16.9%)
85+	3,689 (5.1%)	1,356 (20.3%)	1090 (18.5%)
Race, n (%)[b]			
NHW	55,904 (76.7%)	5,015 (75.0%)	4483 (76%)
NHB	7,316 (10%)	728 (10.9%)	579 (9.8%)
Others	9,620 (13.2%)	941 (14.1%)	837 (14.2%)
Region, n (%)			
Northeast	23,266 (31.9%)	2,152 (32.2%)	1951 (33.1%)
South	22,097 (30.3%)	1,823 (27.3%)	1581 (26.8%)
Midwest	5,380 (7.4%)	447 (6.7%)	415 (7.0%)
West	22,097 (30.3%)	2,262 (33.8%)	1952 (33.1%)
YOST Index[ɣ], n (%)			
Q1	8,313 (11.4%)	901 (13.5%)	739 (12.5%)
Q2	10,606 (14.6%)	1,000 (15%)	873 (14.8%)
Q3	12,951 (17.8%)	1,247 (18.7%)	1092 (18.5%)
Q4	15,671 (21.5%)	1,441 (21.6%)	1303 (22.1%)
Q5	23,137 (31.8%)	1,918 (28.7%)	1731 (29.3%)
Prostate Cancer- Disease-Specific Characteristics			
PSA value			
Valid N (%)	41,298 (56.7%)	3,780 (56.6%)	3,328 (56.4%)
Mean (SD)	16.9 (24.0)	65.4 (37.5)	63.3 (38.7)
Median (IQR)	7.9 [5.5-14.0]	98.0 [23.9 - 98.0]	92.4 [19.6 - 98.0]
Gleason Score			
Valid N (%)	43,970 (60.4%)	2,427 (36.3%)	2,204 (37.4%)
Mean (SD)	7.2 (1.0)	8.6 (0.9)	8.5 (0.9)
Median (IQR)	7 [6 - 8]	9 [8 - 9]	9 [8 - 9]
Survival Estimates[§]			
Death, N (%)	7,429 (10.2%)	3,294 (49.3%)	2,763 (42.8%)
Survival time in months (K-M Estimates)			
Median (95% CI)	NE (NE, NE)	33.7 [32.5, 34.9]	36.4 [34.8, 37.9]

Notes: [b]Died as of Dec. 31, 2020; ¥It uses a weighted combination of seven census variables: average educational level, median household income, poverty rate, median rent, unemployment rate, and employment mix. The first quintile (Q1) represents the lowest socioeconomic status, while the fifth quintile (Q5) represents the highest socioeconomic status

- Demographic characteristics were almost identical between mPC_{claims} and mPC_{SEER} albeit a bit 1-2% lower proportion of mPC in older men aged 85 and above, NHB, and the lower SES status (YOST index Q1, Q2)
- Clinical characteristics, including the distribution of disease severity measures such as PSA and Gleason score, were identified between mPC_{claims} and mPC_{SEER}
- Although a bit lower proportion of mPC_{claims} died, the median time to survival were identical between mPC_{claims} and mPC_{SEER}

Figure 4. Adjusted Odds Ratios (AOR) and 95% Confidence Intervals (CI) of Factors Associated with SEER-Medicare Metastases Diagnosis Discordance (Ref = Concordant) Older Men Diagnosed with Prostate Cancer in 2016-2019



Note: The reference groups are Age 66-69, Non-Hispanic White (NHW), Married, Yost Index Quintile 1, Northeast. The Yost Index's missing values and a risk group based on PSA and Gleason Score are included in the regression but not in the plot.

- Key factors associated with discordance: older age (85+ years), African American race, and low socioeconomic index (Yost Index Q1, Q2)

Subgroup: Bone Metastases using SEER (N=46,512)

- 3,742 and 2,958 had registry and claim-based bone metastasis, respectively
- Accuracy, sensitivity, specificity, PPV and NPV were 96.6%, 68.1%, 99.0%, 86.2%, and 97.3%, respectively
- Key factors associated with discordance: older age, and low socio-economic index

Key study limitations

- Our study primarily includes fee-for-service Medicare beneficiaries. Findings may not be generalizable to Medicare Advantage beneficiaries or commercial insurance beneficiaries
- While mPC_{claims} provides very high confidence in removing true negatives, there was undercoding of mPC in claims, and disparities in coding mainly in individuals with older age (85+) and with low socio-economic status

CONCLUSIONS

- To our knowledge, this is the most up-to-date and largest nationwide cancer registry-linked claims data summarizing the accuracy of claim-based metastatic prostatic cancer diagnosis
- Findings suggest claims-based mPC accurately identified metastatic prostate cancer at diagnosis among older men, albeit with moderate sensitivity. This may be due to changes in the coding practice in the era of reimbursement requirements for metastasis-directed therapies

References

1. NCI-SEER. Available on URL [accessed on 16 April 2025]:https://healthcaredelivery.cancer.gov/seermedicare/considerations/measures.html
2. Lowrance *et al.* JCO 30, e15148-e15148(2012).
3. Sathiakumar *et al.* Med Care. 2017;55(12):e144-e9

Disclosures
This study was funded by Bayer HealthCare Pharmaceuticals Inc. ADR is an employee at Bayer HealthCare Pharmaceuticals.

Abbreviations: AOR, Adjusted Odds Ratio; CI, Confidence Interval; ICD-9 CM, International Classification of Diseases, Ninth Revision, Clinical Modification; IQR, Interquartile Range; K-M, Kaplan-Meier; mPC, metastatic Prostate Cancer; NCI, National Cancer Institute; NE: not evaluable; NHB, Non-Hispanic Black; NHW, Non-Hispanic White; NPV, Negative Predictive Value; PC, Prostate Cancer; PPV, Positive Predictive Value; PSA, prostate specific antigen; RWE, Real-World Evidence; SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results

Presented at the International Society for Pharmacoeconomics and Outcomes Research 2025; May 13–16, 2025; Montreal, QC, Canada

Presenting author: Amit Raval, PhD, M.Pharm (PhPr)
amitkumar.raval@bayer.com