SA27

Comparison of comorbidity indices between an EHR-derived database and claims data among patients with metastatic breast cancer

Alemseged A. Asfaw, PhD[1], Mustafa Steven Ascha, PhD[1], Prakirthi Yerram, PharmD[1], Samantha N. Reiss, PharmD[1], Sarah Brake, PhD[1], and Niquelle Brown Wadé, PhD[1]

[1] Flatiron Health, New York, NY

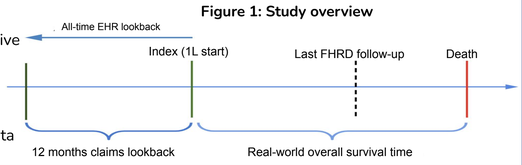
Background

- Electronic health records (EHR) and healthcare claims data are primarily collected for clinical care and healthcare services reimbursement, respectively.[1,2]
- Baseline patient characteristics relevant to cancer diagnosis and treatment are more likely than other characteristics to be captured in an oncology EHR. Characteristics taken from claims data are likely to capture more patient information beyond a specific cancer diagnosis but may require chart review for confirmation.
- While comorbidities are directly documented by clinicians in an EHR, these conditions must be derived from claims data. One derivation algorithm is the Charlson Comorbidity Index (CCI), a weighted sum of specific comorbidities (derived from observed diagnosis and procedure codes).[3]
- Comparisons of EHR-derived and claims data can reveal how each source captures comorbidities. The objective of this study is to compare EHR-derived and claims data on individual comorbidities and CCI.

Methods

- Flatiron Health Research Database (FHRD) is a nationwide oncology EHR-derived longitudinal database comprised of de-identified patient-level structured and unstructured data, curated via technology-enabled abstraction. During the study period, the de-identified data originated from approximately 280 US cancer clinics (~800 sites of care).[4,5]
- Komodo Health is healthcare technology company and its Healthcare Map™ consists of proprietary real-time commercial claims activity data on 330 Million US patients and their interactions with the US healthcare system.
- This retrospective cohort study used a de identified probabilistic matched linked clinico claims data set from the FHRD and the Komodo Healthcare Map. Patients diagnosed with HR+/HER2- metastatic breast cancer (mBC) between 1/2011 and 9/2020 receiving aromatase inhibitor first-line (1L) treatment were selected from the linked FHRD/Komodo Claims Health Database (KHCD) dataset. Claims-based CCI coding algorithms identified comorbidity status using diagnosis and procedure codes, and individual comorbidity status was abstracted from patient charts in EHR (Figure 1).[6]
- Aggregate CCI kappa, sensitivity, specificity, predictive values, and hazard ratio associated with each data source were used to assess alignment between EHR-derived using claims data as the reference standard.[7]

Figure 1: Study overview



Results

- This study included 654 patients from FHRD who had documented claims coverage in KHCD in the year before cohort entry (Figure 2). Patients with any comorbidity had similar baseline characteristics when identified using either data source (Table 1).
- The proportions of patients having a CCI score of 0 were 56% and 55% for FHRD and claims, for CCI score of 1, 24% and 23%, for a CCI score of 2+, 20% and 22%.

Figure 2: Cohort selection

- Kappa of dichotomized CCI (any comorbidity vs none) comparing FHRD to a reference of claims was 0.39, and classification statistics were approximately 0.7 (Table 2). Kappa of multiple CCI categories (0, 1 and 2+) was 0.32 (0.26, 0.37).

Table 2: Concordance and classification statistics

Kappa	Sensitivity	Specificity	PPV	NPV
0.39 (0.31, 0.46)	0.65 (0.60, 0.71)	0.73 (0.68, 0.78)	0.67 (0.61, 0.72)	0.72 (0.67, 0.77)

- Overall, CCI component comorbidities showed comparable prevalence across data sources with notable differences in COPD and mild liver disease (Figure 3). Similar hazards of mortality were observed for CCI scores of 0, 1 and 2+ (Figure 4) between the data sources.

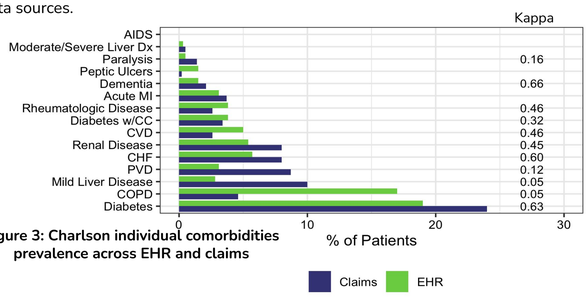


Figure 3: Charlson individual comorbidities prevalence across EHR and claims

Table 1: Baseline patient characteristics

	CCI > 0		CCI = 0	
Characteristic	**Per EHR**	**Per Claims**	**Per EHR**	**Per Claims**
Age at Diagnosis				
18-49	16 (5.5%)	16 (5.9%)	46 (13.1%)	46 (11.5%)
50-64	105 (36%)	94 (35%)	185 (51%)	196 (51%)
65-74	89 (31%)	81 (30%)	74 (20%)	82 (21%)
75 or above	79 (27%)	78 (29%)	60 (16%)	61 (16%)
ECOG Performance Score				
0	65 (22.5%)	67 (24.9%)	130 (35.6%)	128 (33.2%)
1	55 (19.0%)	52 (19.3%)	54 (14.8%)	57 (14.8%)
2+	43 (14.9%)	41 (15.2%)	33 (9.0%)	35 (9.1%)
Unknown	126 (43.6%)	109 (40.5%)	148 (40.5%)	165 (42.9%)
Group Stage				
I	34 (12%)	29 (11%)	43 (12%)	48 (12%)
II	75 (26%)	71 (26%)	104 (28%)	108 (28%)
III	55 (19%)	52 (19%)	60 (16%)	63 (16%)
IV	92 (32%)	76 (28%)	125 (34%)	141 (37%)
Not documented	33 (11%)	41 (15%)	33 (9.0%)	25 (6.5%)
Race				
Asian	8 (3.0%)	9 (3.7%)	12 (3.4%)	11 (3.0%)
Black or African American	29 (11%)	31 (13%)	24 (6.9%)	22 (6.0%)
Other Race	31 (12%)	32 (13%)	48 (14%)	47 (13%)
White	197 (74%)	174 (71%)	263 (76%)	286 (78%)
Unknown	24	23	17	18

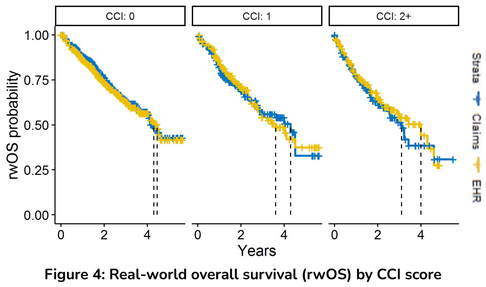


Figure 4: Real-world overall survival (rwOS) by CCI score

Discussion and conclusion

- Flatiron Health EHR-derived clinical and demographic characteristics (age at index, gender, race, region, ECOG status at index, practice type, and stage of disease at diagnosis) were comparable between patients with and without comorbidities per claims versus EHR-derived data.
- Cohort level findings and outcomes were similar in this 1L mBC linked cohort as seen by similar prevalence of comorbidities (Figure 3) and rwOS by CCI score (Figure 4). These findings showed that most of the CCI component comorbidities, with notable exception of COPD and mild liver disease, can be identified from either abstracted FHRD or claims data in 1L mBC population.
- However, there are differences in certain CCI comorbidities, for example COPD and mild liver disease, as seen by kappa concordance shown in figure 3. The observed discordance could be due to reimbursement-incentivized coding and/or oncologists' documentation practices.
- For cohort selection and other use cases requiring high sensitivity, it is strongly recommended to consider using both data sources when possible to identify patient comorbidity status given the discordance at the patient level. However, using both data sources may decrease specificity, so additional sensitivity analyses, such as a stratified analysis as well as incorporating lab results when available, may be helpful to test the robustness of the study results.

Study Limitations

- Patient follow-up around index dates is limited by data availability. KHCD only captures commercial claims data, patients over the age of 65 may have Medicare insurance, which can lead to lack of certain comorbidity data for patients aged 65 or above with Medicare coverage.
- EHR-derived data for the mBC cohort were abstracted from any time prior to 1L start; however, only the window of time from one year prior to 1L start is considered for claims comorbidities derivation.
- Patients were selected based on enrollment in plans whose adjudicated claims data are available which can lead to decrease in sample size. Additional analysis to understand the representativeness of the final study cohort would be valuable.
- Additional investigation to understand reasons for discrepancy were not conducted due to patient privacy restrictions of patient re-identification.

References

1. Kataria S, Ravindran V. Electronic health records: a critical appraisal of strengths and limitations. J R Coll Physicians Edinb. 2020 Sep;50(3):262-268. doi: 10.4997/JRCPE.2020.309. PMID: 32936099.
2. Konrad R, Zhang W, Bjarndóttir M, Proaño R. Key considerations when using health insurance claims data in advanced data analyses: an experience report. Health Syst (Basingstoke). 2019;9(4):317-325. Published 2019 Mar 1. doi:10.1080/20476965.2019.1581433
3. Hude Quan, Bing Li, Chantal M. Couris, Kiyohide Fushimi, Patrick Graham, Phil Hider, Jean-Marie Januel, Vijaya Sundararajan, Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries, American Journal of Epidemiology, Volume 173, Issue 6, 15 March 2011, Pages 676–682, https://doi.org/10.1093/aje/kwq433
4. https://www.medrxiv.org/content/10.1101/2020.03.16.20037143v2
5. https://arxiv.org/abs/2001.09765
6. Hude Quan, Bing Li, Chantal M. Couris, Kiyohide Fushimi, Patrick Graham, Phil Hider, Jean-Marie Januel, Vijaya Sundararajan, Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries, American Journal of Epidemiology, Volume 173, Issue 6, 15 March 2011, Pages 676–682, https://doi.org/10.1093/aje/kwq433
7. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. J Clin Epidemiol. 2012;65:343–349.

Presented at ISPOR Europe 2022, 6-9 November 2022, Vienna, Austria and Virtual. For additional information, contact Alemseged Ayele Asfaw at alemseged.asfaw@flatiron.com.

At the time of the study, all authors report employment at Flatiron Health, Inc., which is an independent subsidiary of the Roche Group, and stock ownership in Roche.