Bias Characterization of Real-World Patients with and without Imaging in a Community Oncology Electronic Health Record-derived Database

Xinye Li[1], Benjamin Ackerman[1], Kelly Magee[1], Jonathan Kern[1], Katherine Tan[1]

[1] Flatiron Health, New York, NY

Background

Radiology imaging is critical to diagnose and monitor response to treatment for patients with cancer. Assessment of real-world response from electronic health record (EHR) documentation alone is reliant upon clinician documentation [1], which limits the ability to apply standardized measurements such as RECIST criteria. With the addition of real-world imaging, a more quantifiable assessment of changes in tumor burden may be possible [2]. However, real-world radiographic imaging data availability and timing for patients may vary due to facility capabilities, institutional standards, etc., and selecting real-world patients based on imaging availability may introduce biases, potentially impacting generalizability of results. Therefore, this study aimed to assess the representativeness of imaging-derived cohorts relative to a broader real-world oncology target population based on available patient baseline characteristics and endpoints.

Methods

Data source

- This study used the nationwide Flatiron Health EHR-derived de-identified Research Database (FHRD), a longitudinal database comprising de-identified patient-level structured and unstructured data, curated via technology-enabled abstraction [3&4].
- During the study period, the de-identified data originated from approximately 280 US cancer clinics (~800 sites of care).
- De-identified imaging metadata were retrospectively retrieved from the Picture Archiving and Communication System (PACS), available for selected practices from the FHRD.

Cohort selection, target population, and imaging samples

- Patients with advanced non-small cell lung cancer (aNCLC) and diffuse large B-cell lymphoma (DLBCL) were selected and further categorized as one of the three study samples. Please see Figure 1 for detailed definitions of target populations and imaging samples.

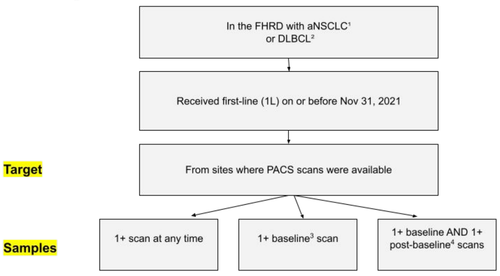
Patient characteristics

- Baseline characteristics: Demographics, biomarker status, clinical characteristics, treatment patterns, follow-up time
- Endpoints (using 1L initiation as index date)
  - Real-world overall survival (rwOS) for both aNSCLC and DLBCL cohorts
  - Real-world progression free survival (rwPFS) for aNSCLC cohort

Statistical analysis

- For each disease cohort (target from Figure 1), our comparisons were based on imaging availability at various time points (samples from Figure 1): any time point, baseline (pre-1L initiation), and post-baseline (during 1L duration).
- Differences in baseline demographic, clinical, and treatment characteristics between groups were measured through the absolute standardized mean difference (ASMD), with a threshold of 0.1 denoting a notable difference [5&6].
- Real-world Overall Survival (rwOS) and real-world Progression-Free Survival (rwPFS) endpoints between groups were compared through Kaplan-Meier estimates. Median survival estimates along with the 95% confidence intervals (CI) were reported.

Figure 1. Illustration of cohort selection



Disease-specific FHRD clinical I/E criteria

[1] aNSCLC:

- Inclusion
  - Diagnosed with with lung cancer (ICD-9 162.x or ICD-10 C34x or C39.9)
  - At least two documented clinical visits, on different days, occurring on or after January 1, 2011
  - Pathology consistent with NSCLC. Diagnosed with Stage IIIB, IIIC, IVA or IVB NSCLC on or after 1/1/2011, or diagnosed with early-stage NSCLC and subsequently develops recurrent or progressive disease on or after 1/1/2011.
- Exclusion
  - Lacking relevant unstructured documents in the FHRD for review by the abstraction team.

[2] DLBCL:

- Inclusion
  - Diagnosed with Non-Hodgkin's Lymphoma (ICD 9: 200x, 202x; ICD 10: C82x, C83x, C84x, C85x, C86x, C88x, C96x)
  - At least two documented clinical visits, on different days, occurring on or after January 1, 2011
  - Has evidence of DLBCL with an initial diagnosis date on or after January 1, 2011
- Exclusion
  - Lacking relevant unstructured documents in the FHRD for review by the abstraction team.

Scan timing definition

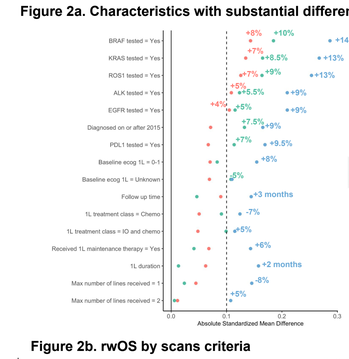
[3] Defined as a CT or PET-CT scan for aNSCLC patients, or PET-CT scan for DLBCL patients within 6 weeks prior to first line (i.e. 1L) initiation.

[4] Defined as a CT or PET-CT scan for aNSCLC patients, or PET-CT scan for DLBCL patients during 1L (non-maintenance portion only).

Results - aNSCLC

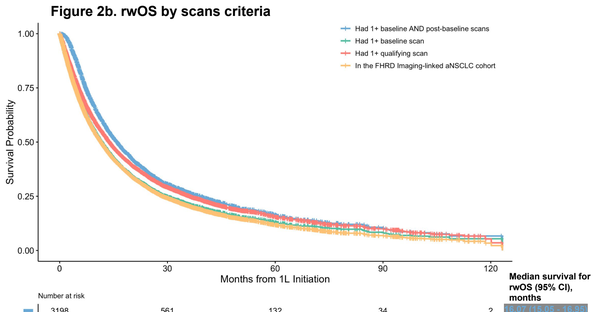
- A total of 11,056 patients met clinical IE criteria and were included in the FHRD imaging-linked aNSCLC target cohort. Of those, 7,493 patients had 1+ scan at any time, 4,619 patients had 1+ baseline scan, and 3,199 patients had 1+ baseline and 1+ post-baseline scans.
- Baseline characteristics compared to the aNSCLC target population
  - Higher biomarker testing rates and better data completeness in **all scans samples**
  - More recent diagnosis, a higher proportion of patients with lower ECOG performance scores among those **with a baseline scan (w/ or w/o a post-baseline scan)**
  - Longer follow-up time, longer 1L duration, more LOT received **if patients had both baseline and post-baseline scans**
- Real-world endpoints were similar between patients with 1+ baseline scan and the target population. The median survivals for both endpoints were slightly longer among patients with a scan at any time point (+2 months for rwOS, +0.5 months for rwPFS), and were significantly longer once post-baseline scans were required (+4.5 months for rwOS, +1 month for rwPFS).

Figure 2a. Characteristics with substantial differences (ASMD > 0.1)



Differences are provided when ASMD is > 0.1

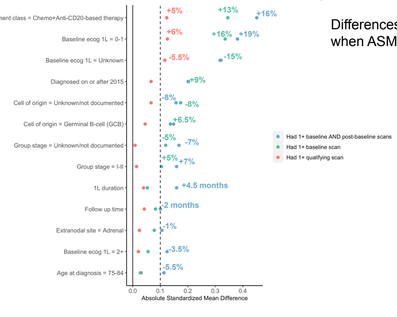
Figure 2b. rwOS by scans criteria



Results - DLBCL

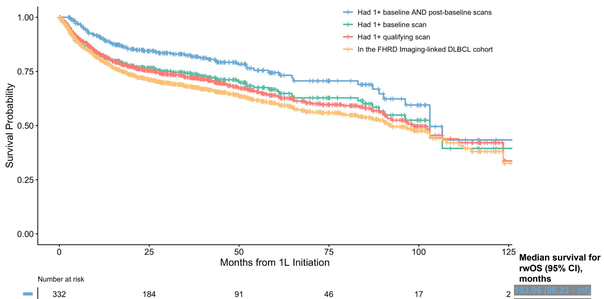
- A total of 1,323 patients met the clinical IE criteria and were included in the FHRD imaging-linked DLBCL target cohort. Of those, 932 patients had 1+ scan at any time, 454 patients had 1+ baseline scan, and 332 patients had 1+ baseline and 1+ post-baseline scans.
- Baseline Characteristics compared to the DLBCL target population
  - Less unknown and more patients with lower ECOG performance scores, more chemo and anti-cd20 therapy in 1L observed in **all scans samples**
  - More recent diagnosis, more with lower stage at diagnosis, more with germinal center B-cell-like (GCB) cell of origin for **patients with a baseline scan (w/ or w/o a post-baseline scan)**
- Patients with 1+ scan at any time point had similar median survival for rwOS compared to the target population. Longer median survivals were observed for patients with 1+ baseline scan (+10 months) or 1+ baseline and post-baseline scans (+10 months).

Figure 3a. Characteristics with substantial differences (ASMD > 0.1)



Differences are provided when ASMD is > 0.1

Figure 3b. rwOS by scans criteria



Conclusion

- Requiring scans at any time point or in the baseline window resulted in the cohort being similar to the broader real-world cohort with respect to baseline characteristics; however, median survival times trended longer for both rwOS and rwPFS endpoints.
- Requiring scans in the post-baseline window may introduce selection and immortal time bias. Patients with post-baseline scans appear healthier, as seen in the baseline characteristics and rwOS and rwPFS estimates.
- Effectively integrating real-world imaging into research studies requires an understanding of the representativeness of imaging-derived cohorts.
- Further work is needed to assess the application of existing methods to account for such selection and immortal bias in this context.

Reference

1. Ma, X., Bellomo, L., Magee, K., Bennette, C., Tymejczyk, O., Samant, M., Tucker, M., Nussbaum, N., Bowser, B., Kraut, J. and Bourla, A., 2022. *Characterization of a Real-World Response Variable and Comparison with RECIST-Based Response Rates from Clinical Trials in Advanced NSCLC*. Retrieved September 19, 2022 from https://link.springer.com/article/10.1007/s12325-021-01659-0
2. Ximran Ma, MS. "Concordance between Ehr-Derived Clinician-Assessed Response and RECIST-Based Imaging Response." *JAMA Network Open*, JAMA Network, 12 May 2022, Retrieved September 19, 2022 from https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2792220.
3. Ma, X., Long, L., Moon, S., Adamson, B. J. S., & Baxi, S. S. (2020, January 1). *Comparison of population characteristics in real-world clinical oncology databases in the US: Flatiron Health, Seer, and NPCR*. medRxiv. Retrieved September 7, 2022, from https://www.medrxiv.org/content/10.1101/2020.03.16.20037143v2
4. Birnbaum, B. (2020, January 13). *Model-assisted cohort selection with bias analysis for generating*. arXiv.org. Retrieved September 7, 2022, from https://arxiv.org/abs/2001.09765
5. Peter C. Austin (2009) Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research, Communications in Statistics - Simulation and Computation, 38:6, 1228-1234, DOI: 10.1080/03610910902859574, Retrieved September 7, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3472075/#b23
6. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *Taylor & Francis*. Retrieved September 14, 2022 from https://www.tandfonline.com/doi/abs/10.1080/00031305.1985.10479383

This study was sponsored by Flatiron Health, Inc., which is an independent subsidiary of the Roche Group.
XL, BA, KM, JK, and KT report employment at Flatiron Health, Inc., which is an independent subsidiary of the Roche Group, and stock ownership in Roche.