

Electronic Health Records vs Administrative Claims to Support Pharmacovigilance Review of Safety Events: A Fit-for-Purpose Evaluation

Lili Zhou,¹ Qianhong Fu,¹ Joanne Tse,¹ Ayad K. Ali¹

¹BeOne Medicines, Ltd, San Carlos, CA, USA

CONCLUSIONS

- This study shows differences in outcomes captured across RWD sources, highlighting the variation in data relevance, quality, and extensiveness between sources
- Infection incidence and all-cause hospitalizations were comparable between Symphony and TriNetX, whereas data suggest that these outcomes were substantially underreported in the oncology practice EHR database. In contrast, the oncology practice EHR database captured mortality data more comprehensively
- These findings highlight the importance of conducting a fit-for-purpose assessment when selecting RWD sources for pharmacovigilance research

INTRODUCTION

- Advances in informatics have expanded the availability of real-world data (RWD) to support pharmacovigilance
- However, fit-for-purpose assessments are needed to evaluate data relevance and quality
- The US Food and Drug Administration M14 guideline, published in March 2026, emphasizes a stepwise, iterative approach to designing pharmacoepidemiologic studies, beginning with the evaluation of whether the available data are fit for purpose¹
- This study compared the incidence of infections, hospitalizations, and mortality in patients with chronic lymphocytic leukemia (CLL) receiving covalent Bruton tyrosine kinase (cBTK) inhibitors across three RWD databases:
 - Symphony Health, an open claims database;
 - TriNetX, a general electronic health records (EHR) database; and
 - An oncology practice EHR database (name withheld due to contractual restrictions)

METHODS

- Adults with CLL receiving first-line (1L) cBTK inhibitor monotherapy, including zanubrutinib, acalabrutinib, and ibrutinib, from January 2019 to December 2023 were identified in the Symphony, TriNetX, and oncology practice EHR databases
- An overview of the three databases is shown in **Figure 1**

Figure 1. Overview of Databases

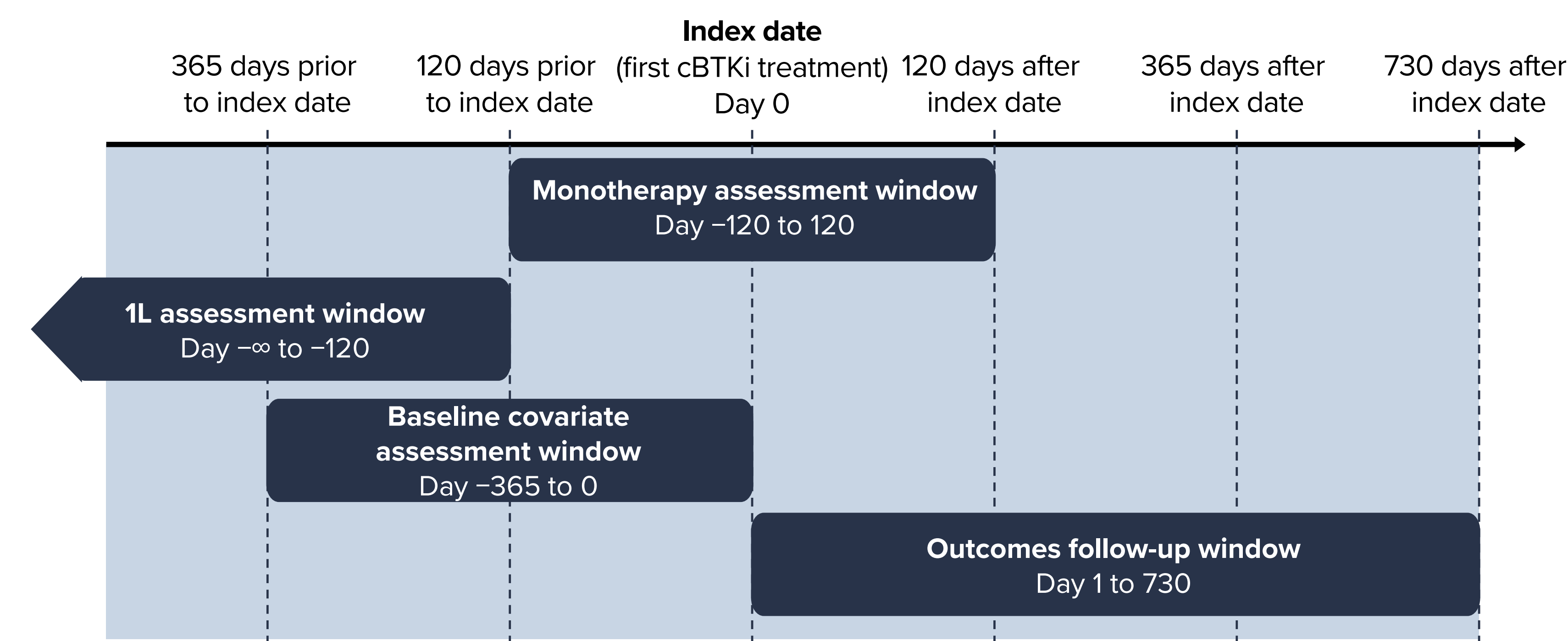
	Symphony Health	TriNetX Dataworks - USA	Oncology practice EHR
Data type	Open claims data	EHR data from HCOs	Oncology practice EHR data
Data source	Healthcare data including medical, hospital, and prescription claims data across the US and its territories	Federated data network from >70 HCOs covering all specialties	Patient records from oncology clinics and academic cancer centers
Strengths	<ul style="list-style-type: none"> • Large sample size • Longitudinal tracking • Detailed prescription data 	<ul style="list-style-type: none"> • Laboratory results and vitals available • Diverse healthcare settings • Mortality data available 	<ul style="list-style-type: none"> • Cancer-specific data, such as LOT, stage, and genomics, available • Mortality data available
Limitations	<ul style="list-style-type: none"> • Lack of clinical details • Lack of mortality data 	<ul style="list-style-type: none"> • Lack of LOT data and ECOG performance status 	<ul style="list-style-type: none"> • Limited to oncology setting • Potential underreporting of events

Abbreviations: ECOG, Eastern Cooperative Oncology Group; EHR, electronic health records; HCO, healthcare organization; LOT, line of therapy.

- The index date was defined as the date of the first recorded treatment
- Outcomes assessed included the incidence of overall infections (excluding COVID-19); specific infection subtypes (eg, pneumonia and shingles), identified using International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) codes; and use of injectable antimicrobials, identified using Healthcare Common Procedure Coding System codes within the 2-year follow-up period

- All-cause hospitalizations and all-cause mortality were also evaluated, where data were available
- Baseline patient demographics, including age, sex, and race and ethnicity, were described
- The study design schema is shown in **Figure 2**

Figure 2. Study Design Schema



Abbreviations: 1L, first line; cBTKi, covalent Bruton tyrosine kinase inhibitor.

RESULTS

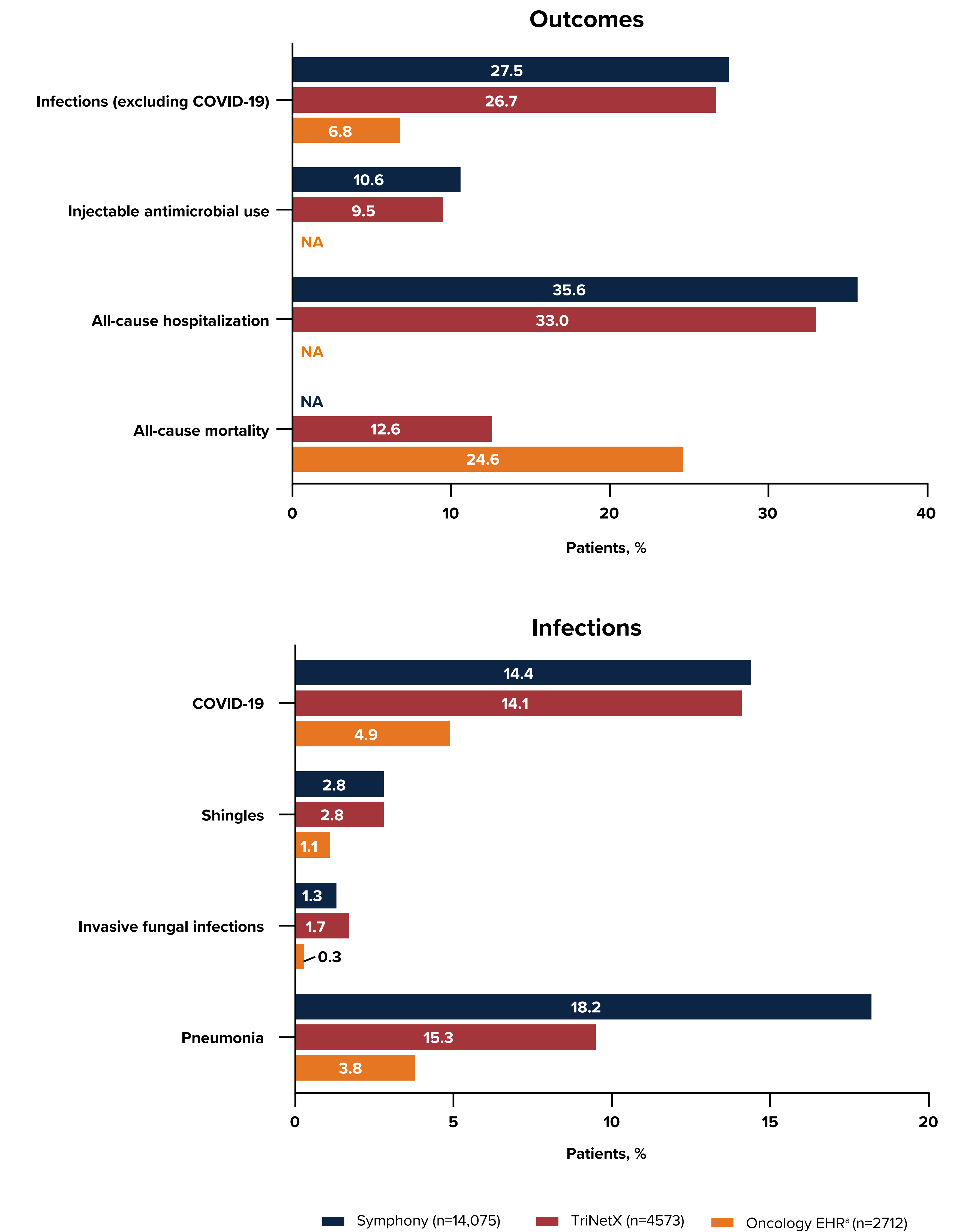
- In the 1L cBTK inhibitor monotherapy cohort, 14,075 patients were identified from Symphony, 4573 from TriNetX, and 2712 from the oncology practice EHR database
- Baseline patient demographics were similar between Symphony and TriNetX (**Table 1**)
 - The mean patient age was 69 years in the Symphony and TriNetX databases, slightly younger than the mean age of 71 years in the oncology practice EHR database
 - Both Symphony and TriNetX had a higher proportion of non-Hispanic White patients compared with the oncology practice EHR database
- The 2-year incidence of overall infections (excluding COVID-19) was comparable between Symphony and TriNetX, whereas the oncology practice EHR database had a lower incidence (27.5% vs 26.7% vs 6.8%, respectively) (**Figure 3**)
- Symphony and TriNetX had similar rates of injectable antimicrobial use (10.6% vs 9.5%, respectively) and all-cause hospitalization (35.6% vs 33.0%); data were not available in the oncology practice EHR database
- All-cause mortality was higher in the oncology practice EHR database compared with TriNetX (24.6% vs 12.6%, respectively); mortality data were not available in Symphony

Table 1. Baseline Demographics

	Symphony (n=14,075)	TriNetX (n=4573)	Oncology EHR database ^a (n=2712)
Age at index, mean (SD), years	69.3 (7.8)	69.6 (9.8)	71.2 (9.7)
Sex, n (%)			
Male	8631 (61.3)	2918 (63.8)	1611 (59.4)
Female	5444 (38.7)	1655 (36.2)	1101 (40.6)
Race and ethnicity, n (%)^b	n=11,347	n=4337	n=2564
White non-Hispanic	9466 (83.4)	3719 (85.8)	2019 (78.7)
Black non-Hispanic	1162 (10.2)	430 (9.9)	234 (9.1)
Asian non-Hispanic	147 (1.3)	77 (1.8)	24 (0.9)
Hispanic	544 (4.8)	95 (2.2)	90 (3.5)
Other	28 (0.2)	16 (0.4)	197 (7.7)

^aName withheld due to contractual restrictions. ^bThe proportions of race and ethnicity were calculated among patients with nonmissing data. Abbreviation: EHR, electronic health records.

Figure 3. Infections, Hospitalizations, and Mortality at the 24-Month Follow-Up



^aName withheld due to contractual restrictions. Abbreviations: EHR, electronic health records; NA, not available.

REFERENCE

1. US Food and Drug Administration. Accessed April 1, 2026. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/m14-general-principles-planning-designing-analyzing-and-reporting-non-interventional-studies-utilize>

ACKNOWLEDGMENTS

The authors thank the patients and their families, investigators, co-investigators, and the study teams at each of the participating centers. This study was sponsored by BeOne Medicines, Ltd. Editorial assistance was provided by Nucleus Global, an Inizio company, and supported by BeOne Medicines.

DISCLOSURES

LZ, QF, JT, AKA: Employment and may own stock: BeOne Medicines, Ltd.