Advancing Real-World Evidence in Diffuse Large B-Cell Lymphoma: Insights from EHR-Derived Data in the UK

Elsie Horne, PhD[1]; Patrycja Pluta, DVM PhD[2]; Christoph Buhl, MD[2]; Tamra Downing, MPH[3]; Harlan Pittell, PhD[3]; Blythe Adamson, PhD, MPH[3]
[1]Flatiron Health UK, London, UK; [2]Flatiron Health Germany, Berlin, Germany; [3]Flatiron Health, New York, NY, USA

Background

- Diffuse large B-cell lymphoma (DLBCL) is the most common aggressive Non-Hodgkin Lymphoma in the UK, and is associated with substantial morbidity and mortality. There is a growing need for region-specific real-world data (RWD) to understand variations in disease characteristics, healthcare delivery, and outcomes
- We aimed to describe the clinical characteristics and treatment patterns of patients with DLBCL in the UK using comprehensive, deidentified electronic health record (EHR) data

Methods

- **Data source:** The UK-based, EHR-derived deidentified Flatiron Health Research Database, comprising longitudinal patient-level data originated from a range of National Health Service (NHS) clinics and curated via technology-enabled abstraction[1,2]
- **Setting:** The study included 1265 patients at least 18 years of age, diagnosed with DLBCL between January 2011 and June 2025
- **Main outcome measures:** Descriptive disease characteristics, treatment patterns, and overall survival (OS) were estimated using Kaplan–Meier methods

Figure 1. Overview of the Process for EHR-Derived RWD Curation
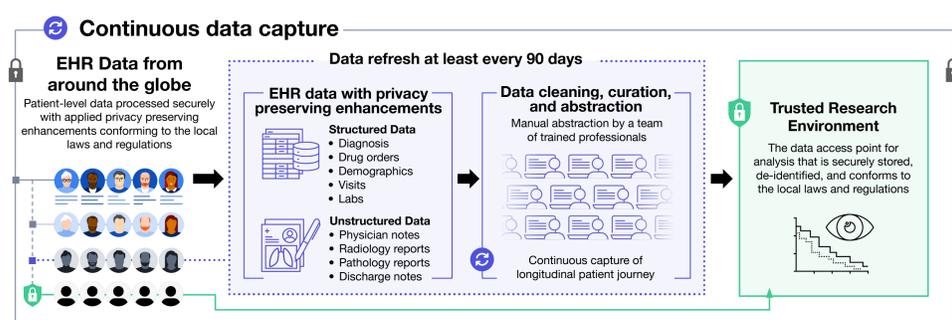
Results

- Data completeness was high: 76% patients had defined cell of origin (COO), 56% of patients had a recorded cytogenetic test with *MYC* being the most frequently tested
- Median OS from initial diagnosis in the full cohort was 6.0 years (95% CI: 5.0-7.1)

### Table 1. Demographics of Patients Included in the Study

Characteristic	Full cohort[a] (n = 1265)	Treated cohort[a] (n = 1035)
Age, median (IQR), y	69 (58, 77)	68 (57, 76)
Disease stage at initial diagnosis, No. (%)		
I	125 (16)	105 (15)
II	105 (14)	95 (14)
III	115 (15)	105 (15)
IV	415 (55)	375 (55)
Unknown	505	335
ECOG at index, No. (%)[b]		
0-1	35 (58)	535 (71)
2+	25 (42)	215 (29)
Unknown	1205	285
Cell of origin, No. (%)		
Germinal center B-cell	495 (52)	425 (54)
Activated B-cell	465 (48)	365 (46)
Unknown	305	245
Extranodal disease, No. (%)		
Yes	905 (71)	715 (69)
No	365 (29)	315 (31)
MYC cytogenetic test at index, No. (%)[c]		
MYC+	105 (15)	55 (15)
MYC-	615 (85)	305 (85)
Not tested	555	675

ECOG, Eastern Cooperative Oncology Group; IQR, interquartile range [a]To ensure privacy protection, all patient counts are rounded up to 10 with 5 subtracted [b]ECOG from 30 days prior to 7 days after index [c]Cytogenetic test results from 30 days prior to 90 days after index

Results (continued)

- From the 1035 treated patients:
  - 65 (6%) received a transplant, and 45 (4%) received CAR T cell therapy.
  - Median OS was 7.3 years (95% CI: 6.1-9.0) from start of treatment
  - Survival was lower in patients with more advanced stage, non-germinal B-cell as COO, and extranodal sites present at initial diagnosis (Figure 3)

Figure 2. Treatment Patterns Across First Three Lines of Therapy
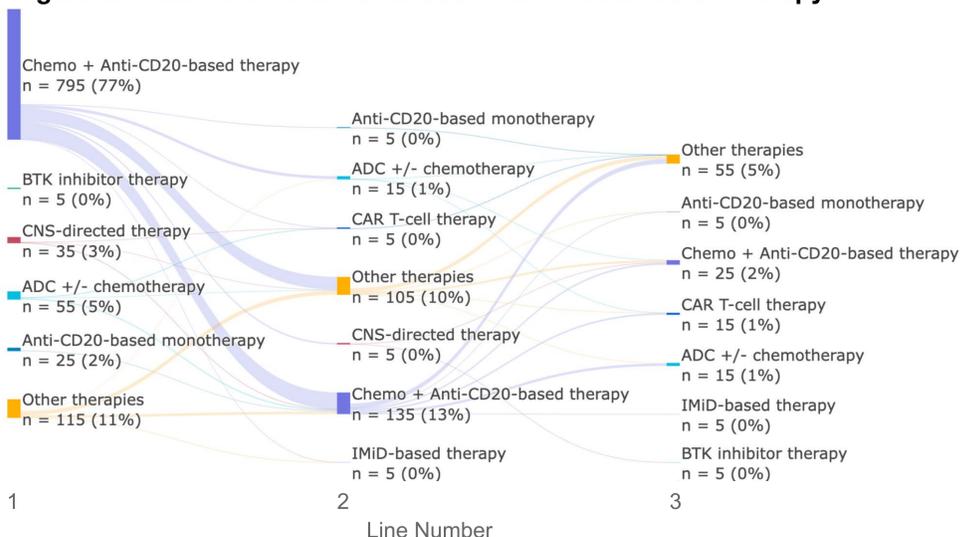
Figure 3. OS Stratified by COO, ECOG, and Presence of Extranodal Sites from Time of Treatment Initiation
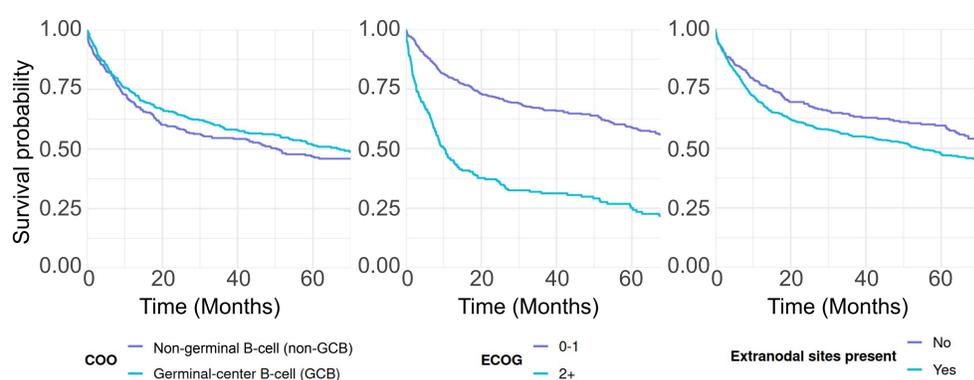
The UK DLBCL dataset reflects **expected clinical patterns** and offers insights to inform and optimise DLBCL management

Future Directions

- Harmonised oncology EHR-derived datasets have been developed at the regional level to enable cross-country comparisons of patients
- The quality and completeness of technology-enabled abstraction data in the UK are high and enable detailed and insightful analyses, going beyond the standard EHR-derived datasets

References

1. Flatiron Health. Database Characterization Guide. Flatiron.com. Published March 18, 2025. Accessed October 1, 2025. https://flatiron.com/database-characterization
2. Adamson B et al. *ESMO Real World Data Digit Oncol.* doi.org/10.1016/j.esmorw.2025.100113

Scan for abstract and digital poster