

Challenges for quality and utilization of real-world data for diffuse large B-cell lymphoma in REALYSA, a LYSA cohort

Tracking no: ADV-2023-010798R1

Herve Ghesquieres (Hopital Lyon Sud, Claude Bernard Lyon 1 University, France) Fanny Cherblanc (Lymphoma Academic Research Organisation, France) Aurelien Belot (Lymphoma Academic Research Organisation, France) Sophie Micon (Roche S.A.S, France) Krmo Bouabdallah (Service d'Hématologie clinique et Thérapie cellulaire, France) Cyril Esnault (Roche S.A.S., France) Luc-Matthieu Fornecker (University of Strasbourg, France) Katia Thokagevistik (Roche S.A.S., France) Maxime Bonjour (Service de Biostatistique-Bioinformatique, Pôle Santé Publique, Hospices Civils de Lyon, France) Fontanet Bijou (Institut Bergonie, France) Corinne Haioun (CHU Henri Mondor, France) Nadine Morineau (HOSPITAL OF LA ROCHE SUR YON, France) Loïc Ysebaert (Institut Universitaire du Cancer de Toulouse-Oncopole, France) Gandhi Damaj (Normandy University, Hematology Institute, France) Benoit Tessoulin (Nantes University Hospital, France) Stéphanie Guidez (CHU de Poitiers,) Franck Morschhauser (CHU Lille, ULR 7365 - GRITA - Groupe de Recherche sur les formes Injectables et les Technologies Associées, France) Catherine Thieblemont (AP-HP, Hôpital Saint-Louis, Hématologie, DMU DHI, F-75010 Paris, France, France) Adrien Chauchet (CHU, France) Remy Gressin (Institute for Advanced Biosciences, INSERM U1209/CNRS UMR 5309/Grenoble Alpes University, France) Fabrice Jardin (INSERM U1245 Unit, Rouen, France) Christophe Fruchart (Service d'Hématologie, France) Gaëlle Labouré (CH Libourne, France) Ludovic Fouillet (CHU Saint-Etienne, France) Pauline Lionne-Huyghe (CH Arras, France) Antoine Bonnet (CH Bretagne Atlantique, France) Laure Lebras (Centre Léon Bérard, France) Sandy Amorim (Hopital Saint Vincent, France) Cecile Leyronnas (Institut Daniel Hollard, France) Gaëlle Olivier (CH niort, France) Romain Guieze (CHU Clermont-Ferrand, France) Roch Houot (CHU Rennes, France) Vincent launay (CH st brieuc, France) Bernard Drénou (CH Mulhouse (GHRMSA), France) Olivier Fitoussi (Polyclinique Bordeaux Nord Aquitaine, France) Laurence Detournignies (CH Roubaix, France) Julie Abraham (CHU Limoges, France) Carole Soussain (Institut Curie, France) Florence Lachenal (CH Pierre Oudot, France) Gian Matteo Pica (Centre Hospitalier Métropole Savoie Chambéry, France) Patrick Fogarty (Lysarc, France) Pascale Cony-Makhoul (LYSARC, France) Adeline Bernier (Lymphoma Academic Research Organisation, France) Sandra Le Guyader-Peyrou (University of Bordeaux, INSERM, Team EPICENE, UMR1219, France) Alain MONNEREAU (University of Bordeaux, INSERM, Team EPICENE, UMR 1219, France) Frédéric Boissard (F. Hoffman-La Roche Ltd, Switzerland) Cédric Rossi (CHU Dijon, France) Vincent Camus (INSERM U1245 Unit, France)

Abstract:

Real-world data are essential to complement clinical trial (CT) data, but major challenges remain, like data quality. REal world dAta in LYmphoma and Survival in Adults (REALYSA) is a prospective non-interventional multicentric cohort started in 2018 (NCT03869619) including patients newly diagnosed with lymphoma in France. Herein is a proof-of-concept analysis on first-line DLBCL patients to (i) evaluate the capacity of the cohort to provide robust data through a multi-step validation process; (ii) assess the consistency of the results; (iii) conduct an exploratory transportability assessment of two recent phase 3 CT (POLARIX, SENIOR). The analysis population comprised DLBCL patients included before March 31st 2021, who received immunochemotherapy. 645 patients were included, for whom 3589 queries were generated, resulting in high data completeness (<4% missing data). Median age was 66 years (19-98) with mostly advanced-stage disease (472; 73%) and high international prognostic index (IPI) score (IPI 2-5, 486; 76%). Treatments were mostly R-CHOP (482; 75%) and R-miniCHOP (86; 13%). Estimated 1-year EFS and OS were 77.9% (95% CI: 73.8-81.4) and 90.0% (95% CI: 86.5-92.5), respectively (median follow-up: 9.9 months). Regarding transportability, when applying trials' main inclusion criteria (age, PS, IPI), outcomes seemed comparable between REALYSA patients and standard arms of POLARIX (1-year PFS 79.8% (95% CI, 75.9-83.6) vs. 79.8% (95% CI, 73.9-84.4)) and SENIOR (1-year EFS à 64.5% (95% CI: 47.8-77.0) vs. 60.0% (95% CI: 50.8-68.1)). With its rigorous data validation process, REALYSA program provides high-quality RWD, thus constituting a platform for numerous scientific purposes.-

Conflict of interest: COI declared - see note

COI notes: H. Ghesquieres: consultancy and advisory boards: Roche, Gilead, BMS, Abbvie S. Micon, C. Esnault, K. Thokagevistik, F. Boissard: Roche Employee F. Morschhauser: consultancy and advisory boards: Roche V. Camus: consultancy and honoraria: Roche, Janssen Celgene-BMS; Research and travel grants: Celgene-BMS.

Preprint server: No;

Author contributions and disclosures: Conception and design: H. Ghesquieres, F. Cherblanc, A. Belot, S. Micon, C. Esnault, K. Thokagevistik, A. Bernier, F. Boissard, V. Camus Provision of study materials or patients: H. Ghesquieres, K.-K. Bouabdallah, L.-M. Fornecker, F. Bijou, C. Haioun, N. Morineau, L. Ysebaert, G. Damaj, B. Tessoulin, S. Guidez, F. Morschhauser, C. Thiéblemont, A. Chauchet, R. Gressin, F. Jardin, C. Fruchart, G. Labouré, L. Fouillet, P. Lionne-Huyghe, A. Bonnet, L. Lebras, S. Amorim, C. Leyronnas, G. Olivier, R. Guieze, R. Houot, V. Launay, B. Drenou, O. Fitoussi, L. Detourmignies, J. Abraham, C. Soussain, F. Lachenal, G. Pica, S. Le Guyader-Peyrou, A. Monnereau, C. Rossi, V. Camus Collection and assembly of data: F. Cherblanc, A. Belot, M. Bonjour, P. Fogarty, P. Cony-Makhoul, A. Bernier Statistical analysis: A. Belot, P. Fogarty Manuscript writing: H. Ghesquieres, F. Cherblanc, A. Belot, A. Bernier Final approval of manuscript: all authors approved the final version of the manuscript and are accountable for all aspects of the work

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Clinical trial registration information (if any):

1 **Challenges for quality and utilization of real-world data for diffuse large B-cell**
2 **lymphoma in REALYSA, a LYSA cohort**

3 **Author names**

4 Hervé Ghesquieres¹, Fanny Cherblanc², Aurélien Belot², Sophie Micon³, Krimeo K.
5 Bouabdallah⁴, Cyril Esnault³, Luc-Matthieu Fornecker⁵, Katia Thokagevistik³, Maxime
6 Bonjour², Fontanet Bijou⁶, Corinne Haioun⁷, Nadine Morineau⁸, Loïc Ysebaert⁹, Gandhi
7 Damaj¹⁰, Benoit Tessoulin¹¹, Stéphanie Guidez¹², Franck Morschhauser¹³, Catherine
8 Thiéblemont¹⁴, Adrien Chauchet¹⁵, Rémy Gressin¹⁶, Fabrice Jardin¹⁷, Christophe Fruchart¹⁸,
9 Gaëlle Labouré¹⁹, Ludovic Fouillet²⁰, Pauline Lionne-Huyghe²¹, Antoine Bonnet²², Laure
10 Lebras²³, Sandy Amorim²⁴, Cécile Leyronnas²⁵, Gaëlle Olivier²⁶, Romain Guieze²⁷, Roch
11 Houot²⁸, Vincent Launay²⁹, Bernard Drenou³⁰, Olivier Fitoussi³¹, Laurence Detourmignies³²,
12 Julie Abraham³³, Carole Soussain³⁴, Florence Lachenal³⁵, Gian Matteo Pica³⁶, Patrick
13 Fogarty², Pascale Cony-Makhoul², Adeline Bernier², Sandra Le Guyader-Peyrou³⁷, Alain
14 Monnereau³⁷, Frédéric Boissard³⁸, Cédric Rossi³⁹, Vincent Camus¹⁷

15 **Author affiliations**

16 ¹Department of Hematology, Hopital Lyon Sud, Claude Bernard Lyon 1 University, Pierre
17 Benite, France

18 ²Lymphoma Academic Research Organisation (LYSARC), Hopital Lyon Sud, Pierre Benite,
19 France

20 ³Roche S.A.S., Boulogne-Billancourt, France

21 ⁴Hematology and Cell Therapy Department, University Hospital of Bordeaux, Bordeaux,
22 France

23 ⁵Institut de Cancérologie Strasbourg Europe (ICANS) and University of Strasbourg,
24 Strasbourg, France

25 ⁶Department of Hematology, Institut Bergonie, Bordeaux, France

26 ⁷Lymphoid Malignancies Unit, AHPH, Hopital Henri Mondor, Creteil, France

27 ⁸CHD Vendée, La Roche-sur-Yon, France

28 ⁹IUCT Oncopole, Toulouse, France

29 ¹⁰Hematology Institute of Basse Normandie, CHU Caen, Caen, France

30 ¹¹Department of Hematology, CHU Nantes, Nantes, France

31 ¹²Department of Hematology, CHU Poitiers, Poitiers, France

32 ¹³Department of Hematology, Université de Lille, CHU Lille, ULR 7365 - GRITA - Groupe de
33 Recherche sur les formes Injectables et les Technologies Associées, Lille, France

34 ¹⁴Université Paris Cité, Assistante Publique – Hôpitaux de Paris, Hôpital Saint-Louis, Service
35 d'Héματο-Oncologie, Paris, France

36 ¹⁵Department of Hematology, CHU Besançon, Besançon, France

37 ¹⁶Department of Hematology, CHU Grenoble, Institute for Advanced Biosciences, INSERM
38 U1209/CNRS UMR 5309/Grenoble Alpes University, Grenoble, France

39 ¹⁷Department of Clinical Hematology, INSERM U1245 Unit, Centre Henri Becquerel, Rouen,
40 France

41 ¹⁸Department of Hematology, CH Dunkerque, Dunkerque, France

42 ¹⁹Department of Hematology, CH Libourne, Libourne, France

43 ²⁰Department of Hematology, CHU Saint Etienne, Saint Etienne, France

44 ²¹Department of Hematology, CH Arras, Arras, France

45 ²²Department of Hematology, CH Bretagne Atlantique, Vannes, France

46 ²³Department of Hematology, Leon Berard Cancer Center, Lyon, France

47 ²⁴Department of Hematology, Hopital Saint Vincent de Paul, Lille

48 ²⁵Department of Hematology, GHM de Grenoble, Grenoble, France

49 ²⁶Department of Hematology, CH Niort, Niort, France

50 ²⁷Department of Hematology, CHU Clermont Ferrand, Clermont Ferrand, France
51 ²⁸Department of Hematology, CHU Rennes, Rennes, France
52 ²⁹Department of Hematology, CH Saint Briec, Saint Briec, France
53 ³⁰Department Hematology, Groupe Hospitalier Mulhouse Sud Alsace, Mulhouse, France
54 ³¹Department of Hematology, Polyclinique Bordeaux Nord Aquitaine, Bordeaux, France
55 ³²Department of Hematology, CH Roubaix, Roubaix, France
56 ³³Department of Hematology, CHU Limoges, Limoges, France
57 ³⁴Department of Hematology, Institut Curie, Saint-Cloud, France
58 ³⁵Department of Hematology, CH Pierre Oudot, Bourgoin-Jallieu, France
59 ³⁶Department of Hematology, CH Metropole Savoie, Chambéry, France
60 ³⁷Registre des Hémopathies Malignes de la Gironde, Institut Bergonié, University of
61 Bordeaux, Inserm, Team EPICENE, UMR 1219, 33000, Bordeaux, France
62 ³⁸F. Hoffmann-La Roche Ltd, Basel, Switzerland
63 ³⁹Department of Hematology, CHU Dijon Bourgogne, Dijon, France
64

65 **Corresponding author**

66 Hervé Ghesquières, MD, PhD, Département d'Hématologie, Hôpital Lyon Sud, Hospices
67 Civils de Lyon, Pierre-Bénite, France; Tel +33 4 78 86 43 01; Fax +33 4 78 86 43 54; Email:
68 herve.ghesquieres@chu-lyon.fr

69

70 **Data sharing statement**

71 The data that support the findings of this study are available from the corresponding author
72 upon reasonable request.

73 **Running-title:** Real-world data for diffuse large B-cell lymphoma

74 **Keywords:** Diffuse large B-cell lymphoma, real-world cohort data, REALYSA, data quality,
75 clinical trial transportability

76 Abstract word count: 250

77 Word count: 4686

78 Table: 2

79 Figures: 3

80 References: 67

81 Supplementary Tables/Figures: 4/2

82

83 This study was presented in Poster form at the 27th Annual Congress of the European
84 Hematology Association (EHA), Vienna, Austria, 9-12 June 2022.

85

86 **Key points:**

- 87
- 88
- 89
- 90
- The REALYSA cohort is a source of real-world data of high quality for lymphoma thanks to a multi-step rigorous data validation process.
 - Effectiveness results on first line DLBCL patients seem consistent with literature and recent clinical trials.

91

92 **Abstract**

93 Real-world data are essential to complement clinical trial (CT) data, but major challenges
94 remain, like data quality. REal world dAta in LYmphoma and Survival in Adults (REALYSA) is
95 a prospective non-interventional multicentric cohort started in 2018 (NCT03869619) including
96 patients newly diagnosed with lymphoma in France. Herein is a proof-of-concept analysis on
97 first-line DLBCL patients to (i) evaluate the capacity of the cohort to provide robust data
98 through a multi-step validation process; (ii) assess the consistency of the results; (iii) conduct
99 an exploratory transportability assessment of two recent phase 3 CT (POLARIX, SENIOR).
100 The analysis population comprised DLBCL patients included before March 31st 2021, who
101 received immunochemotherapy. 645 patients were included, for whom 3589 queries were
102 generated, resulting in high data completeness (<4% missing data). Median age was 66
103 years (19–98) with mostly advanced-stage disease (472; 73%) and high international
104 prognostic index (IPI) score (IPI 2-5, 486; 76%). Treatments were mostly R-CHOP (482;
105 75%) and R-miniCHOP (86; 13%). Estimated 1-year EFS and OS were 77.9% (95% CI:
106 73.8–81.4) and 90.0% (95% CI: 86.5–92.5), respectively (median follow-up: 9.9 months).
107 Regarding transportability, when applying trials' main inclusion criteria (age, PS, IPI),
108 outcomes seemed comparable between REALYSA patients and standard arms of POLARIX
109 (1-year PFS 79.8% (95% CI, 75.9–83.6) vs. 79.8% (95% CI, 73.9–84.4)) and SENIOR (1-
110 year EFS à 64.5% (95% CI: 47.8–77.0) vs. 60.0% (95% CI: 50.8–68.1)). With its rigorous
111 data validation process, REALYSA program provides high-quality RWD, thus constituting a
112 platform for numerous scientific purposes.

113 Introduction

114 In France, lymphomas represent most of hematopoietic cancers, and are the sixth and
115 seventh most frequent cancers in men and women, respectively.¹ Diffuse large B-cell
116 lymphoma (DLBCL) is the most frequent subtype with approximately 5 000 new cases every
117 year.¹ Over the past two decades, the prognosis of DLBCL patients has improved as a result
118 of successive interventional trials investigating the intensity of chemotherapy, combinations
119 of chemotherapy with monoclonal antibodies or targeted therapies, and more recently the
120 evaluation of chimeric antigenic receptor T (CAR-T) cells and bispecific monoclonal
121 antibodies in relapsed/refractory (R/R) patients.² However, while interventional trials remain
122 the gold standard for evaluating new drugs or therapeutic strategies,² recent studies have
123 highlighted limited trial participation.^{3,4} Indeed, less than 10% of patients are included in
124 prospective trials, due to many factors, including trial availability in care center, organ
125 function-based criteria, comorbidities, performance status (PS), age.⁵⁻⁷ Using recent DLBCL
126 prospective trial criteria, Khurana *et al.* estimated that, on the basis of organ function criteria
127 alone, 9-24% of newly diagnosed DLBCL real-world patients would be excluded from trial
128 participation.³ Moreover, Loh *et al.* showed that the number of eligibility criteria in DLBCL
129 trials has increased over the past 30 years, and that less than 50% of real-world patients are
130 actually eligible for the most recent trials.⁴ Lastly, technical requirements (*e.g.* mandatory
131 imaging, tumor biopsy to screen for biomarkers) are time-consuming, can be complex to set-
132 up, and might delay protocol treatment initiation. Recently, in order to modernize trial
133 enrollment criteria with the aim to shorten interval between diagnosis and treatment initiation
134 (DTI), one US group of experts revised 31 eligibility criteria commonly used in DLBCL RCTs
135 in R-CHOP era and identified 13 essential criteria, 9 unnecessary criteria, while no
136 consensus could be reached for the 9 remaining ones.⁸ The DTI is a strong prognostic factor
137 in DLBCL, patients with a short DTI having worse prognostic factors and unfavorable
138 outcomes,⁹ thus creating a potential significant selection bias in complex clinical trials (CT).

139 Overall, these studies highlight the need for a thorough assessment of the generalizability of
140 CT results.

141 To this end, and in a context of ever greater accessibility to electronic health records,
142 increasing focus is being given to real-world data (RWD).^{10,11} With rigorously managed RWD,
143 non-interventional study designs can provide results that are a notable source of additional
144 information to that from randomized clinical trials (RCT).^{12,13} In the field of lymphoma, RWD
145 have, for instance, proven useful in developing clinical scores, defining new outcome
146 endpoints, assessing the role of imaging in patient follow-up, and addressing long-term
147 toxicities.¹⁴⁻¹⁶ Interestingly, some prognostic scores and outcome endpoints were developed
148 using RWD and subsequently validated on data from interventional trials.¹⁷⁻²¹ RWD may
149 come from various sources (*e.g.* institution databases, administrative databases from health
150 insurance, registries, cohorts, or directly from patients via connected electronic devices).^{12,13}
151 The main challenge is to optimize data quality by maximizing information while minimizing
152 missing data, measurement errors or lost to follow-up patients.

153 To generate reliable RWD in lymphoma, the Lymphoma Study Association (LYSA)
154 launched in 2018 an observational multicentric prospective cohort in metropolitan France -
155 REALYSA (REal world dAta in LYmphoma and Survival in Adults) - to evaluate the real-world
156 prognosis of the seven most common lymphoma subtypes.²²

157 In the present study, we report the results of a proof-of-concept analysis on first-line
158 treatment (1L) DLBCL patients in the REALYSA cohort to: (i) evaluate the capacity of a real-
159 world program on lymphoma in France to provide robust data through a specific multi-step
160 data validation system; (ii) assess whether the characteristics of the population, clinical
161 practices and estimated effectiveness are consistent with what is expected in real life; and
162 (iii) conduct an exploratory assessment of the transportability of recent prospective phase 3
163 trials evaluating novel 1L therapeutic agents in DLBCL patients.^{23,24}

164

166 **Methods**

167 **REALYSA cohort**

168 The REALYSA study is registered on ClinicalTrials.gov (NCT number: NCT03869619),
169 approved by ethics committee, and written informed consent was obtained from all patients.
170 Patients were informed of this specific analysis through a dedicated webpage on the LYSA
171 website (<https://lymphoma-research-experts.org/>). The cohort methodology has been
172 described elsewhere.²² Briefly, patients are prospectively recruited in one of the 35
173 hematology centers after signing an informed consent form. Inclusion criteria are the
174 following: aged over 18; diagnosed with lymphoma in the previous 6 months; and before
175 treatment initiation. Patients are managed according to physician's choice and there are no
176 compulsory visits for the study. Clinical and treatment data at diagnosis, for first-line
177 treatment and potential subsequent treatment lines are extracted from medical records.

178 **Data entry checks**

179 Real-time data checks were programmed within the electronic data capture (EDC) system, in
180 order to avoid obvious mistakes in data entry: (i) for all numerical variables, a pre-defined
181 range was programmed and if a value out of pre-defined range is entered, a warning
182 message appears on the screen and the value has to be validated; (ii) for all variables, an
183 expected format is defined, and if the value is out of the expected format, data entry is
184 blocked and the value has to be modified.

185 **Data validation process and quality control**

186 A multi-step data validation process inspired by RCT processes but adapted to the RWD
187 constraints was set up to check the internal validity of the data (Supplemental Figure 1). The
188 objective is to maximize automation of consistency tests to minimize human resources.

189 First (Level 1), data completeness is assessed with automated tests on pre-identified data,
190 within the EDC system. Variables are sorted into three levels of importance that guide the
191 insistence to obtain the data through queries in case of missing data. Simple consistency

192 checks are also performed automatically within the EDC system to identify different types of
193 inconsistency, such as chronological inconsistency (e.g. biopsy date after start date of
194 treatment), discrepancies between treatment response and the Deauville score, or Ann Arbor
195 staging inconsistent with reported nodal localizations. Queries are automatically generated
196 and submitted by the data manager. Some 17 000 completeness checks (305 variables
197 throughout) and over 35 000 basic consistency checks (645 checks reproduced) were
198 programmed.

199 Second (Level 2), more advanced consistency checks are performed using SAS software. All
200 inconsistencies are reviewed by the operational study team and queries generated if deemed
201 necessary. For instance, inconsistencies in the chronology of response evaluation can be
202 identified (e.g. a patient reported in stable disease after being in complete response, see
203 Supplemental Table 1 for additional examples). When consistency checking cannot easily be
204 programmed, a manual review of listings is performed. For example, disease stage for
205 patients with extra-nodal involvement or treatment patterns are reviewed manually. This
206 second level of validation includes 41 advanced checks (baseline characteristics, 19;
207 treatment and follow-up, 22), manual review of 19 free-text entries and a review of two
208 listings.

209 Last (Level 3), to assess the plausibility of patient care pathway given the baseline
210 characteristics, data are reviewed by LYSA clinicians using patient profiles generated using
211 R software (Figure 1). This *ad hoc* tool was developed to automatically provide a summary of
212 each patient (characteristics, therapeutic management, evolution over time), allowing time-
213 efficient data validation. Additional queries can be sent if deemed necessary.

214 Simultaneously, to check the consistency between data in the eCRF versus patients' medical
215 records, external quality control (QC) is being implemented for 5% of the patients (included
216 for ≥ 1 year and considered as validated according to the automatic tests). QC focuses on
217 critical data selected according to 3 criteria: (i) regulatory impact (consent form and
218 inclusion/exclusion criteria), (ii) impact on endpoints (diagnosis date, start date for each

219 treatment line, progression date, date of death), (iii) robustness of the data (an imaging exam
220 date will be chosen over a clinical exam date) (Supplemental Table 2 for full list of variables).
221 A concordance rate per patient/center/sample will be used to assess the overall quality of the
222 database and to trigger corrective actions if deemed necessary.

223 As described in the study protocol article,²² an approach based on continuous improvement
224 was implemented with centers, with various tools and regular meetings, to continuously
225 improve data quality and optimize human time.

226 **DLBCL population**

227 The following patients were considered for the study: patients included in REALYSA before
228 March 31st 2021, diagnosed with DLBCL according to World Health Organization
229 classification 2016, and treated with immunochemotherapy for their first line (1L) (IC-treated
230 population). To focus the analysis on a real-world population, a “non-interventional IC-treated
231 population” (thereafter referred to as “analysis population”) was selected by excluding
232 patients enrolled in a 1L interventional CT.

233 Additionally, to assess the transportability²⁵ of the control arms of two recent phase 3 trials
234 (*i.e.* POLARIX²³ and SENIOR²⁴ trials), “POLARIX-like” (P-L) and “SENIOR-like” (S-L)
235 populations were defined by selecting patients who were treated with the standard of care of
236 the control arm (R-CHOP regimen for P-L and R-miniCHOP regimen for S-L) and met the
237 main inclusion criteria of the trials (POLARIX: age, 18-80 years at treatment initiation;
238 baseline international prognostic index (IPI) score, 2 to 5; ECOG PS score, 0 to 2 / SENIOR:
239 age ≥80 years; Ann Arbor stage, II to IV; PS, 0 to 2).

240 **Endpoints of interest**

241 The primary endpoint was event-free survival (EFS) defined as the time between the start of
242 the first line to either a progression, relapse, new treatment line or death, whichever occurred
243 first. We also investigated progression-free survival (PFS), overall survival (OS) and end of
244 treatment (EOT) response.

245 **Statistical Analysis**

246 Response probabilities were expressed as percentages with 95% confidence intervals (CI)
247 calculated according to Exact Pearson-Clopper method. For PFS, EFS and OS, Kaplan-
248 Meier estimator was used to estimate probabilities of occurrence of a given endpoint at
249 specific time points (with their 95% CI). Median follow-up was estimated using reverse
250 Kaplan-Meier method. To assess the transportability of standard arms of phase 3 trials,
251 patient characteristics and outcomes for the “POLARIX-like” and “SENIOR-like” populations
252 were described. The analysis was conducted by A.Bel. and P.F. in January 2022, using data
253 exported on November 15th, 2021, and with a data cut-off set at June 30th, 2021. Access to
254 primary data was possible for all academic authors.

255

256 **Results**

257 **Data validation and quality report**

258 ***Whole REALYSA population***

259 For the whole REALYSA population, over an 18-month period between 2021 and mid-2022
260 when all validation tests were running — during which 2700 new patients were recruited and
261 1300 patients were in follow-up — 7642 queries were sent (Table 1), with the following
262 distribution (Supplemental Figure 1): 50% were automated queries generated within the EDC
263 software (Level 1), and 50% resulted from advanced consistency testing, using SAS and
264 manual or medical reviews (Levels 2 and 3). Half of the queries (56%) concerned the
265 baseline-data section, 30% treatment data, and 14%, follow-up data. Following a query, data
266 was modified in approximately 75% of cases.

267

268 ***DLBCL population of this study***

269 Regarding the analysis population, 3589 queries were transmitted, among which 99% were
270 answered by the centers before data export. The analysis database comprised information
271 on 100% of patients for most variables, including lymphoma subtype, date of diagnosis, and

272 the presence/absence of extra-nodal involvement (Table 2). Disease stage was missing for
273 only one patient (0.2%). IPI class was available for 96% of patients, and PS score and
274 lactate dehydrogenase (LDH) levels for >97% of patients.

275

276 **DLBCL patient characteristics**

277 The flowchart is presented in Figure 2. Overall, 700 DLBCL patients were selected from the
278 REALYSA cohort. Three patients (age: 87, 87 and 93 years) who had not received IC were
279 excluded (treated with palliative treatment with rituximab monotherapy and oral
280 cyclophosphamide), leaving an IC-treated population of 697 patients. Of these, 52 (7.5%)
281 enrolled in interventional CT were excluded, leaving an analysis population of 645 patients.
282 Patients were recruited from 34 centers across 61 French departments. The majority (543,
283 78% of the IC-treated population) were recruited from University Hospitals, 71 (10%) from
284 General Hospitals, 70 (10%) from Cancer Centers and 13 (2%) from Private Clinics.

285 Overall, patient characteristics in the IC-treated population (N=697) and the analysis
286 population (N=645) were comparable (Table 2). Regarding analysis population, median age
287 was 66 years (range 19-98), 344 patients were male (53%), with advanced-stage disease
288 (Ann-Arbor stage III/IV, 472 patients; 73%), extra-nodal locations (499; 77%), elevated LDH
289 levels (402; 64%) and high IPI score (IPI 2-5, 486; 76%). The main histological subtypes
290 were DLBCL (480; 74%), high-grade B-cell lymphoma (81; 13%) and primary mediastinal B-
291 cell lymphoma (PMBL) (45; 7%). PET/CT scan was performed for most patients for initial
292 work-up (606 out of 642 patients with available data; 94%).

293

294 **Treatment patterns**

295 For almost all patients (630/645; 98%), 1L therapy was anthracycline-based, with either (i) R-
296 CHOP (rituximab, doxorubicin, cyclophosphamide, vincristine, prednisone; 482; 75%), (ii) R-

297 miniCHOP (reduced-dose R-CHOP; 86; 13%), or (iii) high-dose anthracycline-based regimen
298 (62; 10%, of whom 57 (92%) with R-ACVBP [rituximab, doxorubicin, cyclophosphamide,
299 vindesine, bleomycin, prednisone]). Only 15 patients (2%) received a non-anthracycline-
300 based regimen (see Supplemental Table 3 for details). The median age was 65 years (range,
301 20–90) for patients receiving R-CHOP, 83 years (range, 71–95) for patients receiving R-
302 miniCHOP and 40 years (range, 19–67) for patients receiving high-dose anthracycline
303 regimens (Table 2). Among patients treated with high-dose anthracycline regimens (N=62),
304 most were treated in University Hospitals (55; 89%) and one third (21; 34%) had a PMBL
305 subtype. Only nine patients (1.4%) underwent consolidation therapy with autologous stem
306 cell transplantation (ASCT). Of these, five were PMBL patients, two had central nervous
307 system (CNS) involvement at diagnosis and received an R-COPADEM (rituximab, high-dose
308 methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisone) regimen before
309 ASCT, and the remaining two patients received ASCT after R-CHOP in combination with
310 high dose methotrexate according to physician's choice. Radiotherapy was used as
311 consolidation therapy in 19 patients (3%), of whom eight (42%) had a localized disease (I/II)
312 and seven (37%) a gonadal involvement. Regarding the number of cycles of R-CHOP, 278
313 patients (58%) received six cycles and 102 (21%) eight cycles (Supplemental Table 4).
314 Among the 63 patients with an age-adjusted IPI (aaIPI) of 0, 24 (38%) received four cycles of
315 R-CHOP. Overall median DTI was 26.0 days (range, 0–132). Median DTI was similar for
316 patients receiving R-CHOP and R-miniCHOP (26.0 (range, 0–132) and 25.5 (range, 4–84)
317 days, respectively), shorter for patients receiving high-dose anthracycline regimen (19.0 days
318 (range, 0–60)) and longer for patients receiving non-anthracycline based regimen (35.0 days
319 (range, 6–71)).

320

321 **Treatment response**

322 Treatment response at the end of 1L therapy was documented for 603 patients (94%). Of
323 these, 483 (80%, 95% CI: 76.7 - 83.2) had a complete response, 51 (9%, 95% CI: 6.4 - 11.0)

324 a partial response (overall response rate [ORR] of 89%, 95% CI: 85.7 - 91.0), seven (1.2%,
325 95% CI: 0.5 - 2.4) stable disease, and 62 (10%, 95% CI: 8.0 - 13.0) progressive disease. For
326 most patients (556; 92%), treatment response was assessed using PET/CT-scan.

327 **Outcomes**

328 Median follow-up from treatment initiation was 9.9 months (range, 0.4–30.5) for the analysis
329 population. Among the 645 patients included in outcome evaluation, 123 EFS events were
330 documented (new treatment initiation for 15 patients [2.3%], progression/relapse for 86
331 patients [13.3%], death for 22 patients [3.4%]). Median survival was not reached. The 1-year
332 OS rate was 90.0% (95% CI: 86.5 - 92.5). The 1-year EFS rate was 77.9% (95% CI: 73.8 -
333 81.4). The prognostic value of IPI and aalPI²⁶ seems to be reproduced in this real-world
334 dataset (Figure 3), with the limitation of the short follow-up.

335

336

337 **Patient representativeness: Comparison with national data registry**

338 As an indicator of representativeness, the analysis population was compared to national
339 DLBCL incidence data,¹ according to age group and sex (Supplemental Figure 2). It
340 suggests a rather good level of comparability in terms of age distribution, though with a slight
341 under-representation of older patients (over 80), for women in particular.

342

343 **Transportability of DLBCL population data: clinical trials versus REALYSA**

344 A “POLARIX-like” (P-L) population (N=320) was extracted from REALYSA selecting DLBCL
345 patients who fulfilled the main inclusion criteria of the POLARIX trial and were treated with R-
346 CHOP.²³ This P-L population seemed comparable to the R-CHOP arm population of the
347 POLARIX trial (Table 2): median age, 67 versus 66 years; IPI score, 3-5 in 65.7% versus
348 62%, respectively. Median DTI was 26.0 days (range, 0–132) for P-L population and 27 days
349 in the POLARIX R-CHOP arm. The ORR was 89% in real-world P-L population and 84% in
350 the POLARIX R-CHOP arm. The 1-year PFS rates were 79.8% (95% CI, 73.9 - 84.4) in the
351 R-CHOP control arm of POLARIX²³ and 79.8% (95% CI, 75.9 - 83.6) in the patients of the P-
352 L population.

353 Regarding “SENIOR-like” (S-L) population (N=59), characteristics also seemed rather
354 comparable with patients from the control arm of the SENIOR trial²⁴ (Table 2): median age,
355 83.4 vs 83.0 years; PS score, 2-4 in 28.8% versus 28.0%; IPI score, 3-5 in 79.0% versus
356 75.0%, respectively. Median DTI was 27.0 days (range, 5–80) for S-L population and 33
357 days (range, 8–89) for R-miniCHOP arm of SENIOR trial. The 1-year EFS was 64.5% (95%
358 CI: 47.8 - 77.0) and 60.0% (95% CI: 50.8 - 68.1) in the patients of the S-L population and in
359 the R-miniCHOP arm of SENIOR trial²⁴, respectively. The 1-year OS was 78.3% (95% CI:
360 61.4 - 88.5) and 78.5% (95% CI: 70.2 - 84.7) in S-L population and R-miniCHOP arm of
361 SENIOR trial²⁴, respectively.

362

364 **Discussion**

365 This proof-of-concept analysis of REALYSA DLBCL patients demonstrate how a nationwide
366 prospective real-world cohort can provide comprehensive robust data on baseline
367 characteristics of patients and treatment effectiveness. Ensuring data quality in a large
368 prospective observational cohort such as REALYSA is extremely challenging. Tailored
369 processes are essential to ensure effective data management and validation. Furthermore,
370 data collection needs to be regularly updated to keep up with changes in clinical practice. For
371 example, when REALYSA was initiated (2018), CAR T-cell therapy was only used as third-
372 line therapy of DLBCL in CT settings whereas it is now positioned as second-line treatment
373 for R/R DLBCL patients.²⁷⁻²⁹ The eCRF has been tailored to include information regarding
374 the implementation of these new therapies in daily practice. The data validation process is
375 based on experience from CT and adapted to the constraints of high inclusion rates (~140
376 patients per month) from multiple recruitment sites, limited human resources, and the
377 necessity to initiate data analysis before the end of the study. Efforts were focused primarily
378 on internal validation of critical data to limit the number of queries but still maintain an
379 appropriate level of quality. For the analysis population (N=645), 3589 queries were sent to
380 centers, leading to data adjustments in most cases, and consequently to notable
381 improvements in data quality. Despite the inherent limitations of comparing trials with
382 different datasets, it seems that the number of queries generated in our study was similar to
383 the one of an academic phase III trial (LNH09-1B, N=650 patients, 5180 queries) and 7-9
384 times lower than those of industrial phase III trials (REMARC trial,³⁰ N=650 patients, 31 756
385 queries; GAINED trial,³¹ N=671 patients, 26 152 queries). Thus, although the REALYSA
386 validation process may be less stringent than the processes applied to pharmaceutical
387 industry-sponsored CT, it seems rigorous enough to generate meaningful robust data with
388 very few missing data on key variables (<4%). Data validation processes require
389 considerable resources but have a strong impact on data quality. Depending on the situation,
390 dedicating such resources for data validation may not always be feasible, resulting in highly

391 variable levels of RWD data quality. With the recent increase in the use of RWD as a
392 complement to CT data for regulatory decision-making, the challenge is to ensure that data
393 quality is suited to confidently inform on drug use or treatment effectiveness,³²⁻³⁴ in particular
394 if RWD are to be used for marketing authorizations and not only for post-marketing studies.
395 In line with this report, international guidelines stress the importance of clear descriptions of
396 data verification processes.^{35,36} This is however rarely done in the literature.³⁷⁻³⁹ Yet, our
397 process could be improved, and notably in terms of verification of source data (*i.e.* medical
398 records). External QC to compare source data with the eCRF data is scheduled for 2023.
399 Future processes will also include reviews of pathology reports by hematopathologists from
400 LYSA and LymphoPath networks to reinforce diagnostic accuracy.⁴⁰ Such improvements are
401 key to improve data quality and increase our knowledge of lymphoma biology in real-world
402 settings.⁴¹⁻⁴⁴

403 The current analysis shows trends in DLBCL patient care in hematology departments in
404 France. Virtually all patients received curative-intent treatment with only three and 15
405 patients treated without IC and without anthracycline-based chemotherapy, respectively.
406 Most patients (75%) received R-CHOP, 13% R-miniCHOP, and 10% intensive chemotherapy
407 (mainly R-ACVBP regimen), the latter being mainly younger patients, of whom one third had
408 the PMBL subtype.^{45,46} Similar observations were made in the US MER cohort with 92.6%
409 DLBCL patients on IC.⁴³ However, population-based registries showed different
410 observations. One Swedish cohort (2007-2014) reported that 14% of DLBCL patients
411 received non-curative intent therapy (non-anthracycline-based regimen).⁴⁷ Similar results
412 were observed in the British Columbia Cancer registry (16% non-curative intent treatment).⁴⁸
413 In a study conducted by the Danish National Lymphoma Registry (LYFO) on 1011 older
414 DLBCL patients (age, ≥75 years; 2003-2012), Juul *et al.* reported palliative-intent treatment
415 for 21% of patients.⁴⁹ These contrasting results highlight the importance of considering
416 inclusion bias, especially regarding older/frailer patients less likely to be treated by IC.
417 Indeed, we observed differences in age distribution between the studied cohort and data

418 from the French registry with a bias toward a younger population in our cohort. First, this
419 difference may be explained by considering enrolment modalities. Participants in the
420 REALYSA cohort were required to provide written informed consent whereas registries are
421 based on an opt-out system (patients are automatically registered). Clinicians may be
422 reluctant to include frail patients with limited life-expectancy, a poor clinical condition, or
423 cognitive disorders. Additionally, since participation in REALYSA also includes patient-
424 reported outcomes from epidemiology questionnaires, older/frail patients may refuse to
425 participate. This active inclusion process has an impact on patient selection with very unfit
426 patients less likely to be included, thus suggesting that comparisons with registries should be
427 made with caution. To address this selection bias, a specific report is periodically sent to
428 each REALYSA center. It outlines the clinical characteristics of patients included locally and
429 compares them with the global cohort data and national registry data (age, sex, and subtype
430 distribution), thus highlighting the potential selection bias in that center. A nested study
431 conducted in one REALYSA center showed that 54% of all lymphoma patients (151 out of
432 278 patients) referred to this department of hematology over a one-year period were included
433 in REALYSA.⁵⁰ Among non-included patients (n=127), the following reasons of non-inclusion
434 were identified: refusal for 39 (31%) patients, start of a treatment in emergency for 20 (16%)
435 patients, physician's evaluation that inclusion in REALYSA was not feasible for 12 (9%)
436 patients, unknown reason for 56 (44%) patients. The inclusion rate was therefore higher than
437 in CT, but it confirms the need for a clearer understanding of recruitment bias in prospective
438 real-world cohorts. Second, the bias toward a younger population may also be explained by
439 the recruiting centers. With 35 centers, REALYSA cannot ensure coverage of all DLBCL
440 patients in France. There is likely an over-representation of large teaching hospitals in this
441 report (N=543, 77%), though there is a paucity of relevant French data in recent literature.⁵¹
442 Thus, REALYSA most likely reflects lymphoma practice in university hospitals in France
443 rather than overall lymphoma practice in France. Further work based on local initiatives,⁵⁰
444 detailed comparison with registries and potentially with the French National Healthcare Data
445 System may help contextualize these results with recent and more exhaustive data.

446

447 Regarding treatment, results are consistent with current guidelines and reflects the
448 implementation of CT results into routine care. Strategies for reducing treatment intensity
449 have been introduced in international guidelines and implemented into routine care,⁵²⁻⁵⁹ with
450 six cycles of R-CHOP now the standard of care (58% of R-CHOP treated patients), and four
451 cycles recently implemented in the lowest aalPI group (38% of R-CHOP treated patients with
452 aalPI=0). In line with RCT results,⁵² only 19 patients (3%) received consolidation treatment
453 with radiotherapy. Although high-dose chemotherapy followed by ASCT used to be a
454 standard of care for high-risk DLBCL patients in the LYSA group,^{31,60} only nine patients
455 (1.4%) underwent this procedure in this DLBCL cohort. Interestingly, in line with current
456 guidelines, most patients underwent a PET/CT-scan to assess disease-stage at diagnosis
457 (94%) and at end of treatment (92%).^{2,53,61} Collecting accurate information on assessment
458 modalities can be particularly challenging in observational cohorts.^{43,62} The REALYSA cohort
459 data include information on the imaging techniques used. Imaging data can thus be uploaded
460 for specific research projects, and notably interim PET-scans.⁶³ Metabolic imaging and pre-
461 treatment circulating tumor DNA levels⁶⁴⁻⁶⁶ have recently been shown to provide compelling
462 information on patient outcomes. Serial biobanking in the REALYSA cohort ensures the
463 feasibility of such studies.²²

464 Regarding outcomes, since the REALYSA program was initiated in 2018, follow-up is limited
465 so far to ensure robustness on patient outcomes. Nevertheless, the prognostic value of the
466 IPI and aalPI scores could be reproduced on this DLBCL population with a plateau on
467 survival curves observed 24 months after diagnosis as previously reported for DLBCL
468 patients.¹⁴ Estimated 1-year OS in our analysis population was 90.0% [86.5-92.5]. On the
469 national scale based on registry data for patients diagnosed between 2010 and 2015,
470 estimated 1-year OS was 71% [70-72].⁶⁷ Comparison between REALYSA outcomes and
471 national registry-based outcomes must be considered with caution. First, as mentioned
472 above, REALYSA do not cover all hematological centers in France and REALYSA centers

473 are mainly university hospitals. Secondly, this study showed that the inclusion rate of patients
474 with palliative care intent was low. Finally, the follow-up of our study is very short for OS
475 estimations, most likely leading to an overestimation of the overall survival. Of course, the
476 global improvement of OS between these two periods (2018-2021 vs. 2010-2015) could also
477 reflect the improvement of treatments mainly at relapse, since standard of care in first-line
478 therapy did not change between these two periods.

479

480 Data from real-world settings are key to contextualizing the results of CT, in particular
481 regarding the transportability of CT results to the general population. Herein, we conducted
482 an exploratory evaluation of the transportability of 2 clinical trials using data from REALYSA.
483 First, we can see that when applying inclusion criteria of both CTs, POLARIX and
484 SENIOR^{23,24} on REALYSA data, a significant number of patients can be identified from our
485 real-life population, though lower for patients over 80 years of age. The ORR and 1-year PFS
486 seemed comparable between the real-world POLARIX-like population and the POLARIX
487 control arm (88.8% vs 83.8% and 79.8% vs 79.8% respectively), with the limitation of
488 different assessment processes (*i.e.* centralized review in the CT versus local review in the
489 real-world cohort). The 1-year EFS and OS between SENIOR-like population and the
490 SENIOR control arm also seems comparable (64.5% vs 60% and 78.3% vs 78.5%
491 respectively), with the limitation of different imaging techniques used to assess treatment
492 response (*i.e.* PET-scan in S-L population vs CT scan in SENIOR²⁴ control arm). In both
493 cases, there is also the limitation of different follow-up modalities in real-life versus
494 interventional trials (with more standardized imaging assessments in the latter). These data
495 suggest that it is feasible with data from REALYSA to assess transportability of CTs.
496 However, these preliminary results are only descriptive, as a “proof-of-concept” analysis and
497 must be considered with caution considering the short median follow-up. Further work
498 including adjustment on variables known to affect outcomes, using more complex statistical
499 techniques (*e.g.* propensity score) and using data with longer follow-up should be considered
500 to confirm these preliminary results.

501

502 This proof-of-concept study on REALYSA 1L DLBCL patients suggests that this real-world
503 cohort can generate high quality data. The resulting database has minimal missing data on
504 key variables and results are consistent with existing literature in terms of baseline
505 characteristics, treatments, and outcomes. REALYSA is a source of meaningful RWD with
506 significant potential for a multitude of different applications, including a better characterization
507 of the lymphoma population in France, as well as innovative study designs, including—but
508 not restricted to— new outcome endpoints or the creation of synthetic control arms.

509

510 **Acknowledgements**

511

512 The authors would like to thank all the patients and their families for their confidence; the
513 Hospices Civils de Lyon (France), who sponsored the study, as well as the LYSA and
514 LYSARC, who coordinate REALYSA, in collaboration with EPICENE (INSERM). The authors
515 also thank all the investigators and the investigating centers. The REALYSA study is funded
516 by several commercial organizations (Roche, Takeda, Janssen, Amgen, Celgene-BMS). The
517 present study was specifically supported by F Hoffmann-La Roche.

518

519

520 **Authorship and Conflict-of-Interest Statements**

521

522 **Conception and design:** H. Ghesquieres, F. Cherblanc, A. Belot, S. Micon, C. Esnault, K.
523 Thokagevistk, A. Bernier, F. Boissard, V. Camus

524 **Provision of study materials or patients:** H. Ghesquieres, K.-K. Bouabdallah, L.-M.
525 Fornecker, F. Bijou, C. Haioun, N. Morineau, L. Ysebaert, G. Damaj, B. Tessoulin, S. Guidez,
526 F. Morschhauser, C. Thiéblemont, A. Chauchet, R. Gressin, F. Jardin, C. Fruchart, G.
527 Labouré, L. Fouillet, P. Lionne-Huyghe, A. Bonnet, L. Lebras, S. Amorim, C. Leyronnas, G.
528 Olivier, R. Guieze, R. Houot, V. Launay, B. Drenou, O. Fitoussi, L. Detournignies, J.
529 Abraham, C. Soussain, F. Lachenal, G. Pica, S. Le Guyader-Peyrou, A. Monnereau, C.
530 Rossi, V. Camus

531 **Collection and assembly of data:** F. Cherblanc, A. Belot, M. Bonjour, P. Fogarty, P. Cony-
532 Makhoul, A. Bernier

533 **Statistical analysis:** A. Belot, P. Fogarty

534 **Manuscript writing:** H. Ghesquieres, F. Cherblanc, A. Belot, A. Bernier

535 **Final approval of manuscript:** all authors approved the final version of the manuscript and
536 are accountable for all aspects of the work

537

538 H. Ghesquieres: consultancy and advisory boards: Roche, Gilead, BMS, Abbvie

539 S. Micon, C. Esnault, K. Thokagevistk, F. Boissard: Roche Employee

540 F. Morschhauser: consultancy and advisory boards: Roche

541 V. Camus: consultancy and honoraria: Roche, Janssen Celgene-BMS; Research and travel
542 grants: Celgene-BMS.

543 The authors declare that they have no competing interests.

544

545 **ORCID**

546 ORCID : Hervé Ghesquières : <https://orcid.org/0000-0002-3131-3718>

547

548 **References**

549

- 550 1. Le Guyader-Peyrou S, Defossez G, Dantony E, et al. Estimations nationales de l'incidence et de la
551 mortalité par cancer en France métropolitaine entre 1990 et 2018. Volume 2 – Hémopathies
552 malignes. Étude à partir des registres des cancers du réseau Francim. 2019;
- 553 2. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. *N. Engl. J. Med.* 2021;384(9):842–858.
- 554 3. Khurana A, Mwangi R, Nowakowski GS, et al. Impact of Organ Function–Based Clinical Trial
555 Eligibility Criteria in Patients With Diffuse Large B-Cell Lymphoma: Who Gets Left Behind? *J. Clin.*
556 *Oncol.* 2021;39(15):1641–1649.
- 557 4. Loh Z, Salvaris R, Chong G, et al. Evolution of eligibility criteria for diffuse large B-cell lymphoma
558 randomised controlled trials over 30 years. *Br. J. Haematol.* 2021;193(4):741–749.
- 559 5. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-
560 based disparities. *JAMA.* 2004;291(22):2720–2726.
- 561 6. Jenei K, Haslam A, Olivier T, Miljković M, Prasad V. What drives cancer clinical trial accrual? An
562 empirical analysis of studies leading to FDA authorisation (2015–2020). *BMJ Open.*
563 2022;12(10):e064458.
- 564 7. Unger JM, Vaidya R, Hershman DL, Minasian LM, Fleury ME. Systematic Review and Meta-
565 Analysis of the Magnitude of Structural, Clinical, and Physician and Patient Barriers to Cancer
566 Clinical Trial Participation. *JNCI J. Natl. Cancer Inst.* 2019;111(3):245–255.
- 567 8. Harkins RA, Patel SP, Lee MJ, et al. Improving eligibility criteria for first-line trials for patients
568 with DLBCL using a US-based Delphi-method survey. *Blood Adv.* 2022;6(9):2745–2756.
- 569 9. Maurer MJ, Ghesquières H, Link BK, et al. Diagnosis-to-Treatment Interval Is an Important
570 Clinical Factor in Newly Diagnosed Diffuse Large B-Cell Lymphoma and Has Implication for Bias in
571 Clinical Trials. *J. Clin. Oncol.* 2018;36(16):1603–1610.
- 572 10. Concato J, Corrigan-Curay J. Real-World Evidence — Where Are We Now? *N. Engl. J. Med.*
573 2022;386(18):1680–1682.
- 574 11. Concato J, Stein P, Dal Pan GJ, Ball R, Corrigan-Curay J. Randomized, observational,
575 interventional, and real-world—What's in a name? *Pharmacoepidemiol. Drug Saf.*
576 2020;29(11):1514–1517.
- 577 12. El-Galaly TC, Cheah CY, Villa D. Real world data as a key element in precision medicine for
578 lymphoid malignancies: potentials and pitfalls. *Br. J. Haematol.* 2019;186(3):409–419.
- 579 13. Nowakowski G, Maurer MJ, Cerhan JR, Dey D, Sehn LH. Utilization of real-world data in assessing
580 treatment effectiveness for diffuse large B-cell lymphoma. *Am. J. Hematol.* 2023;98(1):180–192.
- 581 14. Maurer MJ, Ghesquières H, Jais J-P, et al. Event-Free Survival at 24 Months Is a Robust End Point
582 for Disease-Related Outcome in Diffuse Large B-Cell Lymphoma Treated With
583 Immunochemotherapy. *J. Clin. Oncol.* 2014;32(10):1066–1073.
- 584 15. Thompson CA, Ghesquieres H, Maurer MJ, et al. Utility of Routine Post-Therapy Surveillance
585 Imaging in Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.* 2014;32(31):3506–3512.
- 586 16. Alabdajabar MS, Durani U, Thompson CA, Constine LS, Hashmi SK. The forgotten survivor: A
587 comprehensive review on Non-Hodgkin lymphoma survivorship. *Am. J. Hematol.*
588 2022;97(12):1627–1637.
- 589 17. Solal-Celigny P. Follicular Lymphoma International Prognostic Index. *Blood.* 2004;104(5):1258–
590 1265.
- 591 18. Federico M, Bellei M, Marcheselli L, et al. Follicular Lymphoma International Prognostic Index 2:
592 A New Prognostic Index for Follicular Lymphoma Developed by the International Follicular
593 Lymphoma Prognostic Factor Project. *J. Clin. Oncol.* 2009;27(27):4555–4562.
- 594 19. Bachy E, Maurer MJ, Habermann TM, et al. A simplified scoring system in de novo follicular
595 lymphoma treated initially with immunochemotherapy. *Blood.* 2018;132(1):49–58.
- 596 20. Casulo C, Byrtek M, Dawson KL, et al. Early Relapse of Follicular Lymphoma After Rituximab Plus
597 Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone Defines Patients at High Risk for
598 Death: An Analysis From the National LymphoCare Study. *J. Clin. Oncol.* 2015;33(23):2516–2522.

- 599 21. Casulo C, Dixon JG, Le-Rademacher J, et al. Validation of POD24 as a robust early clinical end
600 point of poor survival in FL from 5225 patients on 13 clinical trials. *Blood*. 2022;139(11):1684–
601 1693.
- 602 22. Ghesquières H, Rossi C, Cherblanc F, et al. A French multicentric prospective prognostic cohort
603 with epidemiological, clinical, biological and treatment information to improve knowledge on
604 lymphoma patients: study protocol of the “REal world dAta in LYmphoma and survival in adults”
605 (REALYSA) cohort. *BMC Public Health*. 2021;21(1):432.
- 606 23. Tilly H, Morschhauser F, Sehn LH, et al. Polatuzumab Vedotin in Previously Untreated Diffuse
607 Large B-Cell Lymphoma. *N. Engl. J. Med*. 2022;386(4):351–363.
- 608 24. Oberic L, Peyrade F, Puyade M, et al. Subcutaneous Rituximab-MiniCHOP Compared With
609 Subcutaneous Rituximab-MiniCHOP Plus Lenalidomide in Diffuse Large B-Cell Lymphoma for
610 Patients Age 80 Years or Older. *J. Clin. Oncol*. 2021;39(11):1203–1213.
- 611 25. Degtiar I, Rose S. A Review of Generalizability and Transportability. *Annu. Rev. Stat. Its Appl*.
612 2023;10(1):501–524.
- 613 26. International Non-Hodgkin’s Lymphoma Prognostic Factors Project. A predictive model for
614 aggressive non-Hodgkin’s lymphoma. *N. Engl. J. Med*. 1993;329(14):987–994.
- 615 27. Abramson JS, Solomon SR, Arnason JE, et al. Lisocabtagene Maraleucel (liso-cel) Versus Standard
616 of Care (SOC) with Salvage Chemotherapy Followed By Autologous Stem Cell Transplantation
617 (ASCT) As Second-Line (2L) Treatment in Patients with Relapsed or Refractory Large B-Cell
618 Lymphoma (LBCL): Primary Analysis of the Randomized, Phase 3 Transform Study. *Blood*.
619 2022;140(Supplement 1):1581–1583.
- 620 28. Kamdar M, Solomon SR, Arnason J, et al. Lisocabtagene maraleucel versus standard of care with
621 salvage chemotherapy followed by autologous stem cell transplantation as second-line
622 treatment in patients with relapsed or refractory large B-cell lymphoma (TRANSFORM): results
623 from an interim analysis of an open-label, randomised, phase 3 trial. *The Lancet*.
624 2022;399(10343):2294–2308.
- 625 29. Locke FL, Miklos DB, Jacobson CA, et al. Axicabtagene Ciloleucel as Second-Line Therapy for
626 Large B-Cell Lymphoma. *N. Engl. J. Med*. 2022;386(7):640–654.
- 627 30. Thieblemont C, Tilly H, Gomes da Silva M, et al. Lenalidomide Maintenance Compared With
628 Placebo in Responding Elderly Patients With Diffuse Large B-Cell Lymphoma Treated With First-
629 Line Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone. *J. Clin. Oncol*.
630 2017;35(22):2473–2481.
- 631 31. Le Gouill S, Ghesquières H, Oberic L, et al. Obinutuzumab vs rituximab for advanced DLBCL: a
632 PET-guided and randomized phase 3 study by LYSA. *Blood*. 2021;137(17):2307–2320.
- 633 32. US Food and Drug Administration. Framework for FDA’s Real-World Evidence Program. 2018.
634 2018;
- 635 33. Heads of Medicines Agency, European Medicines Agency. HMA-EMA Joint Big Data Taskforce –
636 summary report. 2019;
- 637 34. European Network of HTA. Vision paper on the sustainable availability of the proposed Registry
638 Evaluation and Quality Standards Tool (REQueST). 2019;
- 639 35. Public Policy Committee, International Society of Pharmacoepidemiology. Guidelines for good
640 pharmacoepidemiology practice (GPP): Guidelines for good pharmacoepidemiology practice.
641 *Pharmacoepidemiol. Drug Saf*. 2016;25(1):2–10.
- 642 36. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies
643 in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet*.
644 2007;370(9596):1453–1457.
- 645 37. Arboe B, El-Galaly TC, Clausen MR, et al. The Danish National Lymphoma Registry: Coverage and
646 Data Quality. *PLOS ONE*. 2016;11(6):e0157999.
- 647 38. Lindner L, Weiß A, Reich A, et al. Implementing an automated monitoring process in a digital,
648 longitudinal observational cohort study. *Arthritis Res. Ther*. 2021;23(1):181.
- 649 39. Blacketer C, Defalco FJ, Ryan PB, Rijnbeek PR. Increasing trust in real-world evidence through
650 evaluation of observational data quality. *J. Am. Med. Inform. Assoc*. 2021;28(10):2251–2257.

- 651 40. Laurent C, Baron M, Amara N, et al. Impact of Expert Pathologic Review of Lymphoma Diagnosis:
652 Study of Patients From the French Lymphopath Network. *J. Clin. Oncol.* 2017;35(18):2008–2017.
- 653 41. Smith A, Howell D, Crouch S, et al. Cohort Profile: The Haematological Malignancy Research
654 Network (HMRN): a UK population-based patient cohort. *Int. J. Epidemiol.* 2018;47(3):700–700g.
- 655 42. Roman E, Kane E, Howell D, et al. Cohort Profile Update: The Haematological Malignancy
656 Research Network (HMRN) UK population-based cohorts. *Int. J. Epidemiol.* 2022;51(3):e87–e94.
- 657 43. Cerhan JR, Link BK, Habermann TM, et al. Cohort Profile: The Lymphoma Specialized Program of
658 Research Excellence (SPORE) Molecular Epidemiology Resource (MER) Cohort Study. *Int. J.*
659 *Epidemiol.* 2017;46(6):1753–1754i.
- 660 44. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-
661 world diffuse large B-cell lymphoma population. *Blood.* 2023;141(20):2493–2507.
- 662 45. Camus V, Rossi C, Sesques P, et al. Outcomes after first-line immunochemotherapy for primary
663 mediastinal B-cell lymphoma: a LYSA study. *Blood Adv.* 2021;5(19):3862–3872.
- 664 46. Sibon D, Gisselbrecht C, Molina TJ, et al. Outcome of Patients with Primary Mediastinal Large B-
665 Cell Lymphoma after R-CHOP21, R-CHOP14 and R-ACVBP: A Pooled Analysis of Clinical Trials from
666 Lysa. *Blood.* 2022;140(Supplement 1):1082–1084.
- 667 47. Harrysson S, Eloranta S, Ekberg S, et al. Incidence of relapsed/refractory diffuse large B-cell
668 lymphoma (DLBCL) including CNS relapse in a population-based cohort of 4243 patients in
669 Sweden. *Blood Cancer J.* 2021;11(1):9.
- 670 48. Lee B, Goktepe O, Hay K, et al. Effect of Place of Residence and Treatment on Survival Outcomes
671 in Patients With Diffuse Large B-Cell Lymphoma in British Columbia. *The Oncologist.*
672 2014;19(3):283–290.
- 673 49. Juul MB, Jensen PH, Engberg H, et al. Treatment strategies and outcomes in diffuse large B-cell
674 lymphoma among 1011 patients aged 75 years or older: A Danish population-based cohort
675 study. *Eur. J. Cancer.* 2018;99:86–96.
- 676 50. Le Lan C, Belot A, Golfier C, et al. Evaluation of Participation and Recruitment Bias in a
677 Prospective Real-Life Multicentric Cohort « Real World Data in Lymphoma and Survival in
678 Adults » (REALYSA study) for Newly Diagnosed Lymphoma Patients over One Year in a
679 Hematology Department of Teaching Hospital. *Blood.* 2022;140(Supplement 1):5210–5211.
- 680 51. Le Guyader-Peyrou S, Orazio S, Dejardin O, et al. Factors related to the relative survival of
681 patients with diffuse large B-cell lymphoma in a population-based study in France: does socio-
682 economic status have a role? *Haematologica.* 2017;102(3):584–592.
- 683 52. Lamy T, Damaj G, Soubeyran P, et al. R-CHOP 14 with or without radiotherapy in nonbulky
684 limited-stage diffuse large B-cell lymphoma. *Blood.* 2018;131(2):174–181.
- 685 53. Zelenetz AD, Gordon LI, Chang JE, et al. NCCN Guidelines® Insights: B-Cell Lymphomas, Version
686 5.2021: Featured Updates to the NCCN Guidelines. *J. Natl. Compr. Canc. Netw.*
687 2021;19(11):1218–1230.
- 688 54. Poeschel V, Held G, Ziepert M, et al. Excellent Outcome of Young Patients (18-60 years) with
689 Favourable-Prognosis Diffuse Large B-Cell Lymphoma (DLBCL) Treated with 4 Cycles CHOP Plus 6
690 Applications of Rituximab: Results of the 592 Patients of the Flyer Trial of the Dshnhl/GLA. *Blood.*
691 2018;132(Supplement 1):781–781.
- 692 55. Pfreundschuh M, Schubert J, Ziepert M, et al. Six versus eight cycles of bi-weekly CHOP-14 with
693 or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised
694 controlled trial (RICOVER-60). *Lancet Oncol.* 2008;9(2):105–116.
- 695 56. Sehn LH, Congiu AG, Culligan DJ, et al. No Added Benefit of Eight Versus Six Cycles of CHOP When
696 Combined with Rituximab in Previously Untreated Diffuse Large B-Cell Lymphoma Patients:
697 Results from the International Phase III GOYA Study. *Blood.* 2018;132(Supplement 1):783–783.
- 698 57. Poeschel V, Held G, Ziepert M, et al. Four versus six cycles of CHOP chemotherapy in
699 combination with six applications of rituximab in patients with aggressive B-cell lymphoma with
700 favourable prognosis (FLYER): a randomised, phase 3, non-inferiority trial. *The Lancet.*
701 2019;394(10216):2271–2281.

- 702 58. Sehn LH, Scott DW, Villa D, et al. Long-Term Follow-up of a PET-Guided Approach to Treatment
703 of Limited-Stage Diffuse Large B-Cell Lymphoma (DLBCL) in British Columbia (BC). *Blood*.
704 2019;134(Supplement_1):401–401.
- 705 59. Bologna S, Vander Borgh T, Briere J, et al. Early Positron Emission Tomography Response-
706 Adapted Treatment in Localized Diffuse Large B-Cell Lymphoma (AAIPI=0) : Results Of The Phase
707 3 LYSA LNH 09-1B Trial. *Hematol. Oncol.* 2021;39:.
- 708 60. Haioun C, Mounier N, Emile JF, et al. Rituximab versus observation after high-dose consolidative
709 first-line chemotherapy with autologous stem-cell transplantation in patients with poor-risk
710 diffuse large B-cell lymphoma. *Ann. Oncol.* 2009;20(12):1985–1992.
- 711 61. Tilly H, Gomes da Silva M, Vitolo U, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical
712 Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 2015;26:v116–v125.
- 713 62. Jakobsen LH, Bøgsted M, Brown P de N, et al. Minimal Loss of Lifetime for Patients With Diffuse
714 Large B-Cell Lymphoma in Remission and Event Free 24 Months After Treatment: A Danish
715 Population-Based Study. *J. Clin. Oncol.* 2017;35(7):778–784.
- 716 63. Kanoun S, Tal I, Berriolo-Riedinger A, et al. Influence of Software Tool and Methodological
717 Aspects of Total Metabolic Tumor Volume Calculation on Baseline [18F]FDG PET to Predict
718 Survival in Hodgkin Lymphoma. *PLOS ONE.* 2015;10(10):e0140830.
- 719 64. Kurtz DM, Scherer F, Jin MC, et al. Circulating Tumor DNA Measurements As Early Outcome
720 Predictors in Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.* 2018;36(28):2845–2853.
- 721 65. Alig S, Macaulay CW, Kurtz DM, et al. Short Diagnosis-to-Treatment Interval Is Associated With
722 Higher Circulating Tumor DNA Levels in Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.*
723 2021;39(23):2605–2616.
- 724 66. Meriranta L, Alkods A, Pasanen A, et al. Molecular features encoded in the ctDNA reveal
725 heterogeneity and predict outcome in high-risk aggressive B-cell lymphoma. *Blood*.
726 2022;139(12):1863–1877.
- 727 67. Monnereau A, Cornet E, Maynadié M, et al. Survie des personnes atteintes de cancer en France
728 métropolitaine 1989-2018 – Lymphome diffus à grandes cellules B. Boulogne-Billancourt :
729 Institut national du cancer, 12 p. 2021;
- 730
731
732

Table 1. Number and repartition of queries

	Baseline		Treatment		Follow-up		Whole CRF	
	N	% (Levels)	N	% (Levels)	N	% (Levels)	N	% (Levels)
Level 1 queries	2426	57%	893	39%	480	44%	3799	50%
Level 2 and 3 queries	1839	43%	1385	61%	619	56%	3843	50%
Total	4265	100%	2278	100%	1099	100%	7642	100%
Repartition per section	56%		30%		14%			

Table 2. Demographic and clinical characteristics at baseline for IC-treated population, analysis population (total and according to 1L treatment received), POLARIX-like population, POLARIX R-CHOP treatment arm²³, SENIOR-like population and SENIOR trial²⁴

	IC treated Population N=697	Analysis population N=645					REALYSA "POLARIX-like" population N=320	POLARIX R-CHOP arm ²³ N=439	REALYSA "SENIOR- like" population N=59	SENIOR R-miniCHOP ²³ arm N=127
		Total N=645	R-High dose anthracycline N=62 (9.6%)	R-CHOP N=482 (74.7%)	R-miniCHOP N=86 (13.3%)	R-Non anthracycline based chemotherapy N=15 (2.3%)				
Male sex	371 (53.2%)	344 (53.3%)	33 (53.2%)	260 (53.9%)	40 (46.5%)	11 (73.3%)	174 (54.4%)	234 (53.3%)	32 (54.2%)	56 (44.0%)
Median age at diagnosis (min-max)	67.0 (19-98)	66.3 (19-98)	40.1 (19-67)	65.2 (20-90)	82.5 (71-95)	82.3 (60-98)	67.4 (24-80)	66.0 (19-80)	83.4 (80-95)	83.0 (80-96) ^β
Age at diagnosis										
≤60	238 (34.1%)	235 (36.4%)	61 (98.4%)	173 (35.9%)	0 (0.0%)	1 (6.7%)	78 (24.4%)	131 (29.8%)	0 (0.0%)	0 (0.0%)
]60;80]	373 (53.5%)	325 (50.4%)	1 (1.6%)	301 (62.4%)	18 (20.9%)	5 (33.3%)	242 (75.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
>80	86 (12.3%)	85 (13.2%)	0 (0.0%)	8 (1.7%)	68 (79.1%)	9 (60.0%)	0 (0.0%)	308 (70.2%)	59 (100.0%) ^ε	127 (100.0%) ^ε
ECOG PS[§]										
Data available	680 (97.6%)	628 (97.3%)	61 (98.4%)	470 (97.5%)	84 (97.7%)	13 (86.7%)	320 (100.0%)	439 (100.0%)	59 (100.0%)	127 (100.0%)
2-4	128 (18.8%)	125 (19.9%)	7 (11.5%)	79 (16.8%)	33 (39.3%)	6 (46.2%)	56 (17.5%)	75 (17.1%)	17 (28.8%)	36 (28.0%)
Main histological subtypes										
DLBCL	521 (74.7%)	480 (74.4%)	32 (51.6%)	369 (76.6%)	69 (80.2%)	10 (66.7%)	244 (76.3%)	367 (83.6%)	47 (79.7%)	NA
PMBL	46 (6.6%)	45 (7.0%)	21 (33.9%)	24 (5.0%)	0 (0.0%)	0 (0.0%)	9 (2.8%)	0 (0.0%)	0 (0.0%)	
HGBL	85 (12.2%)	81 (12.6%)	5 (8.1%)	61 (12.7%)	11 (12.8%)	4 (26.7%)	45 (14.1%)	50 (11.4%)	9 (15.2%)	
Other large B-cell lymphoma	45 (6.5%)	39 (6.0%)	4 (6.4%)	28 (5.8%)	6 (7.0%)	1 (6.7%)	22 (6.8%)	22 (5.0%)	3 (5.1%)	
Low grade component at diagnosis	42 (6.0%)	41 (6.4%)	2 (3.2%)	31 (6.4%)	7 (8.1%)	1 (6.7%)	20 (6.3%)	NA	5 (8.5%)	NA
PET/CT performed at diagnosis										
Data available	694 (99.6%)	642 (99.5%)	62 (100.0%)	480 (99.6%)	85 (98.8%)	15 (100.0%)	318 (99.4%)	NA	58 (98.3%)	NA
Yes	658 (94.8%)	606 (94.4%)	62 (100.0%)	464 (96.7%)	70 (82.4%)	10 (66.7%)	308 (96.9%)	NA	47 (81.0%)	NA
Ann-Arbor stage[§]										
Data available	696 (99.9%)	644 (99.8%)	62 (100.0%)	481 (99.8%)	86 (100.0%)	15 (100.0%)	320 (100.0%)	439 (100.0%)	59 (100.0%)	127 (100.0%)
III-IV	518 (74.4%)	472 (73.3%)	41 (66.1%)	350 (72.8%)	70 (81.4%)	11 (73.3%)	291 (90.9%)	387 (88.2%)	52 (88.1%)	105 (83.0%)
IPI class[§]										
Data available	671 (96.3%)	619 (96.0%)	60 (96.8%)	466 (96.7%)	81 (94.2%)	12 (80.0%)	320 (100.0%)	439 (100.0%)	57 (96.6%)	126 (99.2%)
0-1	155 (23.1%)	152 (24.6%)	23 (38.3%)	121 (26.0%)	7 (8.6%)	1 (8.3%)	0 (0.0%)	0 (0.0%)	3 (5.3%)	
2	153 (22.8%)	141 (22.8%)	15 (25.0%)	110 (23.6%)	12 (14.8%)	4 (33.3%)	110 (34.3%)	167 (38.0%)	9 (15.8%)	32 (25.0%)
3	167 (24.9%)	148 (23.9%)	18 (30.0%)	110 (23.6%)	19 (23.5%)	1 (8.3%)	210 (65.7%)	272 (62.0)	16 (28.1%)	94 (75.0%)
4-5	196 (29.2%)	178 (28.8%)	4 (6.7%)	125 (26.8%)	43 (53.1%)	6 (50.0%)			29 (50.9%)	
Age-adjusted IPI										
Data available	662 (95.0%)	610 (94.6%)	59 (95.2%)	458 (95.0%)	81 (94.2%)	12 (80.0%)	315 (98.4%)	NA	57 (96.6%)	NA
0	77 (11.6%)	76 (12.5%)	2 (3.4%)	66 (14.4%)	8 (9.9%)	0 (0.0%)	0 (0.0%)		4 (7.0%)	
1	213 (32.2%)	197 (32.3%)	22 (37.3%)	153 (33.4%)	17 (21.0%)	5 (41.7%)	102 (32.4%)		14 (24.6%)	
2	277 (41.8%)	244 (40.0%)	30 (50.8%)	181 (39.5%)	30 (37.0%)	3 (25.0%)	173 (54.9%)		25 (43.9%)	
3	95 (14.4%)	93 (15.2%)	5 (8.5%)	58 (12.7%)	26 (32.1%)	4 (33.3%)	40 (12.7%)		14 (24.6%)	
Extra-nodal involvement**	540 (77.5%)	499 (77.4%)	39 (62.9%)	372 (77.2%)	75 (87.2%)	13 (86.7%)	276 (86.3%)	NA	53 (89.8%)	NA
Elevated LDH at baseline										
Data available	680 (97.6%)	628 (97.4%)	60 (96.8%)	471 (97.7%)	83 (96.5%)	14 (93.3%)	315 (98.4%)	438 (99.8%)	57 (96.6%)	NA
Yes	441 (64.9%)	402 (64.0%)	52 (86.7%)	284 (60.3%)	58 (69.9%)	8 (57.1%)	226 (71.7%)	284 (64.8%)	39 (68.4%)	

FIGURE LEGENDS

Figure 1 Automated patient profile

The patient profile is divided into 4 main parts. In the upper left part, the patient and tumor characteristics at inclusion are described. In the upper right part, the involvements (nodal and extra-nodal) at diagnosis are reported. Red circles are automatically located on the man/woman (depending on patient's sex) to represent nodal involvements. Nodal involvements reported in the "other" section of the eCRF appear in the red box in the upper right part. Extra-nodal involvements are detailed in the blue box in the lower right part. In the graph on the left, the longitudinal information per line(s) of treatment is indicated, with a gray horizontal bar per line of treatment (the treatment cycles are symbolized by black vertical lines), with the evaluations of the responses (circle above the line with color code according to the response) and the events (progression (inverted red triangle), adverse events (red cross), death (crossed out circle)). Finally, on the lower right part, follow-up information is reported.

Figure 2. Study flowchart

DLBC, Diffuse Large B-Cell lymphoma; IC, Immuno-Chemotherapy, 1L, first-line treatment.

Figure 3. EFS according to IPI (A) and aalPI (B)

IPI: International Prognostic Index; aalPI: age-adjusted International Prognostic Index; EFS: Event-Free Survival

TABLE LEGEND

Table 1. Number and repartition of queries

Table 2. Demographic and clinical characteristics at baseline for IC-treated population, analysis population (total and according to 1L treatment received), POLARIX-like population, POLARIX R-CHOP treatment arm²³, SENIOR-like population and SENIOR trial²⁴

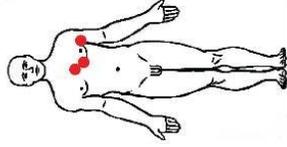
*IC, Immunochemotherapy; DLBCL, diffuse large B-cell lymphoma; PMBL: Primary Mediastinal B-cell Lymphoma; HGBL: High Grade B-cell Lymphoma; **extra-nodal involvement includes spleen; [§]when missing data is reported, the percentage of available data is reported, and the distribution of the variable is expressed as a percentage of available data; IPI, International prognostic score; LDH, lactate dehydrogenase; R-CHOP (rituximab, doxorubicin, cyclophosphamide, vincristine, prednisone and adapted protocol for toxicity reasons); R-mini-CHOP, reduced dose R-CHOP; R-Non-anthracycline based chemotherapy regimen: R-CEOP (cyclophosphamide, etoposide, vincristine, prednisone), R-CEP (cyclophosphamide, etoposide, prednisone), R-COP (cyclophosphamide, vincristine, prednisone), R-COP + Gemcitabine, R-CVP (cyclophosphamide, vincristine, prednisone), R-CVP+Gemcitabine, R-GEMOX (Gemcitabin, Oxaliplatin), R-Ifosfamide, Etoposide; [£]≥80

years; ^βmedian age for both arms (data not available for R-miniCHOP arm only). If there is no “Data available” mention, it means that data are available for 100% of patients.

Figure 1

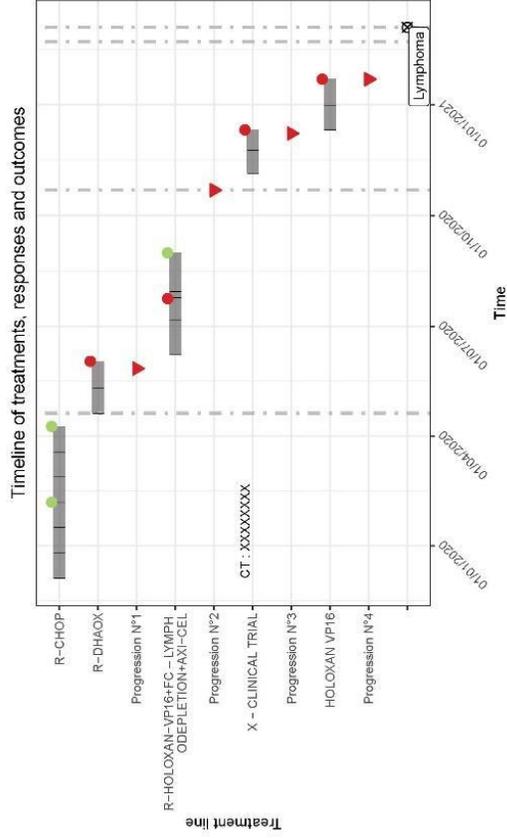
Patient: 1991810XXXXXX (Data export date: XX/XX/XXXX & profile creation date XX/XX/XXXX)
 Nodal involvement at diagnosis

Center	Sex	Date of diagnosis	Lymphoma type	Lymphoma subtype	Pre phase	Perf. status	Stage	Age at diagnosis	Prognostic index
XXXXXXXXXX	Male	XXXXXXXXXX	DLBCL	96003 Diffuse large B-cell lymphoma, Germinal center B-cell type	Yes	0	IV	72	4



Other nodal location
NA

Extra-nodal location
Lung
Bone
Pancreas



Follow-up information

Visit	Theoretical date	Actual Date of visit	Status at visit
Follow-up 6 Months	08/05/2020	20/04/2020	Partial Response
Follow-up 12 Months	07/11/2020	22/10/2020	Progressive Disease
Follow-up 18 Months	09/05/2021	22/02/2021	Progressive Disease
Death	NA	06/03/2021	Death

Figure 2

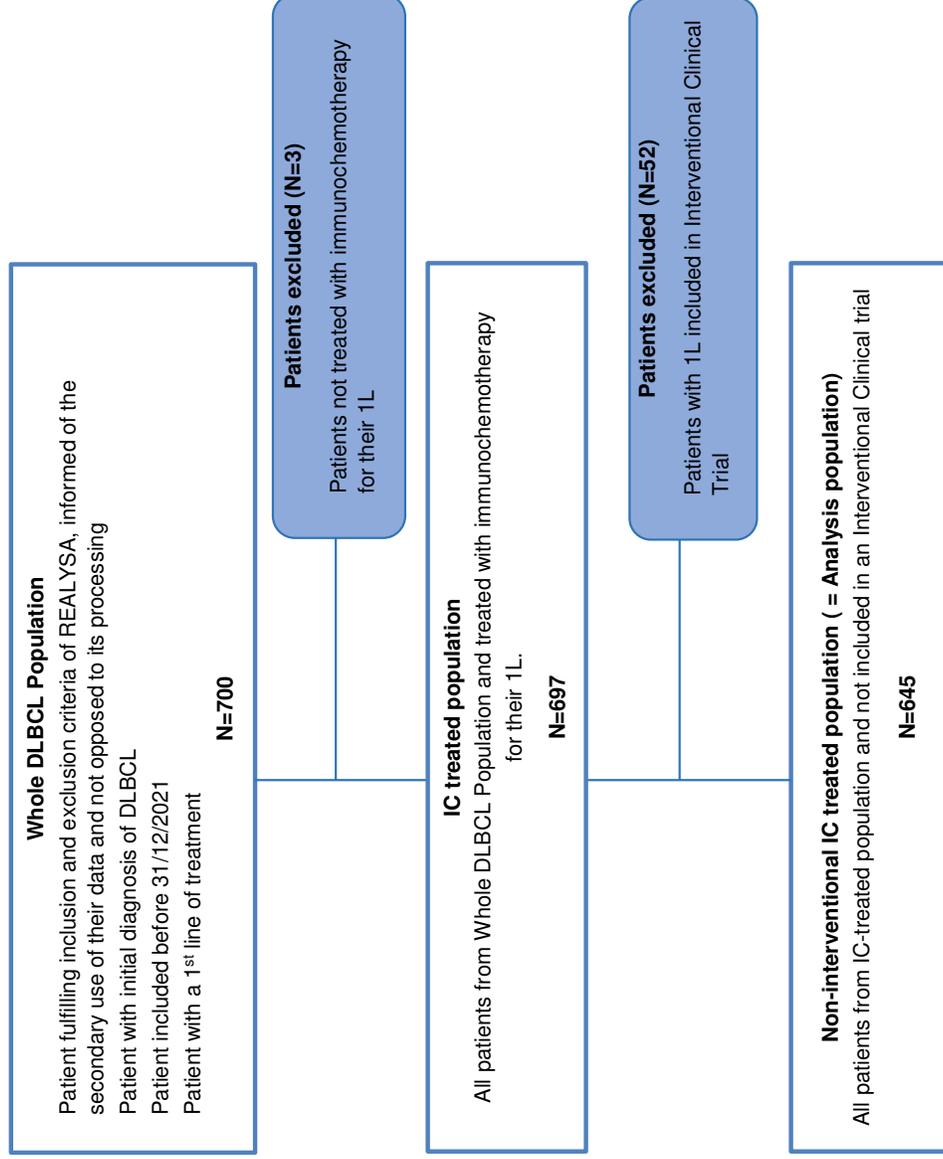


Figure 3

