

Supplemental Table 1. Examples of Level 2 tests

Type of test	Content of test or condition for review	Examples
<p>Inconsistency test : Only patients that appear inconsistent according to the test are reviewed</p>	<p>Inconsistencies between the reported involved sites and the biopsy site or the stage.</p>	<p>- patient without extranodal involvement, nodal involvement restricted to only one side of the diaphragm and a stage different from 2) - patient with a mediastinal biopsy but mediastinal sites not involved</p>
	<p>Patients with aggressive lymphoma not treated 6 months after diagnosis</p>	<p>DLBCL patient not treated at 12 months of diagnosis</p>
	<p>Inconsistency in response chronology</p>	<p>patient in stable disease after being in complete response</p>
	<p>Inconsistency of responses indicated on different pages</p>	<p>patient reported in progression on one page of the eCRF and in partial response in another page, on the same date</p>
	<p>Inconsistency between the reason for death and the last assessment</p>	<p>patient in complete response and reported death cause is lymphoma</p>
	<p>End-of-line assessment missing</p>	<p>Patient without any response assessment 6 months after the last cycle</p>
	<p>Potentially Incomplete treatment line : patient are reviewed if less than 4 cycles are reported for the treatment line to detect partial entry</p>	<p>Patients with only 2 cycles of R-CHOP reported in 1st line</p>
<p>Manual Scientific reviews : <i>all patients responding to the condition are reviewed</i></p>	<p>19 free-text entries to check and encode</p>	<p>Other method of evaluation: free text is manually encoded into a restricted list of item to allow grouping for data interpretation (e.g. echography, endoscopy,...)</p>
	<p>Patient with several lines of treatment to verify the appropriate change of line</p>	<p>Check that R-CHOP cycles are reported as line 1 and that the following R-DHAC cycles are reported as line 2</p>
	<p>Patients with 1 extra-nodal involvement to check the stage</p>	<p>Patient with 1 extra-nodal involvement can be stage II(E) or IV depending on the contiguity of the nodal and extra-nodal sites.</p>

Supplemental Table 2. Variables from the eCRF who will be validated against medical records for quality control

Section	Item to verify
BASELINE	Informed consent signature date (Numeric)
	Genetic studies informed consent signature date (Numeric)
	Last name
	First name
	Sex
	Date of birth (Numeric)
	Date of initial diagnosis (biopsy) (Numeric)
	Initial pathological diagnosis
	DLBCL associated with a low grade lymphoma at diagnosis
	Precise sub-type from which DLBCL is derived
	Ann Arbor stage
Prephase	Prephase ?
	Date of prephase (si prephase=yes)
	Treatment of pre-phase
TREATMENT(s) - Lymphoma X Line	NB: The first 3 lines of treatment are to be reviewed within the QC
	Patient included in an experimental trial study for this treatment line
	Trial name (if LYSARC study)
	First date of treatment phase (numeric)
	First date of RT (numeric)
	Date of transplant (Numeric)
	Date of second transplant (Numeric)
	Start date of maintenance
	End of line evaluation imaging date
	End of line evaluation imaging type
Progression n°X	Date of progression
	Transformation
Last Follow-up completed	Actual date of contact
	Type of last contact
Secondary Malignancy n°X	Date of diagnosis of other malignancy
	Description of other malignancy
End of study	Date status
	Patient status
	Date of death
	Cause of death

Supplemental Table 3. Classes and subclasses of drugs used as first-line treatment

Main class, N (%) Subclass	Non-interventional IC-treated population N=645	
R-CHOP	482	(74.7%)
R-CHOP	318	(49.3%)
R-CHOP + HD-MTX	80	(12.4%)
R-CHOP like	65	(10.1%)
R-CHOP like + HD-MTX	18	(2.8%)
R-CHOP + Ibrutinib	1	(0.2%)
R-miniCHOP	86	(13.3%)
R-miniCHOP	68	(10.5%)
R-miniCHOP like	15	(2.3%)
R-miniCHOP + MTX	3	(0.5%)
R-high-dose anthracycline regimen	62	(9.6%)
R-ACVBP	57	(8.8%)
R-COPADEM like	3	(0.5%)
R-COPADEM	2	(0.3%)
R-non-anthracycline-based chemotherapy	15	(2.3%)
R-CVP	4	(0.6%)
R-CEOP	2	(0.3%)
R-CEP	2	(0.3%)
R-COP	2	(0.3%)
R-GEMOX	2	(0.3%)
R-COP + Gemcitabine	1	(0.2%)
R-CVP + Gemcitabine	1	(0.2%)
R-Ifosfamide-vepeside	1	(0.2%)

Note: percentages for subclasses of treatment are expressed as percentages of the total number of patients.

HD-MTX, high-dose methotrexate

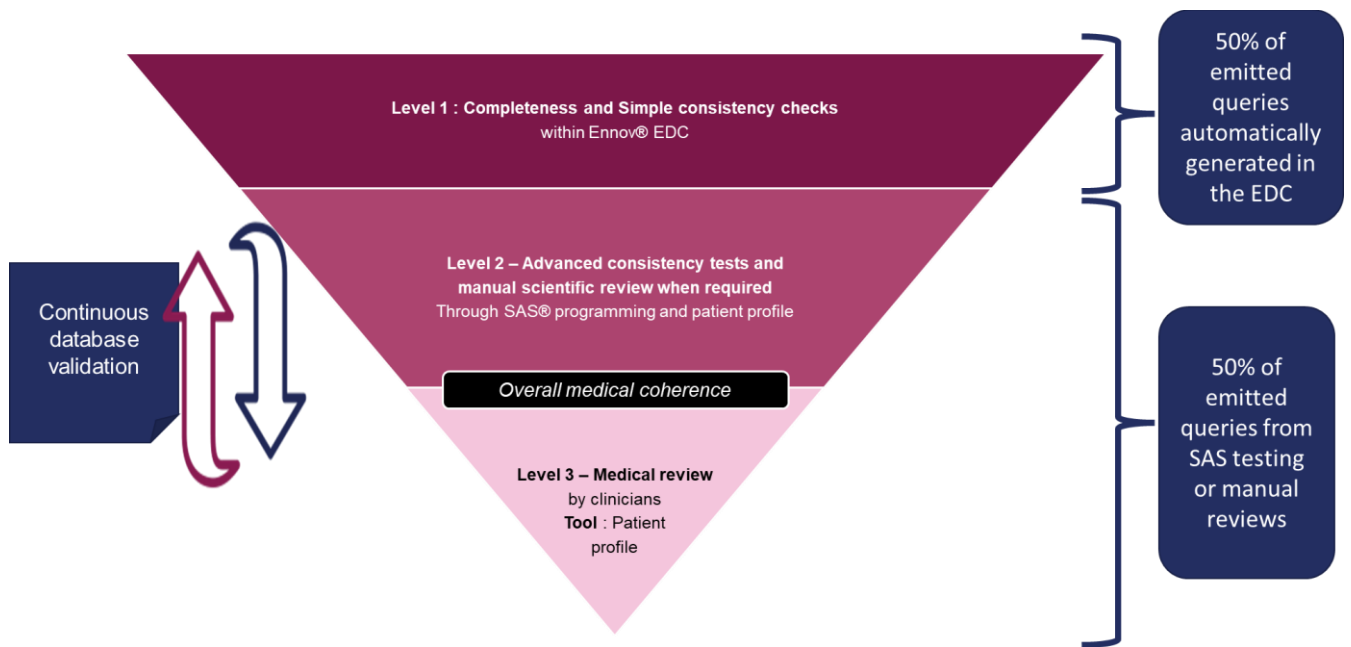
Note on automated data classification: For visualization and analyses purposes, treatments collected from the eCRF at the cycle level were classified. From the detailed eCRF data, an ad hoc SAS program restitutes an intention-to-treat standard protocol. Here, broad classes of treatment therapies were defined, based on the reported treatment and potential complementary information in the pre-treatment multi-disciplinary meeting report, with a treatment “intention” approach. Prophylaxis for lysis syndrom and corticosteroids are not considered for classification and are not reported. The addition of high dose methotrexate is reported as ‘MTX’. Herein, broad classes of therapies have been defined : (i) R-High Dose anthracycline Regimen (R-ACVBP [Rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone], R-COPADEM [Rituximab, methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisone]); (ii) R-CHOP regimen (R-CHOP [Rituximab, doxorubicin, cyclophosphamide, vincristine, prednisone], R-CHOP + MTX, R-CHOP + Ibrutinib); (iii) R-miniCHOP (reduced-dose CHOP) regimen (R-miniCHOP, R-miniCHOP + MTX); (iv) 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. If the protocols are adapted for toxicity reasons, they are reported as “protocol-like”. For example, a patient treated with R-CHOP for whom vincristine has been stopped after a couple of cycles will be reported as “R-CHOP like”.

Supplemental Table 4. Repartition of the number of CHOP cycles

	Non-interventional IC-treated population Treated with R-CHOP All aalPI N = 482		Non-interventional IC-treated population Treated with R-CHOP Restricted to aalPI = 0 N = 63	
R-CHOP cycles	N*	%	N*	%
1	2	0.4%	0	0
2	3	0.6%	0	0
3	6	1.2%	1	1.6%
4	52	11%	24	38.1%
5	30	6%	1	1.6%
6	278	58%	35	55.6%
7	9	2%	0	0
8	102	21%	2	3.2%
Total	482	100%	63	100%

*Only cycles with R-CHOP were considered, Rituximab or Methotrexate alone were not counted as cycles.

Supplemental Figure 1. REALYSA 3-steps validation process



EDC: Electronic Data Capture system; SAS: software (Statistical Analysis System)

Supplemental Figure 2. Distribution of patients by age group in our study (in grey) versus national incidence (in blue) for (A) non-interventional IC-treated population (N=697); and for (B) analysis population (N=645).

As a reading example: men in the 85–90-year-old age group represent 7.7% of the national incident cases, but 2.9% of the analysis population. In contrast, the 50–55-year-old age group represents around 5.7% of the national incident cases, but 9.3% of the analysis population.

