A Methodologic Solution to Missing Deauville Scores Using Imaging Report Data to Classify Lymphoma Treatment Response in Real-World Data

PT29

Scott Swain, Kristin M. Zimmerman Savill, Andrew Klink, Parisa Asgarisabet, Alexandrina Balanean, Harlen Hays, Jill Kaufman, Lindsay McAllister, Courtney Omary, Hsing-Ting Yu, Bindu Kalesan, JaLyna Laney, Bruce Feinberg

Cardinal Health, Dublin, Ohio, USA

BACKGROUND

- Outcome misclassification is an important source of bias when comparing real-world data (RWD) to clinical trial data because classification methods are inherently different.
- To address outcome misclassification, Cardinal Health developed a standardized real-world methodology, rwLugano, using Lugano 2014, to classify lymphoma outcomes using RWD.
- The Deauville score evaluates metabolic activity of lymphoma on positron emission tomography–computed tomography (PET-CT) scans and is a necessary component of the Lugano classification. The score is a standardized method for evaluating the metabolic activity of lymphoma using PET-CT scans and the classification is based on the uptake of radioactive glucose analog, fluorodeoxyglucose (FDG).[1] Based on the amount of FDG absorbed, tumor activity is categorized using a 5-point scale (5PS) (**Table 1**).
- In clinical practice, Deauville scores of 1-3 indicate disease response, whereas scores of 4-5 suggest marginal to no response.[2]
  - Deauville score is often missing in patient charts (imagining report or physician notes), whereas the standardized uptake value (SUV) of metabolically active tissues is often present.
  - SUV is a measure of FDG uptake in tissue expressed as a number indicating brightness in a PET-CT.
  - Accordingly, SUV can be extracted, and used to calculate Deauville score. However, SUV of non-cancer sites (i.e., background tissue, mediastinum, and liver) is often missing in PET-CT reports.
- We developed an algorithm to calculate a real-world Deauville score (rwDeauville) by comparing tumor SUV to literature-based values for background tissue (SUV=1), mediastinum (SUV=3), and liver (SUV=5) (**Table 2**).
- We compared rwLugano derived using reported Deauville scores versus rwDeauville scores to assess agreement.

Table 1. Lugano Classification of Response (Simplified)[3]

Deauville/Lugano 5PS	Change from baseline	New lesions	Bone marrow	Lugano response
1, 2, or 3	Reduced	No	No	CR
4 or 5	Reduced	No	Reduced	PR
4 or 5	No charge	No	No change	NR
4 or 5	Increased	No	Yes	PD
Any	Any	Yes	Any	PD

Abbreviations: CR: complete response; PR: partial response; NR: no response; PD: progressive disease

Table 2. Deauville and rwDeauville Classification

Deauville	Description	SUV algorithm for Deauville*	SUV algorithm for rwDeauville**
1	No uptake or no residual uptake (when used interim)	$SUV_{tumor} = 1$	$SUV_{tumor} = 1$
2	Slight uptake, but ≤ blood pool (mediastinum)	$SUV_{tumor} \leq SUV_{mediastinum}$	$1 < SUV_{tumor} \leq 3$
3	Uptake > mediastinal, but ≤ uptake in the liver	$SUV_{mediastinum} < SUV_{tumor} \leq SUV_{liver}$	$3 < SUV_{tumor} \leq 5$
4	Uptake slightly to moderately higher than liver	$SUV_{liver} < SUV_{tumor} \leq (SUV_{liver})x2$	$5 < SUV_{tumor} \leq 10$
5	Markedly increased uptake or any new lesion (on response evaluation)	$SUV_{tumor} > (SUV_{liver})x2$	$SUV_{tumor} > 10$

OBJECTIVE

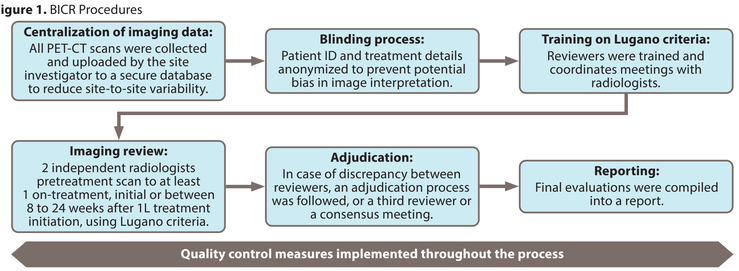
- To develop an algorithm accounting for missing Deauville scores to assess performance of lymphoma treatment response classified via rwLugano-derived response.

METHODS

Study design

- A multicenter, retrospective chart review study conducted at 6 sites within the Cardinal Health Oncology Practice Research Network (PRN), a consortium of US-based community oncologists and hematologists.
- The study included patients ≥18 years old with histologically confirmed, diffuse large B-cell lymphoma (DLBCL) treated with chemoimmunotherapy as first-line (1L) therapy.
- Each participating PRN site selected consecutive patients, starting with the earliest eligible, at each practice.
  - Deidentified data, captured via an electronic case report form (eCRF).
  - Digital PET-CT scans were deidentified upon upload to a secure platform.
- All study materials were reviewed by a central Institutional Review Board.
- Study endpoints:
  - CR (primary endpoint) – complete response
  - PR – partial response
  - SD/NR – stable disease/no response
  - PD – progressive disease
  - ORR – overall response rate
- Standardized blinded independent centralized review (BICR) of PET-CT scans using Lugano 2014 criteria was conducted as follows:

Figure 1. BICR Procedures



METHODS

Eligibility Criteria

Inclusion criteria

- Adults with a diagnosis of DLBCL (with histologic confirmation) between 2015 and 2022.
- Treated with an anthracycline-containing chemoimmunotherapy regimen that includes an anti-CD20 monoclonal antibody.
- PET-CT images available at baseline (within 8 weeks prior to 1L therapy initiation) and initial response assessment scan (within 8-24 weeks of initiating 1L therapy).
- At least 6 months follow-up from 1L therapy initiation, including eligible patients who died within this period.

Exclusion criteria

- Central nervous system (CNS) involvement at the time of DLBCL diagnosis.
- Treated for other malignancies during 1L therapy.
- Enrolled in a clinical trial during 1L therapy.

Treatment Response Assessment

- rwLugano was derived from Lugano 2014 criteria by using the abstracted EMR data associated with imaging reports and scans (**Table 1**).
- rwLugano was calculated using 2 methods:
  1. reported Deauville score
  2. calculated rwDeauville score

Statistical Analysis

- We assessed agreement (%) and concordance (Cohen's kappa [κ]) between Deauville scores recorded in PET-CT reports at baseline and initial DLBCL treatment response and rwDeauville scores calculated using the algorithm in Table 1. We also assessed agreement and concordance between rwLugano calculated at initial response assessment using recorded Deauville scores versus rwDeauville scores.

RESULTS

- A total of 178 patients with DLBCL (105 [59%] male, 73 [41%] female, 137 [77%] White, 13 [7%] Hispanic) diagnosed at a mean age of 66 years (**Table 3**).
- Deauville score and tissue SUV were frequently missing from baseline scans (n=79 [44.4%]) (**Table 4**) and first response assessment (n=30 [16.9%]) (**Table 4**).
- Among baseline PET-CT reports with Deauville data available for comparison (n=99), rwDeauville demonstrated low concordance at 68% agreement, κ=0.38.3 (95% CI 0.22-0.54).
- Among initial response PET-CT reports with Deauville data available for comparison (n=148), rwDeauville demonstrated high concordance at 87.6% agreement, κ=0.82 (95% CI 0.74-0.90).
- The 2 methods used to calculate rwLugano demonstrated high concordance at 93.9% agreement, κ=0.77 (95% CI 0.63-0.91).

Table 3. Baseline Patient Demographic and Clinical Characteristics (N=178)

Participant characteristics	N=178
Age at diagnosis (years), mean (SD)	66.4 (12.8)
Sex, n (%)	
Male	105 (59.0)
Female	73 (41.0)
Race, n (%)	
American Indian or Alaska Native	0 (0)
Asian	5 (2.8)
Black or African American	18 (10.1)
Native Hawaiian or Other Pacific Islander	0 (0)
White	137 (77.0)
Unknown	18 (10.1)
Ethnicity, n (%)	
Hispanic or Latino	13 (7.3)
Not Hispanic or Latino	143 (80.3)
Unknown	22 (12.4)
Duration of follow-up from 1L therapy initiation (months), median (P25-P75)	25.6 (16.8-43.8)
Ann Arbor stage at 1L therapy initiation among patients with known stage, n (%)	154 (86.5)
Stage I	24 (15.6)
Stage II	42 (27.3)
Stage III	38 (24.7)
Stage IV	50 (32.5)

Table 4. Deauville Scores and rwDeauville Scores at Baseline (within 8 weeks prior to 1L therapy) and First Response Assessment (N=178)

Deauville scores	Baseline	1st response
Reported (among patients with Deauville score)		
n, %	99 (55.6)	148 (83.2)
1	9 (9.1)	98 (66.2)
2	9 (9.1)	17 (11.5)
3	9 (9.1)	12 (8.1)
4	10 (10.1)	12 (8.1)
5	62 (62.6)	9 (6.1)
Calculated (among patients for whom Deauville/complete SUV data were not reported)		
n, %	79 (44.4)	30 (16.9)
1	0 (0)	11 (36.7)
2	1 (1.3)	3 (10.0)
3	5 (6.3)	7 (23.3)
4	10 (12.7)	8 (26.7)
5	63 (79.7)	1 (3.3)

RESULTS

Figure 2. Agreement Between Physician-reported Deauville and rwDeauville Scores at Baseline (n=99)

Agreement between physician-reported Deauville and rwDeauville scores among patients with Deauville in baseline scans available (n=99)	rwDeauville				
Physician-reported Deauville (among patients having score)	**1**	**2**	**3**	**4**	**5**
1	2	3	1	0	3
2	0	0	1	2	6
3	0	0	1	1	7
4	0	0	1	2	7
5	0	0	0	1	61

	Kappa/Coefficient/ Percentage	95% CI-Lower limit	95% CI-Upper limit
Overall percent agreement (total concordance/total sample)	**68.04%**		
Weighted kappa for 5-level Deauville (κ, 95% CI), *P*-value	**0.38**	**0.22**	**0.54**

Figure 3. Agreement Between Physician-reported Deauville and rwDeauville Scores at First Response Assessment (n=148)

Agreement between physician-reported Deauville and rwDeauville scores among patients with Deauville in initial scans available (n=148)	rwDeauville				
Physician-reported Deauville (among patients having score)	**1**	**2**	**3**	**4**	**5**
1	92	1	3	1	1
2	1	13	2	1	0
3	1	1	7	2	1
4	0	0	3	7	2
5	0	0	0	1	8

	Kappa/Coefficient/ Percentage	95% CI-Lower limit	95% CI-Upper limit
Overall percent agreement (total concordance/total sample)	**87.59%**		
Weighted kappa for 5-level Deauville (κ, 95% CI), *P*-value	**0.82**	**0.74**	**0.90**

Figure 4. Agreement Between Physician-reported Deauville and rwDeauville Scores at First Response Assessment (n=148)

Agreement between rwLugano derived using reported Deauville scores and rwDeauville at first response assessment (n=148)	rwLugano derived using rwDeauville			
rwLugano derived using reported Deauville	**CR**	**PR**	**SD/NR**	**PD**
CR	121	4	0	2
PR	3	14	0	0
SD/NR	0	0	0	0
PD	0	0	0	4

	Kappa/Coefficient/ Percentage	95% CI-Lower limit	95% CI-Upper limit
Overall percent agreement (total concordance/total sample)	**93.92%**		
Weighted kappa for 5-level Deauville (κ, 95% CI), *P*-value	**0.77**	**0.63**	**0.91**

CONCLUSIONS

- As Lugano cannot be calculated without Deauville, we developed an algorithm to calculate rwDeauville that in turn would permit calculation of rwLugano when Deauville score is missing from the medical record.
- The rwDeauville construct did not perform well at baseline when SUV measures tend to be high and variable.
- The rwDeauville algorithm estimated missing Deauville score and background SUV data, with high accuracy at first response assessment.
- Calculation of rwDeauville facilitated completeness of the novel rwLugano classification, a standardized real-world methodology based on Lugano 2014 criteria, to classify lymphoma outcomes using RWD for all study patients.

REFERENCES

1. Kim HJ, Lee R, Choi H, Paeng JC, Cheon GJ, Lee DS, et al. Application of quantitative indexes of FDG PET to treatment response evaluation in indolent lymphoma. *Nucl Med Mol Imaging*. 2018;52:342–349. doi: 10.1007/s13139-018-0543-8.
2. Li YH, Zhao YM, Jiang YL, Tang S, Chen MT, Xiao ZZ, et al. The prognostic value of end-of-treatment FDG-PET/CT in diffuse large B cell lymphoma: comparison of visual Deauville criteria and a lesion-to-liver SUV(max) ratio-based evaluation system. *Eur J Nucl Med Mol Imaging*. 2022;49:1311–1321. doi: 10.1007/s00259-021-05581-z.
3. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32(27):3059-3068.

ACKNOWLEDGEMENTS

This project was supported by the Oncology Center of Excellence, Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) contract number **75F40121C00134** totaling $749,362 with 100 percent funded by FDA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS, or the U.S. Government.

OTHER RELATED POSTERS PRESENTED AT ISPOR

- (**CO201**) Feasibility of Using Positron Emission Tomography–Computed Tomography (PET-CT) Scans from Real-World Medical Record Data to Support Lymphoma Treatment Response Assessment
- (**MSR101**) A Novel Methodology for Assessing Response to Lymphoma Treatment in Real-World Studies – Real-World Lugano (rwLugano)

Presented at ISPOR 2024 - The Professional Society for Health Economics and Outcomes Research
May 5-8, 2024
Atlanta, Georgia, USA

Corresponding author: Scott Swain, PhD, MPH
scott.swain@cardinalhealth.com

To learn more, scan the QR Code or visit cardinalhealth.com/rwe-abstracts

CardinalHealth™