An automated quantification algorithm for evaluating total metabolic tumor volume in patients with FDG-avid lymphomas using a deep learning model

Tao Xu,[1] Skander Jemaa,[2] Manas Kumar,[2] Sandhya Balasubramanian,[2] Aisha Shamas-Din,[1] Svetlana Lyalina,[2] Souhila Ounadjela,[2] Bilal Malik,[2] Joe Lee,[2] Salvador Figueroa Morales,[1] Tina Nielsen,[1] Richard AD Carano,[2] Lale Kostakoglu,[3] William Capra[2]

[1]F. Hoffmann-La Roche Ltd, Basel, Switzerland; [2]Genentech, Inc, South San Francisco, CA, USA; [3]Department of Radiology and Medical Imaging, University of Virginia, Charlottesville, VA, USA

Poster number: 245

Introduction

- Total metabolic tumor volume (TMTV) is a quantitative radiological assessment of total tumor burden that can be derived from 18F-fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) scans.[1]
- Although the TMTV metric has prognostic value in diffuse large B-cell lymphoma (DLBCL)[2,3] and follicular lymphoma (FL),[4] methodological limitations mean that it is rarely used in clinical practice.[1,5–7]
- Accordingly, we are developing a model based on deep learning[8–10] to automate the detection and segmentation of lesions and the quantification of TMTV in people with FDG-avid lymphoma.

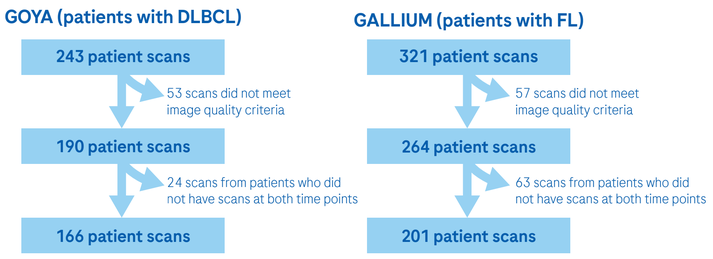
Aim

- To evaluate model performance and identify clinical and technical factors that may influence the accuracy of lesion detection and TMTV quantification in patients with DLBCL or FL.

Methods

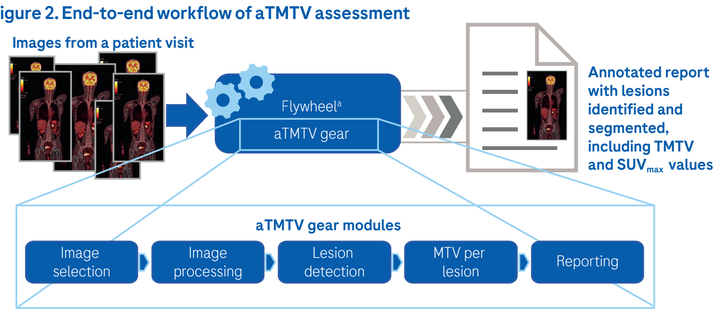
- The model was trained using retrospective trial data, comprising baseline and post-treatment FDG-PET/CT scans from 836 adults with DLBCL, from the phase 3 GOYA study (ClinicalTrials.gov identifier: NCT01287741).
- The test set included baseline and post-treatment scans from 166 adults with DLBCL (an independent hold-out set from GOYA) and from 201 adults with advanced FL collected as part of the GALLIUM study (ClinicalTrials.gov identifier: NCT01332968; **Figure 1**).
  - Although the GALLIUM study enrolled participants with advanced indolent non-Hodgkin lymphoma, 86% of participants had been diagnosed with FL.
- To evaluate model performance, images in the test set were assessed by expert readers using semi-automated software (manual TMTV [mTMTV]) and by the algorithm (automated TMTV [aTMTV]; **Figure 2**).
- Pearson's correlation coefficient (r) was used to assess the relationship between aTMTV and mTMTV, and measurement bias was assessed using the slope and intercept of weighted Deming regression.
- Lesion detection performance was assessed by sensitivity (true positive rate: the proportion of mTMTV-detected lesions identified by aTMTV) and precision (positive predictive value: the proportion of aTMTV-detected lesions identified by mTMTV). Sensitivity and precision values were calculated across the cohort and at the patient level (**Figure 3**).
- Subgroup analyses evaluated algorithm performance for TMTV quantification among patient populations with different demographics and clinical characteristics, and across images from different PET/CT scanner manufacturers.

Figure 1. Data selection for the test set (combined data set, N = 367)



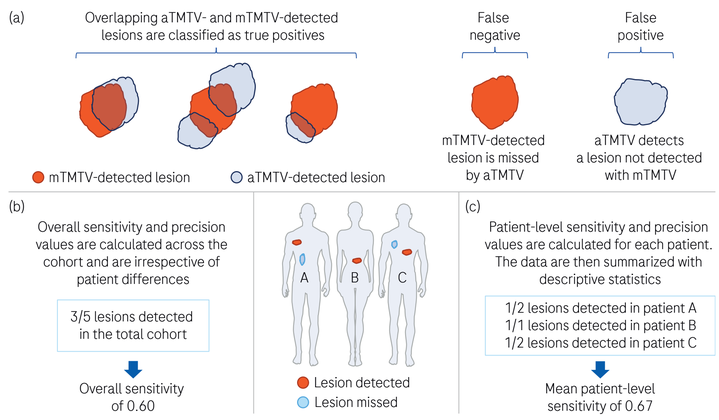
DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma.

Figure 2. End-to-end workflow of aTMTV assessment



aTMTV is an investigational device currently in development.
[a]Flywheel is an R&D data management platform that captures, curates and computes imaging and associated biomedical data to accelerate discovery and product development.
aTMTV, automated TMTV; MTV, metabolic tumor volume; SUV_{max}, maximum standardized uptake value; TMTV, total metabolic tumor volume.

Figure 3. Classification of true positives, false negatives and false positives in (a) lesion detection performance, (b) calculation of overall sensitivity and (c) patient-level sensitivity



aTMTV, automated TMTV; mTMTV, manual TMTV; TMTV, total metabolic tumor volume.

Results

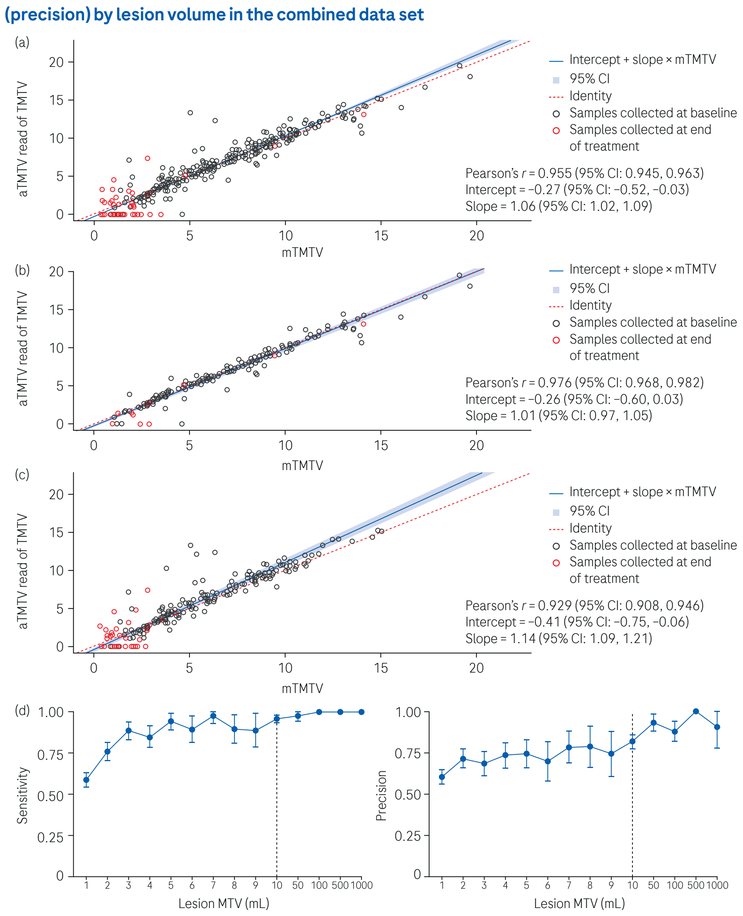
- At baseline, most patients were under 65 years old and the majority had an Eastern Cooperative Oncology Group Performance Status of 0 or 1 (**Table 1**).
- aTMTV quantification highly correlated with mTMTV in the combined data set and in the DLBCL and FL subgroups (Pearson's $r > 0.9$; **Figures 4a–c**).

Table 1. Patient characteristics at baseline

Characteristic, n (%)	Patients with DLBCL (GOYA study) n = 166	Patients with FL (GALLIUM study) n = 201
Sex, female	85 (51.2)	123 (61.2)
Age group, years		
< 65	98 (59.0)	139 (69.2)
≥ 65	68 (41.0)	62 (30.8)
BMI, kg/m^2		
< 18.5	9 (5.4)	4 (2.0)
≥ 18.5 to < 25	67 (40.4)	97 (48.3)
≥ 25 to < 30	56 (33.7)	64 (31.8)
≥ 30	33 (19.9)	34 (16.9)
Unknown	1 (0.6)	2 (1.0)
Ethnicity		
Hispanic or Latin	14 (8.4)	19 (9.5)
Not Hispanic or Latin	146 (88.0)	158 (78.6)
Not reported or unknown	6 (3.6)	24 (11.9)
ECOG Performance Status		
0 or 1	148 (89.2)	195 (97.0)
2 or 3	18 (10.8)	6 (3.0)
Presence of bulky disease	60 (36.1)	93 (46.3)
Elevated baseline LDH levels	90 (54.2)	62 (30.8)
> 1 extra-nodal site	113 (68.1)	120 (59.7)
Ann Arbor stage		
I or II	42 (25.3)	15 (7.5)
III or IV	124 (74.7)	185 (92.0)
Unknown	0 (0)	1 (0.5)
Bone marrow involvement		
Yes	17 (10.2)	108 (53.7)
No	148 (89.2)	88 (43.8)
Indeterminate	1 (0.6)	5 (2.5)
Cell of origin		
ABC	31 (18.7)	N/A
GCB	77 (46.4)	N/A
Unclassified	26 (15.7)	N/A
Unknown	32 (19.3)	N/A
Double-hit lymphoma		
Yes	3 (1.8)	N/A
No	68 (41.0)	N/A
Unknown	95 (57.2)	N/A
Double-expressor lymphoma (MYC and BCL2)		
Yes	36 (21.7)	N/A
No	65 (39.2)	N/A
Unknown	65 (39.2)	N/A

ABC, activated B-cell like; BMI, body mass index; DLBCL, diffuse large B-cell lymphoma; ECOG, Eastern Cooperative Oncology Group; FL, follicular lymphoma; GCB, germinal center B-cell-like; LDH, lactate dehydrogenase; N/A, not available.

Figure 4. Deming regression fit between mTMTV and aTMTV read of TMTV in cubic root in (a) the combined data set, (b) DLBCL, (c) FL and (d) sensitivity and positive predictive values (precision) by lesion volume in the combined data set



95% CIs are calculated using the bootstrap (quantile) method.
aTMTV, automated TMTV; CI, confidence interval; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; mTMTV, manual TMTV; MTV, metabolic tumor volume; TMTV, total metabolic tumor volume.

- The slope and intercept of weighted Deming regression analyses revealed only moderate levels of systematic bias in the combined data set or the DLBCL subgroup; in the FL subgroup, the algorithm slightly overestimated TMTV (**Figures 4a–c**).
- The mean ± standard deviation (SD) difference between aTMTV and mTMTV was 0.10 ± 1.15 in the combined data set, −0.17 ± 0.93 in the DLBCL subgroup and 0.4 ± 1.83 in the FL subgroup.
- Overall sensitivity and precision for lesion detection were both more than 0.8 in the combined data set.
  - Performance was slightly lower in lesions that were 10 mL or smaller (mean sensitivity, 0.67; mean precision, 0.72) than for lesions greater than 10 mL (mean sensitivity and precision > 0.95; **Figure 4d**).
- In patients with DLBCL, mean ± SD patient-level sensitivity and precision of aTMTV were 0.80 ± 0.29 and 0.81 ± 0.26, respectively. In patients with FL, they were 0.74 ± 0.31 and 0.71 ± 0.31, respectively.
- aTMTV and mTMTV were strongly correlated in all demographic and disease characteristic subgroups (Pearson's $r \geq 0.89$), except the 'ethnicity not reported' subgroup ($r = 0.59$), and irrespective of the PET/CT scanner manufacturer used to obtain the image ($r \geq 0.94$; **Table 2**).

Table 2. Correlation and bias between aTMTV and mTMTV across patient subgroups and across images from different scanner manufacturers

Subgroup	n	Pearson's r	Intercept (95% CI)	Slope (95% CI)
Sex				
Male	159	0.92 (0.90, 0.94)	−0.08 (−0.50, 0.31)	1.04 (0.98, 1.11)
Female	208	0.97 (0.96, 0.98)	−0.38 (−0.72, −0.05)	1.07 (1.03, 1.12)
Age group, years				
< 65	237	0.95 (0.93, 0.96)	−0.38 (−0.67, −0.09)	1.07 (1.03, 1.12)
≥ 65	130	0.97 (0.95, 0.98)	−0.08 (−0.54, 0.36)	1.04 (0.98, 1.11)
BMI, kg/m^2				
≥ 18.5 to < 25	164	0.98 (0.96, 0.98)	−0.11 (−0.56, 0.29)	0.99 (0.93, 1.05)
≥ 25 to < 30	120	0.97 (0.95, 0.98)	−0.54 (−1.31, 0.14)	1.03 (0.94, 1.13)
≥ 30	67	0.98 (0.95, 0.99)	0.17 (−0.79, 0.74)	0.99 (0.92, 1.13)
Ethnicity				
Hispanic or Latin	33	1.00 (0.99, 1.00)	0.38 (−0.24, 1.03)	1.00 (0.93, 1.06)
Not Hispanic or Latin	304	0.95 (0.94, 0.96)	−0.27 (−0.53, −0.01)	1.05 (1.02, 1.10)
Not reported	30	0.59 (0.23, 0.80)	−1.65 (−5.22, −0.15)	1.33 (1.05, 2.00)
ECOG Performance Status				
0 or 1	343	0.94 (0.92, 0.95)	−0.37 (−0.63, −0.11)	1.08 (1.04, 1.13)
2 or 3	24	0.99 (0.97, 1.00)	−0.53 (−1.78, 0.53)	1.01 (0.93, 1.13)
Presence of bulky disease				
Yes	153	0.97 (0.96, 0.98)	−0.08 (−0.53, 0.36)	1.04 (0.99, 1.09)
No	214	0.89 (0.86, 0.92)	−0.43 (−0.82, −0.10)	1.09 (1.02, 1.17)
LDH level				
Normal	215	0.89 (0.86, 0.92)	−0.62 (−1.01, −0.30)	1.16 (1.10, 1.25)
Elevated	152	0.97 (0.95, 0.98)	−0.18 (−0.60, 0.25)	1.02 (0.97, 1.07)
Number of extra-nodal sites				
0–1	134	0.96 (0.94, 0.97)	−0.54 (−0.90, −0.23)	1.10 (1.05, 1.17)
> 1	233	0.95 (0.94, 0.96)	−0.07 (−0.41, 0.27)	1.03 (0.99, 1.08)
Ann Arbor stage				
I or II	57	1.00 (1.00, 1.00)	−0.69 (−1.59, −0.01)	1.11 (1.03, 1.25)
III or IV	309	0.94 (0.93, 0.95)	−0.18 (−0.44, 0.07)	1.05 (1.02, 1.09)
Bone marrow involvement				
Yes	125	0.92 (0.89, 0.95)	−0.46 (−0.91, −0.01)	1.11 (1.05, 1.18)
No	236	0.97 (0.96, 0.98)	−0.12 (−0.40, 0.15)	1.02 (0.98, 1.07)
Image scanner manufacturer				
CPS Innovations	30	0.98 (0.96, 0.99)	0.11 (−0.67, 0.82)	1.04 (0.96, 1.14)
GE Medical Systems	139	0.97 (0.96, 0.98)	−0.38 (−0.79, 0.00)	1.06 (1.01, 1.12)
Siemens	167	0.94 (0.92, 0.95)	−0.15 (−0.52, 0.24)	1.05 (1.00, 1.12)
Philips Medical Systems	22	0.99 (0.97, 0.99)	−0.52 (−0.97, 0.05)	1.10 (1.02, 1.19)

Patient subgroups with missing values and those with fewer than 10 patients were omitted from the subgroup analysis.
aTMTV, automated TMTV; BMI, body mass index; CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; LDH, lactate dehydrogenase; mTMTV, manual TMTV; TMTV, total metabolic tumor volume.

Conclusions

- **Although the aTMTV model was solely trained with DLBCL data, good performance and acceptable levels of bias for the quantification of TMTV were observed when tested using FL or DLBCL data.**
- **Higher variability among expert readers in the determination of small lesions may be a contributing factor for reduced algorithm performance for lesions 10 mL or smaller; this will be addressed in future studies.**
- **Good generalizability was observed among patients with different demographics and clinical characteristics, and across images from different PET/CT scanner manufacturers, which indicates that the model may have utility in a real-world setting.**
- **With further optimization and clinical validation, this model may provide a novel automated approach for lesion segmentation and TMTV quantification to inform the management of patients with FDG-avid lymphoma.**

References

1. Barrington SF, Meignan M. *J Nucl Med* 2019;60:1096–10. **2.** Sasanelli M *et al. Eur J Nucl Med Mol Imaging* 2014;41:2017–22. **3.** Kostakoglu L *et al. Blood* 2017;130(Suppl 1):824. **4.** Meignan M *et al. J Clin Oncol* 2016;34:3618–26. **5.** Al Tabaa Y *et al. Cancers (Basel)* 2021;13:5222. **6.** Capobianco N *et al. J Nucl Med* 2021;62:30–6. **7.** Weisman AJ *et al. Radiol Artif Intell* 2020;2:e200016. **8.** Jemaa S *et al. J Digit Imaging* 2020;33:888–94. **9.** Jemaa S *et al. Blood* 2019;134(Suppl 1):4666. **10.** Jemaa S *et al. Cancer Imaging* 2022;22:39.

Acknowledgments

This study was sponsored by F. Hoffmann-La Roche Ltd, Basel, Switzerland. Medical writing support was provided by Kayleigh Dodd PhD of PharmaGenesis Cardiff, Cardiff, UK, and funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland in accordance with Good Publication Practice (GPP 2022) guidelines (www.ismpp.org/gpp-2022).

Disclosures

TX, SL, SFM and TN are employees of F. Hoffmann-La Roche AG and hold stock/stock options. SJ is an employee of, and has patents issued or pending with Genentech, Inc. and holds stock/stock options with F. Hoffmann-La Roche AG. MK and SO are employees of Genentech, Inc. SB is an employee of Genentech, Inc. and holds stock/stock options in Hoffmann-La Roche AG and Genentech, Inc. AS-D is an employee of F. Hoffmann-La Roche AG. BM and WC are employees of Genentech, Inc. and hold stock/stock options in F. Hoffmann-La Roche AG. JL is an employee of Genentech, Inc. and holds stock/stock options in F. Hoffmann-La Roche AG, YungShin Global Holding and Yung Zip Chemical. RADC is an employee of Genentech, Inc., has patents issued or pending with, and holds stock/stock options in F. Hoffmann-La Roche AG. LK is a consultant for F. Hoffmann-La Roche AG and Genentech, Inc.

Presented at the 17th International Conference on Malignant Lymphoma (ICML), June 13–17, 2023, Lugano, Switzerland
Previously presented at the 2023 European Hematology Association (EHA) Hybrid Congress, June 8–15, 2023, Frankfurt, Germany