Definition of bulky disease in early stage Hodgkin lymphoma in computed tomography era: prognostic significance of measurements in the coronal and transverse planes

Anita Kumar,[1] Irene A. Burger,[2] Zhigang Zhang,[3] Esther N. Drill,[3] Jocelyn C. Migliacci,[1] Andrea Ng,[4] Ann LaCasce,[5] Darci Wall,[6] Thomas E. Witzig,[7] Kay Ristow,[7] Joachim Yahalom,[8] Craig H. Moskowitz,[1] and Andrew D. Zelenetz[1]

[1]Lymphoma Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA; [2]Department Medical Radiology, University Hospital Zurich, Switzerland; [3]Biostatistics and Epidemiology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA; [4]Department of Radiation Oncology, Brigham and Women's Hospital and Dana-Farber Cancer Institute, Boston, MA, USA; [5]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; [6]Department of Hematology, Mayo Clinic, Rochester, MN, USA; [7]Department of Hematology, Mayo Clinic, Rochester, MN, USA; and [8]Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

 doi:10.3324/haematol.2016.14184

Received: January 14, 2016.

Accepted: July 5, 2016.

Pre-published: July 6, 2016.

Correspondence: kumara2@mskcc.org

Supplement

Methodology for Determining Optimal Cutoff for Disease Bulk in Early Stage Hodgkin Lymphoma

Table S1. Univariate analysis of continuous transverse and coronal max diameters for RFS

	Relapse-free survival		
	HR	95% CI	P value
Transverse maximal diameter	1.22	1.09, 1.37	<0.001
Coronal maximal diameter	1.17	1.07, 1.27	<0.001

Identifying the optimal cutoff of transverse and coronal max diameters for RFS (separately):

Due to the small number of deaths (only 6 in total), we do not use overall survival time but only RFS time. Previous studies have variably defined tumor bulk as a prognostic indicator, ranging from 5-10cm. Based on the previous results and quantiles of the data (we insist that the cut-off points are between the 10^{th} and 90^{th} percentiles), we pre-determine some cut-off points for either transverse or coronal max diameters (see table below). We examine their significance levels using log-rank tests (for correlating with RFS).

Of note, we used a Cox proportional hazards model and log-rank tests to correlate various cut-off points with RFS. We did not use ROC analysis which requires a dichotomous endpoint. That is, whether or not patient developed progression at a fixed time point (i.e. 4-year RFS). We sought to examine the time to relapse, which includes not only the event status, but also when the event occurred. Therefore, a Cox proportional hazards model was used, not a ROC analysis.

The cut-off point resulting in the maximal significance level (i.e., smallest p-value) will be identified as the optimal cut-off, whose p-value will also be adjusted by the maximal chi-square method due to the fact that we have looked at multiple tests.[1]

We see that the optimal cutoff point for transverse max diameter is 7·0 which yields a p-value of 0·025. After the adjustment this p-value is around 0.046. The optimal cutoff point for coronal max diameter is 10·5 which yields a p-value of 0·0092. After the adjustment this p-value is around 0·014. Note that for the coronal max diameter several other cut-off points give close p-values: 6·0, 6·5, 7·0, 9·5, 11.5 and 12·0. As a matter of fact, if we use the concordance probability (reference 2) instead of the log-rank test p-value to identify the cut-off points, 7·0 will still be the optimal choice for transverse max diameter, but for coronal max diameter 4.5 through 8.5 and 9.5are better than 10·5. This implies that an optimal cut-off point for coronal max diameter cannot be determined clearly statistically, and several options are available.

Table S2. Transverse and coronal max diameter cutoff points and correlation with RFS.

Transverse	**P-value**	**# of pts > cutoff**		**Coronal**	**P-value**	**# of pts > cutoff**
3·0	0·938	168		4·0	0·208	161
3·5	0·535	152		4·5	0·151	147
4·0	0·766	148		5·0	0·126	136

4·5	0·482	133		5·5	0·066	130
5·0	0·177	121		6·0	0·017	119
5·5	0·332	111		6·5	0·015	109
6·0	0·121	100		7·0	0·012	91
6·5	0·070	88		7·5	0·066	80
7·0	***0·025***	**73**		8·0	0·091	75
7·5	0·055	57		8·5	0·054	64
8·0	0·081	46		9·0	0·112	55
8·5	0·174	37		9·5	0·019	45
9·0	0·086	33		10·0	0·034	41
9·5	0·062	26		***10·5***	***0·0092***	***36***
				11·0	0·042	30
				11·5	0·016	27
				12·0	0·0094	20

Seeking for a "combined" criterion:

We combined the two diameters (transverse and coronal) and identify a better predictor. Although 7·0 and 10·5 have been identified as the optimal cut-off points individually, their combination may not be the best in terms of the predictive capacity (as mentioned above, 10·5 for coronal max diameter, though having the smallest log-rank p-value, does not even provide the best predictive capacity separately). Therefore we look at several options in case there exist better predictor than "Transverse > 7·0 OR Coronal > 10·5". The table below shows the findings. Note that here the predictive capacity is evaluated quantitatively by the concordance probability.[2]

Table S3. Combined Criterion for disease bulk

Combined Criterion	# of pts > cutoff	Concordance Probability
Transverse > 7·0 OR Coronal > 4.5	72	0.574
Transverse > 7·0 OR Coronal > 5.0	72	0.581
Transverse > 7·0 OR Coronal > 5.5	72	0.601
Transverse > 7·0 OR Coronal > 6·0	70	0.629
Transverse > 7·0 OR Coronal > 6·5	68	0.622
Transverse > 7·0 OR Coronal > 7·0	***63***	***0.654***
Transverse > 7·0 OR Coronal > 7.5	58	0.643
Transverse > 7·0 OR Coronal > 8.0	57	0.629
Transverse > 7·0 OR Coronal > 8.5	51	0.619
Transverse > 7·0 OR Coronal > 9·5	41	0.615
Transverse > 7·0 OR Coronal > 10·5	34	0.619
Transverse > 7·0 OR Coronal > 11·5	25	0.619
Transverse > 7·0 OR Coronal > 12·0	19	0.620

The above results show that "Transverse > 7·0 OR Coronal > 7·0" is the best predictor for progression. We caution here that this choice is not associated with a significance level using re-sampling methods, nor is it based on any calibration (it is sort of a "pick the winner" strategy).

References:

1. Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. *Statistics in medicine.* Jan 15 2000;19(1):113-132.
2. Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika.* Dec 2005;92(4):965-970.

Figure S1. Relapse-free survival by presence of bulky disease with traditional definition (>10cm in transverse plane) in the combined modality therapy (CMT) group only.