RESEARCH ARTICLE

Manual Contouring Based Volumetric Evaluation for Colorectal Cancer with Liver Limited Metastases: A Comparison with RECIST

WJ Fang[1,2], KO Lam[2], SCY Ng[2], CW Choi[3], DLW Kwong[2], SS Zheng[1]*, VHF Lee[2]*

Abstract

Background: To compare response evaluation criteria in solid tumours (RECIST) and volumetric evaluation (VE) for colorectal cancer with liver-limited metastasis. Patients and Methods: VE of liver metastases was performed by manual contouring before and after chemotherapy on 45 pairs of computed tomography (CT) images in 36 patients who suffered from metastatic colorectal cancer (mCRC) with liver metastasis only. Cohen kappa was used to compare the agreement between VE and RECIST. Pearson correlation was performed for their comparison after cubic root transformation of the aggregate tumor volumes. Logistic regression was done to identify clinical and radiographic factors to account for the difference which may be predictive in overall response (OR). Results: There were 16 partial response (PR), 23 stable disease (SD) and 6 progressive disease (PD) cases with VE, and 14 PR, 23 SD and 8 PD with RECIST. VE demonstrated good agreement with RECIST (κ=0.779). Discordant objective responses were noted in 6 pairs of comparisons (13.3%). Pearson correlation also showed excellent correlation between VE and RECIST (r^2=0.966, p<0.001). Subgroup analysis showed that VE was in slightly better agreement with RECIST for enlarging lesions than for shrinking lesions (r^2=0.935 and r^2=0.780 respectively). No factor was found predictive of the difference in OR between VE and RECIST. Conclusions: VE exhibited good agreement with RECIST. It might be more useful than RECIST in evaluation shrinking lesions in cases of numerous and conglomerate liver metastases.

Keywords: Manual contouring - volumetric evaluation - RECIST - liver metastases - colorectal cancer

Asian Pac J Cancer Prev, **14 (7)**, 4151-4155

Introduction

Colorectal cancer is one of the leading cancer causes of mortality in the world and liver is the predominant site of metastases in majorities of this disease (Shin et al., 2012). Evaluation of the metastatic liver nodules after treatment including chemotherapy is of crucial importance which can certainly lead us to earlier and more effective salvage treatment in an attempt to prolong survival. Response Evaluation Criteria in Solid Tumours (RECIST) has become the most frequently used response criteria for solid tumours (William et al., 2013). Nevertheless, there are shortcomings and criticism regarding its use in a number of clinical circumstances such as gastrointestinal stromal tumor (GIST) which has created inevitable inter-observer variability and non-reproducible judgments (Yankelevitz et al., 2000; Reeves et al., 2007; Heussel CP et al., 2007; Benjamin et al., 2007; Plathow et al., 2008; Mantatzis et al., 2009; Galizia et al., 2011). As RECIST is primarily based on measurement of the longest diameters of the target lesions in the transverse section, it is often a clinical dilemma for radiologists to determine the longest diameters when the target lesions merge to form a conglomerate mass or split into even smaller lesions after treatment (Sohaib et al., 2000; Reiner et al., 2009). Instead, volumetric evaluation (VE) has gained a wider interest and acceptance as an alternative assessment method (Yankelevitz et al., 2000; Winer et al., 2003; Lemke et al., 2006; Vogl et al., 2008; Gavrielides et al., 2009; Frauenfelder et al., 2011; Galizia et al., 2011). An important theoretical advantage of VE is that it offers more accurate measurement and reflection of overall tumor burden in an organ which is with less inter-observer variability and better than the mere measurement of the maximum diameters of up to five indicator lesions per organ (Prasad et al., 2002; Tran et al., 2004; Reiner et al., 2009). With regard to the similarities and potential discrepancies between these two assessment criteria, we carried out a retrospective study to evaluate their agreement and correlation, in an attempt to derive a more reliable assessment method of tumour response after systemic therapy for patients suffering from colorectal cancer with liver metastases only.

*[1]First Affiliated Hospital, School of Medicine, Zhe Jiang University, China, [2]Department of Clinical Oncology, Queen Mary Hospital, The University of Hong Kong, [3]Department of Community Medicine and School of Public Health, The University of Hong Kong, Hong Kong, China *For correspondence: vhflee@hku.hk, shusenzheng@zju.edu.cn*

Table 1. Category of Objective Response Determined by RECIST and VE

Category	RECIST	VE
Complete response (CR)	Tumor disappearance	Tumor disappearance
Partial response (PR)	≥30% reduction in sum of maximal diameters	≥65% reduction in total volume of all metastatic lesions
Stable disease (SD)	Between that for partial response and stable disease	Between that for partial response and stable disease
Progressive disease (PD)	≥20% increase in sum of maximal diameters	≥73% increase in total volume of all metastatic lesions

Table 2. Patient Characteristics (n=36)

Characteristic	
Median age, years (range)	59.5 (41-77)
Gender	Number (%)
Male	22 (61.1)
Female	14 (38.9)
Location of primary tumor	
Colon	21 (58.3)
Rectum	13 (36.1)
Colon+Rectum	2 (5.6)
Chemotherapy regimen	
FOLFOX4	5 (13.9)
XELOX	8 (22.2)
FOLFOX4 + cetuximab	9 (25.0)
XELOX + cetuximab	13 (36.1)
FOLFIRI + cetuximab	1 (2.8)

FOLFOX4, oxaliplatin + fluorouracil + folinic acid; XELOX, oxaliplatin + capecitabine; FOLFIRI, irinotecan + fluorouracil + folinic acid

Materials and Methods

Study design and treatment schedule

Patients who had regular interval CT scans during treatment with 1st line chemotherapy +/- cetuximab as response assessment for their metastatic colorectal cancer (mCRC) with liver-limited disease were eligible for this study for tumour response using RECIST and VE. Unenhanced and dual-phase (arterial and porto-venous) contrast CT scans were performed by a 64 multidetector CT scanner (GE Medical Systems, Milwaukee, USA) at baseline and then after 3-4 cycles of chemotherapy +/- cetuximab. The CT images were then transferred to Eclipse Treatment Planning System version 8.0 (Varian Medical Systems, Palo Alto, CA, USA), a commercialized radiotherapy planning system for manual contouring. VE was performed by manual contouring of all liver metastases at baseline and after chemotherapy +/- cetuximab on every slice of image captured in porto-venous phase. The whole liver was also contoured in all patients as well. The contouring process was performed by two oncologists who had received prior training in radiology and radiation oncology. The volumes of the whole liver, all liver metastases and the residual normal liver (difference between the whole liver and liver metastases) were then obtained. The longest diameter of each target lesion (up to five lesions) on the liver was determined by electronic calipers and summed up for RECIST evaluation. VE criteria as well as RECIST were listed in Table 1 as described by previous literature (Therasse et al., 2006; Eisenhauer et al., 2009).

Patients

Thirty six patients treated with 1st line palliative chemotherapy +/- cetuximab (at the discretion of treating physicians when there was no KRAS mutation) for their mCRC with liver metastases only at the Department of Clinical Oncology, Queen Mary Hospital between January 2008 to December 2010 were retrospectively reviewed. Ethics approval from local institutional review board was sought prior to study commencement.

Statistical Analysis

Statistical Package for Social Sciences (SPSS) version 20.0 was used for the statistical analysis. A two-sided p-value smaller than or equal to 0.05 was considered statistically significant. The agreement between VE and RECIST was examined using Cohen kappa which were then sub-classified into five categories: poor (κ=0-0.20), fair (κ=0.21-0.40), moderate (κ=0.41-0.60), good (κ=0.61-0.80) and excellent (κ=0.81-1.00) as previously mentioned (Therasse et al., 2000; Mantatzis et al., 2009). Logistic regression was performed to investigate for factors, including age, sex, use of cetuximab, location of primary tumour, number of liver metastasis, largest diameter of the liver metastasis and the ratio of liver metastases to the volume of whole liver, which may be predictive of difference in OR between VE and RECIST. We further analysed the relationship between RECIST and VE. First of all, the aggregate volumes of liver metastases were obtained. If the total volume of the all metastatic liver target lesions were 20% larger than the baseline, it would be recorded as 1.2. Similarly if their total volumes were 20% smaller than the baseline, it would be 0.8. As VE is a three-dimensional measurement based the aggregate volumes of liver metastases while RECIST is a uni-dimensional measurement, we performed cubic root transformation of the data obtained from VE to facilitate reasonable comparisons with RECIST (Kundel and Polansky, 2003; Armato et al., 2004). This is well supported by the fact that the definitions of partial response (PR), stable disease (SD) and progressive disease (PD) in VE are the cubes of their counterparts in RECIST (Eisenhauer et al., 2009). The relationship between VE and RECIST was subsequently analyzed by Pearson correlation.

Results

There were 45 pairs of CT images for comparison in total (Table 2).

23 out of 36 patients received cetuximab in addition to systemic chemotherapy due to the absence of KRAS mutation (De Gramont et al., 2000; Oxnard et al., 2006; Cassidy et al., 2008; Bokemeyer et al., 2009; Van et al., 2009). Seven patients had 4 CT scans and one had 6 CT scans in total for further evaluation of tumor response after additional chemotherapy +/- cetuximab while the rest of 28 patients had only 2 sets (baseline and after treatment)

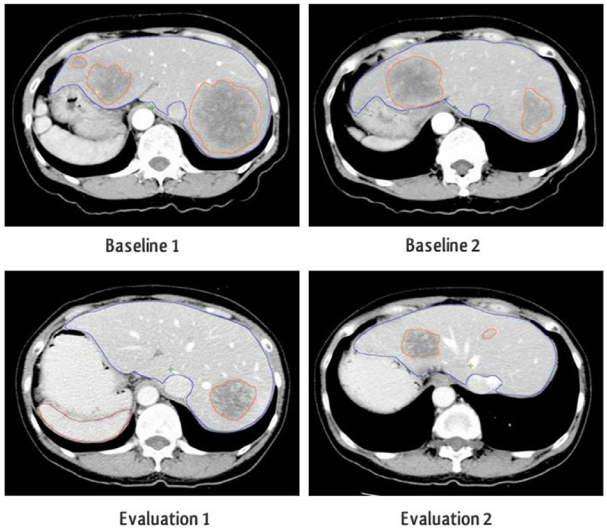

Figure 1. A patient with Several Metastatic Liver Nodules on Different Slices at Baseline (upper panel) and 3 Months after Systemic Chemotherapy (lower panel). Stable disease (SD) was noted in RECIST while partial response (PR) was confirmed in VE

Table 3. Comparison of OR Between RECIST and VE

Methods and OR		VE			
		PR	SD	PR	Total
RECIST	PR	13	1	0	14
	SD	3	20	0	23
	PD	0	2	6	8
	Total	16	23	6	45

Cohen Kappa = 0.779 (good agreement)

of CT scans. No patients had complete response (CR) after treatment. The mean number of liver metastases were 5.91 (range 1-16). Fourteen patients had more than 5 liver metastatic nodules and six patients had more than 10 liver metastatic nodules (maximum number of liver metastatic nodules was 16). Only the biggest and most discrete 5 lesions in these eleven patients were selected as target lesions according to RECIST but all liver metastases were contoured and their volumes obtained for VE.

Whole population analysis

There were 16 PR, 23 SD and 6 PD in VE, and 14 PR, 23 SD and 8 PD in RECIST, respectively. Cohen-kappa analysis revealed that VE was in good agreement with RECIST (κ=0.779) (Table 3).

Discordant OR between VE and RECIST was noticed in 6 pairs of comparison. Of these, SD was noted in RECIST while it was regarded as PR in VE in 3 pairs of comparisons. This was exemplified in Figure 1 in which a patient had stable disease revealed by RECIST but it would be considered as partial response by VE. The disproportionate asymmetrical shrinkage of tumours in the other two dimensions while maintaining the stability of the longest diameter of the target lesions accounted for PR by VE while still SD by RECIST. On the contrary, PR was noted in RECIST but it was SD in VE in one pair of comparison. The other 2 pairs of discordance exhibited PD in RECIST but SD in VE. After cubic root transformation of data in VE, Pearson correlation displayed an excellent correlation between VE and RECIST (r^2=0.966, p<0.001). Logistic regression failed to identify any predictive factor to account for the difference in OR between VE and RECIST.

Subgroup analysis

As mentioned, there were altogether 39 pairs of comparisons in which the effect evaluation of liver metastases is the same in both RECIST and VE. Interestingly, for the enlarging lesions confirmed including 6 pairs of PD and 4 pairs of SD, VE showed better correlation with RECIST (r^2=0.935, p<0.001) while the correlation was weaker (though still strong) for shrinking lesions including 13 pairs of PR and 16 pairs of SD (r^2=0.780, p<0.001).

Discussion

There have been various previous studies comparing volumetric measurements against uni-dimensional measurements with contrasting results (Prasad et al., 2002; Tran et al., 2004; Heussel et al., 2007; Mantatzis et al., 2009). The study conducted by Tran et al and Heussel et al demonstrated good agreement (discordance rate 3.3% and 13% respectively) while Prasad et al revealed fair agreement only (discordance rate 31.6%) (Prasad et al., 2002; Tran et al., 2004; Heussel et al., 2007). Our study result was in accord with that published by Mantatzis et al (Mantatzis et al., 2009). In his study, magnetic resonance imaging (MRI) was used to evaluate tumour response in 57 patients with liver metastases from colorectal cancer. RECIST and VE were in good agreement when comparing patients' overall response and individual lesions (κ=0.735 and κ=0.741 respectively compared with κ=0.779 in our study). However, our discordance rate was 13.3% which was better than that in his study. One of the major differences between our study and Mantatzis' study was that we performed cubic root transformation of the data in VE before performing correlation analysis with RECIST. It was because VE is a three-dimensional measurement in which its data should be cubic-root transformed in order to allow meaningful comparisons with the uni-dimensional data in RECIST. In fact, we demonstrated excellent correlation between VE and RECIST after transformation was conducted. And we also produced highly comparable results by using multidetector CT scanner taking less than 20 seconds for a scan, obviating the need of at least 30 minutes for a MRI scan (Maughan et al., 2011).

Though manually demanding and time-consuming taking at least 15 minutes to clearly differentiate tumours from adjacent structures like bile ducts, portal vein and inferior vena cava and complete volumetric contouring for one set of images, volumetric measurement should be praised for its superiority of accuracy and reliability over RECIST, as found to be especially practicable and reproducible in our patients with liver metastases only. Evolution of novel medical imaging tools may soon overcome this tedious and cumbersome manual task (Gavrielides et al., 2009). Additional manual modifications of the contours of the target lesions are still necessary

despite this novel auto-segmentation technique.

As VE reflects the actual change in size of the lesions and it has good agreement with RECIST shown in our study and previous literature, it makes us feel comfortable and confident to use the simpler method RECIST in daily clinical practice. However, our subgroup analysis offered us some new information. Of note, it was found that VE had better agreement with RECIST for enlarging lesions than for shrinking lesions. This interesting finding may imply that VE may be more useful in evaluating shrinking lesions of liver metastasis because RECIST was less sensitive as compared with VE. Of course, the number of patients in these subgroups was relatively small, which was one of the limitations of our study. Further studies are warranted to confirm this finding.

In conclusion, manual contouring based VE showed good agreement with RECIST for evaluation of tumour response of liver metastases in mCRC. Of note, it might be more sensitive especially in monitoring numerous, conglomerate and shrinking lesions in liver metastases, if the time of radiologists and oncologists is not a concern.

Acknowledgements

This study was supported by the National S&T Major Project (No. 2012ZX10002017) and National Basic Research Program (973 Program: 2009CB522403). We would also like to express our gratitude to Foundation for National Natural Science Foundation of China (No. 81121002, No. 81272679) and Major Scientific Project of Zhejiang Province (2012C13014-2) for their support to our study. The author(s) declare that they have no competing interests.

References

Armato SG, 3rd, Oxnard GR, MacMahon H, et al (2004). Measurement of mesothelioma on thoracic CT scans: a comparison of manual and computer-assisted techniques. *Med Phys*, **31**, 1105-15.

Benjamin RS, Choi H, Macapinlac HA, et al (2007). We should desist using RECIST, at least in GIST. *J Clin Oncol*, **25**, 1760-4.

Bokemeyer C, Bondarenko I, Makhson A, et al (2009). Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer. *J Clin Oncol*, **27**, 663-71.

Cassidy J, Clarke S, Diaz-Rubio E, et al (2008). Randomized phase III study of capecitabine plus oxaliplatin compared with fluorouracil/folinic acid plus oxaliplatin as first-line therapy for metastatic colorectal cancer. *J Clin Oncol*, **26**, 2006-12.

De Gramont A, Figer A, Seymour M, et al (2000). Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *J Clin Oncol*, **18**, 2938-47.

Eisenhauer EA, Therasse P, Bogaerts J, et al (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*, **45**, 228-47.

Frauenfelder T, Tutic M, Weder W, et al (2011). Volumetry: an alternative to assess therapy response for malignant pleural mesothelioma? *Eur Respir J*, **38**, 162-8.

Gavrielides MA, Kinnard LM, Myers KJ and Petrick N(2009). Noncalcified lung nodules: volumetric assessment with thoracic CT. *Radiology*, **251**, 26-37.

Galizia MS, Tore HG, Chalian H and Yaghmai V(2011). Evaluation of hepatocellular carcinoma size using two-dimensional and volumetric analysis: effect on liver transplantation eligibility. *Acad Radiol*, **18**, 1555-60.

Heussel CP, Meier S, Wittelsberger S, et al (2007). Follow-up CT measurement of liver malignoma according to RECIST and WHO vs. volumetry. *RoFo*, **179**, 958-64.

Kundel HL and Polansky M(2003). Measurement of observer agreement. *Radiology*, **228**, 303-8.

Lemke AJ, Brinkmann MJ, Schott T, et al (2006). Living donor right liver lobes: preoperative CT volumetric measurement for calculation of intraoperative weight and volume. *Radiology*, **240**, 736-42.

Laghi A(2007). Multidetector CT (64 Slices) of the liver: examination techniques. *Eur Radiol*, **17**, 675-83.

Mantatzis M, Kakolyris S, Amarantidis K, Karayiannakis A and Prassopoulos P (2009). Treatment response classification of liver metastatic disease evaluated on imaging. Are RECIST unidimensional measurements accurate? *Eur Radiol*, **19**, 1809-16.

Maughan TS, Adams RA, Smith CG, et al (2011). Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet*, **377**, 2103-14.

Oxnard GR, Armato SG, 3rd and Kindler HL(2006). Modeling of mesothelioma growth demonstrates weaknesses of current response criteria. *Lung Cancer*, **52**, 141-8.

Prasad SR, Jhaveri KS, Saini S, et al (2002). CT tumor measurement for therapeutic response assessment: comparison of unidimensional, bidimensional, and volumetric techniques initial observations. *Radiology*, **225**, 416-9.

Plathow C, Klopp M, Thieke C, et al (2008). Therapy response in malignant pleural mesothelioma-role of MRI using RECIST, modified RECIST and volumetric approaches in comparison with CT. *Eur Radiol*, **18**, 1635-43.

Reeves AP, Biancardi AM, Apanasovich TV, et al (2007). The Lung Image Database Consortium (LIDC): a comparison of different size metrics for pulmonary nodule measurements. *Acad Radiol*, **14**, 1475-85.

Reiner CS, Karlo C, Petrowsky H, et al (2009). Preoperative liver volumetry: how does the slice thickness influence the multidetector computed tomography- and magnetic resonance-liver volume measurements? *J Comput Assist Tomogr*, **33**, 390-7.

Sohaib SA, Turner B, Hanson JA, et al (2000). CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *Br J Radiol*, **73**, 1178-84.

Shin SJ, Ahn JB, Choi JS, et al (2012). Implications of clinical risk score to predict outcomes of liver-confined metastasis of colorectal cancer. *Surg Oncol*, **21**, e125-30.

Therasse P, Arbuck SG, Eisenhauer EA, et al (2000). New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst*, **92**, 205-16.

Tran LN, Brown MS, Goldin JG, et al (2004). Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. *Acad Radiol*, **11**, 1355-60.

Vogl TJ, Naguib NN, Eichler K, et al (2008). Volumetric evaluation of liver metastases after thermal ablation: long-term results following MR-guided laser-induced thermotherapy. *Radiology*, **249**, 865-71.

Van Cutsem E, Kohne CH, Hitre E, et al (2009). Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med*, **360**, 1408-17.

Winer-Muram HT, Jennings SG, Meyer CA, et al(2003). Effect of varying CT section width on volumetric measurement of lung tumors and application of compensatory equations. *Radiology*, **229**, 184-94.

William WN, Jr., Pataer A, Kalhor N, et al (2013). Computed tomography RECIST assessment of histopathologic response and prediction of survival in patients with resectable non-small-cell lung cancer after neoadjuvant chemotherapy. *J Thorac Oncol*, **8**, 222-8,

Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B and Henschke CI (2000). Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology*, **217**, 251-6.