

Dose-effect relationships of holmium-166 radioembolization in colorectal cancer

Caren van Roekel*¹, Remco Bastiaannet¹, Maarten L.J. Smits, Rutger C. Bruijnen, Arthur J.A.T. Braat, Hugo W.A.M. de Jong, Sjoerd G. Elias, Marnix G.E.H. Lam

* Corresponding author

¹ Contributed equally

Institution from which the work originated:

University Medical Center Utrecht, Utrecht University

Heidelberglaan 100

3584 CX Utrecht

The Netherlands

Corresponding author:

Caren van Roekel

+31 88 75 746 55

j.vanroekel@umcutrecht.nl

University Medical Center Utrecht

Heidelberglaan 100

3584 CX Utrecht

The Netherlands

ORCID: [0000-0002-1247-8103](https://orcid.org/0000-0002-1247-8103)

Running title: Dose-effect Ho-166 radioembolization

ABSTRACT

INTRODUCTION Radioembolization is a treatment option for colorectal cancer (CRC) patients with inoperable, chemorefractory hepatic metastases. Personalized treatment requires established dose thresholds. Hence, the aim of this study was to explore the relation between dose and effect (i.e. response and toxicity) in CRC patients treated with holmium-166 (^{166}Ho) radioembolization. **MATERIALS AND METHODS** CRC patients treated in the HEPAR II and SIM studies were analyzed. Absorbed doses were estimated using the activity distribution on post-treatment ^{166}Ho -SPECT/CT. Metabolic response was assessed using the change in total lesion glycolysis on ^{18}F FDG-PET/CT between baseline and three-months follow-up. Toxicity between treatment and three months was evaluated according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5, and its relation with parenchymal-absorbed dose was assessed using linear models. The relation between tumor-absorbed dose and patient- and tumor-level response was analyzed using linear mixed-models. Using a threshold of 100% sensitivity for response, the threshold for a minimal mean tumor-absorbed dose was determined and its impact on survival was assessed. **RESULTS** Forty patients were included. The median parenchymal-absorbed dose was 37 Gy (range 12-55 Gy). New CTCAE grade ≥ 3 clinical and laboratory toxicity were present in eight and seven patients, respectively. For any clinical toxicity (highest grade per patient), the mean difference in parenchymal dose (Gy) per step increase in CTCAE grade category was 5.75 (95% confidence interval (CI) 1.18-10.32). On a patient level, metabolic response was: complete response (CR) $n=1$, partial response (PR) $n=11$, stable disease (StD) $n=17$ and progressive disease (PD) $n=8$. The mean tumor-absorbed dose was 84% higher in patients with CR/PR than in patients with PD (95%CI: 20-180%). Survival for patients with a mean tumor-absorbed dose >90 Gy was significantly better than for patients with a mean tumor-absorbed dose <90 Gy (hazard ratio=0.16, 95%CI 0.06-0.511). **CONCLUSION** A significant dose-response relationship in CRC patients treated with ^{166}Ho -radioembolization was established and a positive association between toxicity and parenchymal dose was found. For future

patients, it is advocated to use ^{166}Ho -scout to select patients and personalize the administered activity targeting a mean tumor-absorbed dose of >90 Gy and a parenchymal dose <55 Gy.

KEY WORDS: radioembolization, holmium, dosimetry

INTRODUCTION

Colorectal cancer (CRC) is one of the most common types of cancer worldwide (1). The liver is the first site of hematogenous spread and 70-80% of patients with hepatic metastases are deemed unresectable because of tumor size, location, multifocality, or inadequate hepatic reserve (2). Hence, the majority of patients with metastatic CRC cannot be cured. Palliative treatment generally consists of several lines of systemic chemotherapy. If the available chemotherapeutic options fail, treatment with radioembolization should be considered for patients with liver-only or liver-dominant disease (3).

During radioembolization, radioactive microspheres are delivered intra-arterially to hepatic tumors. The rationale of this treatment is to administer a high local radiation dose to the tumors, while relatively sparing the healthy liver parenchyma by using the predominant arterial blood flow to tumors. Currently, three types of microspheres are available: yttrium-90 (^{90}Y) resin (SIRspheres[®], Sirtex), ^{90}Y glass (TheraSphere[®], BTG/Boston Scientific) and holmium-166 (^{166}Ho) microspheres (Quiremspheres[®], Quirem Medical).

One advantage of ^{166}Ho -radioembolization is that treatment can be preceded by a scout dose of the same microspheres, using only limited activity (250 MBq). This ^{166}Ho -scout has proven to be a more accurate predictor of the distribution of the treatment dose (4). Another advantage is that ^{166}Ho -microspheres can be visualized by both MRI and SPECT/CT (5). The safety and efficacy of ^{166}Ho -radioembolization was determined in the HEPAR and SIM studies (6-9). In these studies, activity calculation was based on a whole-liver absorbed dose of 60 Gy. To allow for personalized, or optimized treatment, reference levels for efficacy and toxicity are needed (10). Hence, the aims of this study were to determine the relationship between dose and toxicity and to determine the relation between dose and metabolic response, in CRC patients who were treated with ^{166}Ho -radioembolization.

MATERIALS AND METHODS

Patients

This was a retrospective analysis of CRC patients who were treated with ^{166}Ho -radioembolization in the HEPAR II (NCT01612325 (6)) and the SIM (NCT02208804 (8)). Before study inclusion, all patients provided written informed consent. The institution's Medical Ethics Committee approved both studies. The CRC patients of the HEPAR II study were already part of a preliminary mixed tumor-type cohort analysis and were also included in this CRC-only analysis (11).

Treatment Procedures

During work-up, laboratory and clinical examinations were performed and patients underwent multiphasic liver CT and ^{18}F FDG-PET/CT at a median of 16 days before treatment (range 6-42). Pre-treatment activity calculation was performed using a method similar to the medical internal radiation dosimetry (MIRD) method (12). The injected activity (IA) to reach an average absorbed dose of 60 Gy in the target volume was calculated as (7):

$$IA_{(MBq)} = \text{target volume weight (kg)} * 3780 \left(\frac{MBq}{kg} \right)$$

The IA was not adjusted for lung shunt fraction, in line with the instructions for use for Quiremspheres. Since the abundance of gamma photons invokes detector dead-time, patients underwent a quantitative ^{166}Ho -SPECT/CT to assess the therapeutic dose distribution three to five days after treatment. An aortic blood pool activity-based threshold and a volume restriction of \geq five mL were used to automatically define tumors on baseline ^{18}F FDG-PET/CT using the software package ROVER (ABX GmbH, Radeberg, Germany). Liver contours were manually delineated. Tumor- and liver contours were transferred to post-

treatment quantitative ^{166}Ho -SPECT/CT reconstructions using automatic rigid CT-to-CT co-registration (13). Tumor- and parenchymal-absorbed doses were assessed using the activity in transferred contours (Figure 1). All activity was assumed to be present in the target volume. Further details are provided in the supplemental data (Methods used for response estimation).

Toxicity Evaluation

The emergence of clinical toxicity between treatment and three months post-treatment was recorded, with exception of clinical adverse events during the first week after treatment, to allow for distinction between adverse events due to embolization and adverse events due to radiation. Laboratory toxicity between treatment and three months post-treatment was evaluated using the following parameters: albumin, alkaline phosphatase (AP), alanine aminotransferase (ALAT), aspartate aminotransferase (ASAT), bilirubin and gamma-glutamyltransferase (GGT). Common terminology criteria for adverse events (CTCAE) version 5.0 was used for grading (14). Since version 5.0 allows for higher values of laboratory parameters when these were already abnormal at baseline, relative change in laboratory values between baseline and three months follow-up was calculated as well. Furthermore, presence of ascites and encephalopathy (as part of radioembolization-induced liver disease (REILD)) was determined at three-months follow-up.

Efficacy Evaluation

Metabolic response to treatment was evaluated on ^{18}F FDG PET/CT at three-month follow-up. Tumors were automatically defined based on standardized uptake value (SUV) and lesion total lesion glycolysis (TLG) was obtained. To avoid misidentification, baseline and follow-up images were evaluated in parallel. Metabolic response of hepatic lesions was defined based on the change in TLG between baseline and follow-up, according to the PERCIST criteria (15). Hepatic tumor response was also assessed according to the Response evaluation criteria in solid tumours (RECIST) version 1.1 (16).

Statistical Analyses

Patient demographics and treatment characteristics were summarized using descriptive analyses. The associations between CTCAE toxicity grade and parenchymal-absorbed dose and relative change in laboratory parameters and parenchymal-absorbed dose were assessed using linear regression. For clinical relevance, CTCAE grading of any clinical and laboratory toxicity was also dichotomized in the following categories: grade 0/I/II versus grade III/IV/V and analyzed using logistic regression. All toxicity analyses were adjusted for response to therapy, previous treatment and tumor load (defined as percentage involvement of the liver by tumors).

The relationship between tumor-absorbed dose and response was analyzed using a linear mixed-effects regression model with tumor-absorbed dose (log-transformed) as the dependent variable. Analyses were adjusted for previous treatment and tumor load. More details of this method can be found in the supplemental data (Elaboration on statistical analysis). An ROC analysis, accounting for clustered data, was done to determine the discriminatory power of tumor dose in response estimation (17). Using a threshold of 100% sensitivity for response (CR/PR), the threshold for a minimal mean tumor-absorbed dose was determined and its impact on survival was assessed. The agreement between response according to PERCIST and according to RECIST was analyzed using Cohen's kappa, with disagreements weighted according to their squared distance from perfect agreement.

Overall survival was defined as the interval between treatment and death from any cause. Cox regression models were made using Firth's correction for small sample bias (18). Analyses were adjusted for tumor load, parenchymal dose and the presence of extrahepatic disease at baseline. Analyses were performed using R statistical software, version 3.6.2 for Windows. We report effect estimates with associated 95% confidence intervals (CIs) and corresponding two-sided p-values.

RESULTS

Forty patients were included, with a total of 133 hepatic lesions. Three patients did not have follow-up imaging for tumor-response assessment and were only included in the survival- and toxicity analyses. Patient- and treatment characteristics are summarized in Table 1.

Toxicity

The median parenchymal-absorbed dose was 37 Gy (range 12-55 Gy). Toxicity incidence during three months post-treatment and CTCAE grades are summarized in Table 2. New grade ≥ 3 clinical toxicity was present in eight patients (20%) and new grade ≥ 3 laboratory toxicity was present in seven patients (17.5%). There was one patient (2.5%) who developed REILD, evidenced by hyperbilirubinemia, hypoalbuminemia and ascites, without evidence of progression or biliary obstruction. The mean parenchymal-absorbed dose of this patient was 34 Gy.

The results of the linear clinical toxicity regression analyses suggested a positive association between higher parenchymal dose and increase in CTCAE grade clinical toxicity (Table S1). The mean difference in parenchymal dose for patients with CTCAE grade 0/1/2 any clinical toxicity versus CTCAE grade 3/4/5 was 11.6 Gy (95%CI 3.4-19.7, $p=0.0070$). The odds ratio for CTCAE grade 3/4/5 any clinical toxicity versus CTCAE grade 0/1/2 per 10 Gy increase in parenchymal dose was 7.62 (95%CI 1.95-249.03, $p=0.0063$) (Table S2).

For laboratory toxicity, the results of the linear regression analyses for both the CTCAE grades and the relative change in laboratory parameters showed that a higher parenchymal-absorbed dose is related with an increase in laboratory toxicity (Table S3a-b and Figures 2a-f).

Efficacy

Solely based on the metabolic response of measurable hepatic metastases at baseline, there was one patient with CR, PR was reached in 11 patients, 17 patients had StD and PD was found in eight patients at three-months follow-up. On a lesion level, CR occurred in 23 lesions, PR in 20 lesions, StD in 49 lesions and 23 lesions were progressive. A significant dose-response relation was found on patient- and tumor-level. The mean tumor-absorbed dose was 77% higher in patients with CR/PR than in patients with PD (95%CI: 18-164%, $p=0.011$) and the mean absorbed dose was 95% higher in lesions with CR than in lesions with PD (34-188%, $p=0.00065$) (Table 3). Mean absorbed doses per response category are visualized in Figure 3. Based on ROC analysis, the ability of tumor-absorbed dose to discriminate between patients with and without metabolic response was 0.671 (95%CI: 0.54-0.80) and the ability of mean tumor-absorbed dose per patient to differentiate between responders and non-responders was 0.698 (95%CI: 0.45-0.95) (Figure 4a,b). At a mean tumor-absorbed dose threshold with 100% sensitivity (95%CI: 48-100%) for CR/PR at a patient level (90 Gy), specificity was 38% (95%CI: 21-56%) . At a tumor-level, without accounting for clustered data, sensitivity was 100% using a tumor-absorbed dose threshold of 80 Gy (95%CI: 74-100%) and specificity was 41% (95%CI: 31-51%). Agreement between PERCIST and RECIST was minimal, with $\kappa=0.345$ (95%CI 0.14-0.55). Anatomic response was lower than metabolic response in 15 cases (40.5%) and higher in seven cases (18.9%).

Survival

Median overall survival was 10.7 months (95% CI: 7.2 – 13.4). Survival was significantly different in patients without a metabolic response (including the development of new intra- or extrahepatic lesions) versus patients with a metabolic response (HR=2.34, 95% CI 1.09 - 5.69, $p=0.029$). After adjusting for tumor load, extrahepatic disease at baseline and parenchymal dose, the HR for non-responders was 2.54 (95% CI: 1.13 - 6.52, $p=0.023$). Median overall survival in responders was 14.8 months (95% CI 14.2 - ∞ , $n=8$) versus 8.6 months (95% CI 6.4 – 13.4 months, $n=29$) in non-responders

(Figure 5a-b). Furthermore, there was a significant difference in overall survival between patients with a mean tumor-absorbed dose >90 Gy versus a mean tumor-absorbed dose <90 Gy (HR=0.16, 95%CI 0.06-0.511), p=0.0031), Figure 6.

DISCUSSION

Building on the establishment of a dose-response relationship in patients treated with ¹⁶⁶Ho-radioembolization by Bastiaannet et al. (11), this study explored the dose-response relationship in a homogenous population of patients with CRC only. Furthermore, dose-toxicity relationships were studied. Our results suggest a positive association between a higher parenchymal-absorbed dose and increase in CTCAE grade toxicity, both for clinical and laboratory toxicity. Furthermore, our data unveils, both at a lesion- and at a patient-level, a significant dose-response relationship. Also, a mean tumor-absorbed dose >90 Gy – the minimal mean tumor-absorbed dose in the group of patients with CR/PR - was associated with a significantly longer survival.

In this study, treatment with radioembolization was well tolerated. The most frequent clinical adverse events were CTCAE grade 1-2 abdominal pain, nausea and fatigue. These adverse events are well-known side effects of treatment with radioembolization (19). One patient of our study died of hepatic failure. This safety profile is compliant with the results of the MORE study, which showed that treatment with ⁹⁰Y-resin radioembolization is safe in a patient population highly comparable to ours, namely CRC patients who received several lines of prior chemotherapy (20). A study on the safety of ⁹⁰Y-glass radioembolization in CRC patients showed similar results, with the most frequent side effects being fatigue, abdominal pain and nausea (21). The incidence of grade ≥3 laboratory toxicity is also comparable between the three types of microspheres (20,21).

Regarding efficacy, the metabolic response rate (CR/PR) at a tumor level was 36%, comparable to previous dose-response data on resin microspheres in a similar patient cohort, treated in the same hospital (22). In other studies, higher tumor-response rates up to 75% were found, with dose thresholds of 46 and 60 Gy (23,24). However, it is difficult to compare these studies with our study, as those patients were less heavily pretreated or received concomitant systemic therapy (23,24). Also, their thresholds cannot be compared to our data, as there are major differences between microsphere types in specific activity, size, number of particles and half-life. There was only minimal agreement in response between the PERCIST and the RECIST assessments. In 15 patients of our study, RECIST underestimated response according to PERCIST. This finding is in accordance with other studies comparing these response assessments after radioembolization (25,26). Metabolic response assessment is not hampered by the presence of necrosis, cystic changes and hemorrhage, as can be the case with size evaluation on transaxial images (27). Moreover, several studies found that changes in functional metrics, such as TLG, were related with overall survival and were more accurate predictors than anatomic changes (22,27,28).

Although the majority of our patients underwent ≥ 2 prior lines of systemic treatment, the response rate seems suboptimal. Before treatment, patients with CRC are currently selected based on clinical criteria, such as WHO performance status and progressive disease after several lines of chemotherapy (29). In case patients are deemed eligible for treatment with radioembolization, a second selection criterion should be the activity distribution based on either ^{99m}Tc -MAA or ^{166}Ho -scout. Based on the results of this study, we would argue that patients should only be selected for treatment if there is a favorable activity distribution with a sufficient mean tumor-absorbed dose >90 Gy and a parenchymal-absorbed dose of <55 Gy. Although a causal relationship cannot be claimed solely based on these observational data, the findings of this study suggest that below a mean tumor-absorbed dose of 90 Gy, metabolic response seems unlikely. However, since the discriminatory power of absorbed dose for response is limited, this number should be used with caution.

The need for personalized dosimetry is widely accepted, with several studies showing a dose-response relationship in CRC patients treated with ^{90}Y -resin radioembolization (22-24,28). There also is growing evidence for the possibility of improving treatment outcomes by using personalized treatment planning in radioembolization (10,11). However, thus far, the DOSISPHERE study was the only study implementing personalized radioembolization planning in a prospective clinical study, investigating the tumor-absorbed dose and response rate in HCC patients using a standard versus a personalized dosimetric approach with ^{90}Y glass microspheres. Preliminary results showed that both the response rates and tumor-absorbed doses were significantly higher in the personalized dosimetry arm (30).

Strengths of this study are the homogenous patient population, the standardized methods for tumor delineation, the use of a mixed-effects regression model accounting for clustered data, and the analyses of both safety and efficacy. This study also has several limitations. First, it is a single-center retrospective evaluation and there is a level of subjectivity in identifying the response of existing lesions, possibly leading to inter-operator variations in estimated doses. Second, the sample size was limited and due to the low incidence of toxicity, there was not enough data to draw a strong conclusion on the maximum tolerable parenchymal dose. Furthermore, the discriminatory value of absorbed dose for response estimation is limited and a causal dose-response relationship cannot be claimed based on these observational data. Hence, the reference values obtained should be interpreted with utmost caution and only be used as a direction. The rigid co-registrations used in this study are likely affected by differences in patient positioning, differences in breath-hold policy and the relatively low resolution of the low-dose CT of the SPECT/CT. The resulting (local) errors are likely to propagate as underestimated tumor doses and (slightly) overestimated parenchymal doses, which contributes to the error in each response category, decreasing statistical power.

In future studies on radioembolization in CRC patients, personalized dosimetry should be used. By using dosimetry-based optimized treatment planning, treatment doses can be tailored to the individual patient to acquire a maximum response while minimizing the chance of toxicity. As the incidence of toxicity was low, it is difficult to establish an absolute threshold for a maximum parenchymal dose. At the same time, it is likely that the parenchymal-absorbed dose threshold is different for each individual patient, dependent on many clinical characteristics. We therefore advise a pragmatic and clinically feasible approach, with activity calculation in order to obtain a sufficient tumor-absorbed dose and a parenchymal-absorbed dose of up to 55 Gy, dependent on individual patient characteristics. With a median parenchymal-absorbed dose of 37 Gy and a maximum of 55 Gy, this was proven to be a safe approach, with only one case of REILD. Furthermore, those patients for whom no meaningful mean tumor-absorbed dose (>90 Gy) can be reached at an acceptable parenchymal-absorbed dose threshold should be excluded from radioembolization treatment. On a tumor-level, based on our results, treatment strategy should be adjusted to guarantee a tumor-absorbed dose of at least 80 Gy for every tumor. Partition-modeling and multiple injection positions can be used to reach that objective. In other words, planning should be primarily based on applying a safe parenchymal-absorbed dose threshold, and selection of patients on a sufficient tumor-absorbed dose.

CONCLUSION

In CRC patients treated with ^{166}Ho -radioembolization, a positive association between tumor-absorbed dose and metabolic response was established. Survival for patients with a mean tumor-absorbed dose >90 Gy was significantly better than for patients with a mean tumor-absorbed dose <90 Gy. There also was a positive association between parenchymal-absorbed dose and both laboratory and clinical toxicity.

A treatment approach with selection of patients based on the activity distribution of the ^{166}Ho -scout and personalized treatment activity calculation is advocated.

DISCLOSURE

Conflicts of interest:

ML is a consultant for Boston Scientific and Terumo. MS and AB have served as speakers for BTG and Terumo.

The Department of Radiology and Nuclear Medicine of the UMC Utrecht receives royalties from Quirem Medical and research support from Terumo and Quirem Medical.

No other potential conflicts of interest relevant to this article exist.

KEYPOINTS

Question: What is the relation between dose and effect (i.e. response and toxicity) in CRC patients treated with holmium-166 radioembolization?

Pertinent findings: A significant dose-response relation was established. A positive relation was found between parenchymal-absorbed dose and toxicity. A mean tumor-absorbed dose >90 Gy was associated with improved overall survival.

Implications for patient care: For future patients, it is advocated to use ^{166}Ho -scout to select patients and personalize the administered activity targeting a mean tumor-absorbed dose of >90 Gy.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394-424.
2. Donadon M, Ribero D, Morris-Stiff G, Abdalla EK, Vauthey JN. New paradigm in the management of liver-only metastases from colorectal cancer. *Gastrointest Cancer Res.* 2007;1:20-27.
3. Network NCC. NCCN Guidelines Version 1.2018 Colon Cancer. https://oncolife.com.ua/doc/nccn/Colon_Cancer.pdf. Accessed September 12, 2019.
4. Smits MLJ, Dassen MG, Prince JF, et al. The superior predictive value of (166)Ho-scout compared with (99m)Tc-macroaggregated albumin prior to (166)Ho-microspheres radioembolization in patients with liver metastases. *Eur J Nucl Med Mol Imaging.* 2020;47:798-806.
5. Reinders MTM, Smits MLJ, van Roekel C, Braat A. Holmium-166 Microsphere Radioembolization of Hepatic Malignancies. *Semin Nucl Med.* 2019;49:237-243.
6. Prince JF, van den Bosch M, Nijsen JFW, et al. Efficacy of radioembolization with (166)Ho-microspheres in salvage patients with liver metastases: a phase 2 study. *J Nucl Med.* 2018;59:582-588.
7. Smits MLJ, Nijsen JFW, van den Bosch MAAJ, et al. Holmium-166 radioembolisation in patients with unresectable, chemorefractory liver metastases (HEPAR trial): a phase 1, dose-escalation study. *The Lancet Oncology.* 2012;13:1025-1034.
8. van den Hoven AF, Prince JF, Bruijnen RC, et al. Surefire infusion system versus standard microcatheter use during holmium-166 radioembolization: study protocol for a randomized controlled trial. *Trials.* 2016;17:520.
9. Braat AJAT, Bruijnen RCG, van Rooij R, et al. Additional holmium-166 radioembolisation after lutetium-177-dotatate in patients with neuroendocrine tumour liver metastases (HEPAR PLuS): a single-centre, single-arm, open-label, phase 2 study. *The Lancet Oncology.* 2020;21:561-570.
10. Kafrouni M, Allimant C, Fourcade M, et al. Retrospective voxel-based dosimetry for assessing the ability of the body-surface-area model to predict delivered dose and radioembolization outcome. *J Nucl Med.* 2018;59:1289-1295.

11. Bastiaannet R, van Roekel C, Smits MLJ, et al. First evidence for a dose-response relationship in patients treated with (166)Ho-radioembolization: a prospective study. *J Nucl Med*. 2019;61:608-612.
12. Salem R, Thurston KG. Radioembolization with 90Yttrium microspheres: a state-of-the-art brachytherapy treatment for primary and secondary liver malignancies. Part 1: Technical and methodologic considerations. *J Vasc Interv Radiol*. 2006;17:1251-1278.
13. Klein S, Staring M, Murphy K, Viergever MA. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196-205.
14. Cancer IN. Common terminology criteria for adverse events (CTCAE) v5.0. Cancer Therapy Evaluation Program; 2018.
15. O JH, Lodge MA, Wahl RL. Practical PERCIST: a simplified guide to PET response criteria in solid tumors 1.0. *Radiology*. 2016;280:576-584.
16. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247.
17. Obuchowski NA. Nonparametric analysis of clustered ROC curve data. *Biometrics*. 1997;53:567-578.
18. Heinze G, Dunkler D. Avoiding infinite estimates of time-dependent effects in small-sample survival studies. *Stat Med*. 2008;30:6455-6469.
19. Riaz A, Awais R, Salem R. Side effects of yttrium-90 radioembolization. *Front Oncol*. 2014;4:198.
20. Kennedy AS, Ball D, Cohen SJ, et al. Multicenter evaluation of the safety and efficacy of radioembolization in patients with unresectable colorectal liver metastases selected as candidates for (90)Y resin microspheres. *J Gastrointest Oncol*. 2015;6:134-142.
21. Hickey R, Lewandowski RJ, Prudhomme T, et al. 90Y Radioembolization of Colorectal Hepatic Metastases Using Glass Microspheres: Safety and Survival Outcomes from a 531-Patient Multicenter Study. *J Nucl Med*. 2016;57:665-671.
22. van den Hoven AF, Rosenbaum CE, Elias SG, et al. Insights into the Dose-Response Relationship of Radioembolization with Resin 90Y-Microspheres: A Prospective Cohort Study in Patients with Colorectal Cancer Liver Metastases. *J Nucl Med*. 2016;57:1014-1019.

- 23.** Flamen P, Vanderlinden B, Delatte P, et al. Multimodality imaging can predict the metabolic response of unresectable colorectal liver metastases to radioembolization therapy with Yttrium-90 labeled resin microspheres. *Phys Med Biol.* 2008;53:6591-6603.
- 24.** Levillain H, Duran Derijckere I, Marin G, et al. (90)Y-PET/CT-based dosimetry after selective internal radiation therapy predicts outcome in patients with liver metastases from colorectal cancer. *EJNMMI Res.* 2018;8:60.
- 25.** Jongen JMJ, Rosenbaum C, Braat M, et al. Anatomic versus metabolic tumor response assessment after radioembolization treatment. *J Vasc Interv Radiol.* 2018;29:244-253 e242.
- 26.** Sager S, Akgun E, Uslu-Besli L, et al. Comparison of PERCIST and RECIST criteria for evaluation of therapy response after yttrium-90 microsphere therapy in patients with hepatocellular carcinoma and those with metastatic colorectal carcinoma. *Nucl Med Commun.* 2019;40:461-468.
- 27.** Bastiaannet R, Lodge MA, de Jong H, Lam M. The Unique Role of Fluorodeoxyglucose-PET in Radioembolization. *PET Clin.* 2019;14:447-457.
- 28.** Willowson KP, Hayes AR, Chan DLH, et al. Clinical and imaging-based prognostic factors in radioembolisation of liver metastases from colorectal cancer: a retrospective exploratory analysis. *EJNMMI Res.* 2017;7:46.
- 29.** Dendy MS, Ludwig JM, Kim HS. Predictors and prognosticators for survival with Yttrium-90 radioembolization therapy for unresectable colorectal cancer liver metastasis. *Oncotarget.* 2017;8:37912-37922.
- 30.** Garin E. A multicentric and randomized study demonstrating the impact of MAA based dosimetry on tumor response in SIRT for HCC. <https://www.eventscribe.com/2019/GEST/fsPopup.asp?Mode=presInfo&PresentationID=521844>. Accessed December 12, 2019.

TABLES AND FIGURES

TABLE 1. Baseline patient and treatment characteristics	
Characteristic	N (%) or median + range
Gender	
Male	25 (62.5)
Female	15 (37.5)
Age (y)	64 (37-84)
WHO performance score	
0	28 (70)
1	11 (27.5)
2	1 (2.5)
Previous therapy*	
Locoregional (liver)	
EBRT	2 (5)
Metastasectomy	5 (12.5)
Radiofrequency ablation	3 (7.5)
Lines of prior systemic treatment	
1	8 (20)
2	20 (50)
3	7 (17.5)
4	5 (12.5)
Extrahepatic disease before treatment	
Lymph node	10 (25)
Lung	10 (25)
No	23 (57.5)
Liver volume (mL)	1987 (1272-3167)
Metabolic tumor volume (mL)	320 (26-1446)
Fractional tumorload	0.15 (0.01-0.49)
Radioembolization treatment	
Whole-liver	39 (97.5)
Lobar (right lobe only)	1 (2.5)
Administered activity (MBq)	6387 (3822-12386)

*No patient received synchronous systemic treatment.

TABLE 2. CTCAE grading of new clinical toxicity per patient during three months after treatment					
Toxicity	CTCAE grade I	CTCAE grade II	CTCAE grade III	CTCAE grade IV	CTCAE grade V
Abdominal pain	16	10	4		
Nausea	15	9	2		
Fatigue	21	10	2		
Anorexia	10	5			
Dyspnea	4	1			
Fever	7	1	1		
Ascites	1		2		
Flu like symptoms	2	1			
Malaise	4	1			
Hepatic failure			1		1*
Weight loss	2				
Chest pain	1	2			
Vomiting	9	5			
Dyspepsia	1	1			
Metal taste	3				
Contrast allergy	1	2			
Hematoma	1				
Diarrhea	1				
Constipation	4				
Upper GI tract bleeding			1		
Limb edema	2				
Dizziness	1				
Chills	2				
Any clinical toxicity	13	19	7		1
Lowered albumin	9	4			
Elevated ALAT	24	1	1		
Elevated AP	4	14	2		
Elevated ASAT	28	2			
Elevated bilirubin	2	1		2	
Elevated GGT	5	15	5		
Any laboratory toxicity	7	23	5	2	

CTCAE scores of new toxicity (highest CTCAE grades per clinical symptom or laboratory value are represented). *Radioembolization-induced liver disease.

TABLE 3. Percentage change in mean absorbed dose (Gy) per response category (95%CI)					
	Progressive disease	Stable disease	Partial response	Complete response*	
<i>Patient-level</i>	n=8	n=17	n=11	n=1	
Unadjusted	<i>reference</i>	53.8 (5.6 - 24.2)	74.6 (18.6 - 57.6)	-	$P_{trend}=0.012$
Adjusted†	<i>reference</i>	62.0(10.4 - 136.0)	77.3 (18.3 - 163.6)	-	$P_{trend}=0.019$
<i>Patient-level‡</i>	n=23	n=6	n=7	n=1	
Unadjusted	<i>reference</i>	29.8 (-15.1 - 98.6)	44.4 (1.4 - 106.0)	-	$P_{trend}=0.041$
Adjusted†	<i>reference</i>	18.7 (-24.3 - 85.4)	38.1 (-5.8 - 101.9)	-	$P_{trend}=0.12$
<i>Tumor-level</i>	n=23	n=49	n=20	n=23	
Unadjusted	<i>reference</i>	31.1 (-3.2 - 78.8)	71.5 (17.1 - 150.4)	95.2 (34.7 - 183.6)	$P_{trend}=0.0003$ 0
Adjusted	<i>reference</i>	35.2 (0.2 - 87.5)	72.2 (16.6 - 151.3)	94.8 (33.9 - 188.4)	$P_{trend}=0.0006$ 8

Interpretation at tumor level: the average dose is 95.23% higher in CR than PD (95%CI 4.69-183.62%).

*as there was only one patient with complete metabolic response, the categories complete response and partial response were taken together at a patient level. † The analyses were adjusted for previous treatment and tumorload or tumor volume (tumor-level analyses). ‡including the development of new lesions, in which case patients were categorized as having progressive disease.

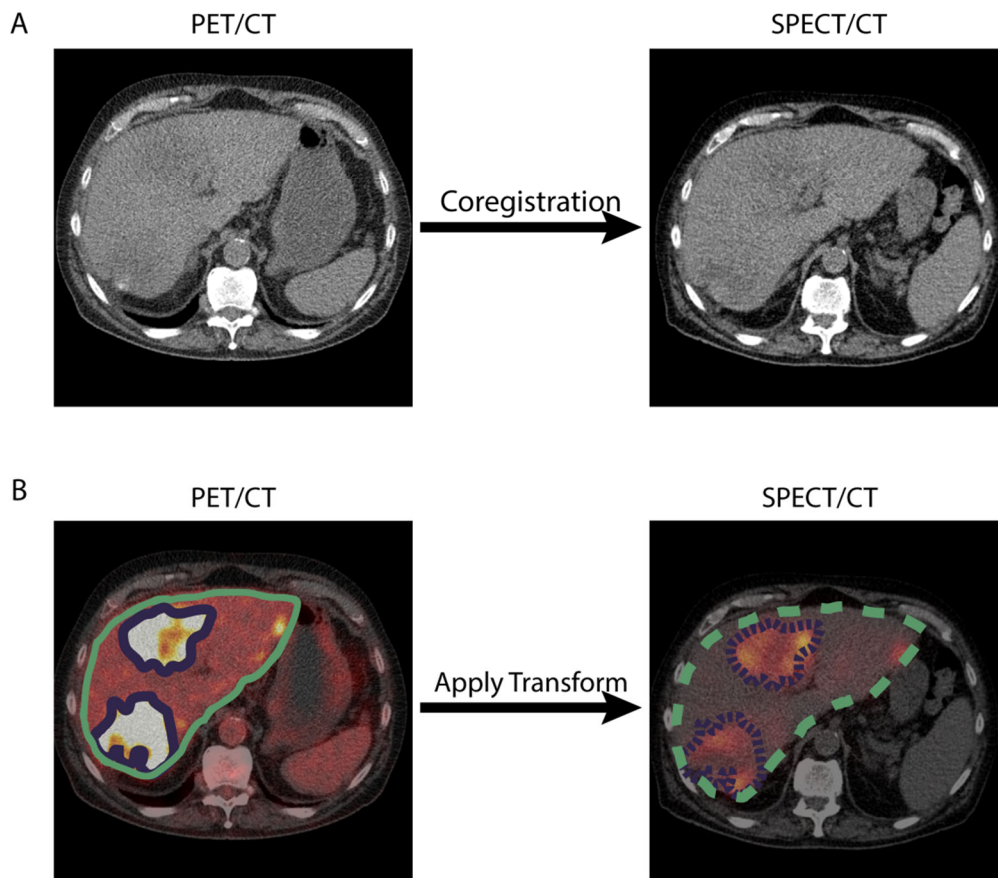


Figure 1. Example of tumor delineation and absorbed-dose estimation. Using the liver contour, the low-dose CT of the PET was matched to the low-dose CT of the SPECT (A). The tumors were automatically defined using a threshold. Liver- and tumor contours were transferred from the PET/CT to the SPECT/CT and absorbed doses were calculated (B).

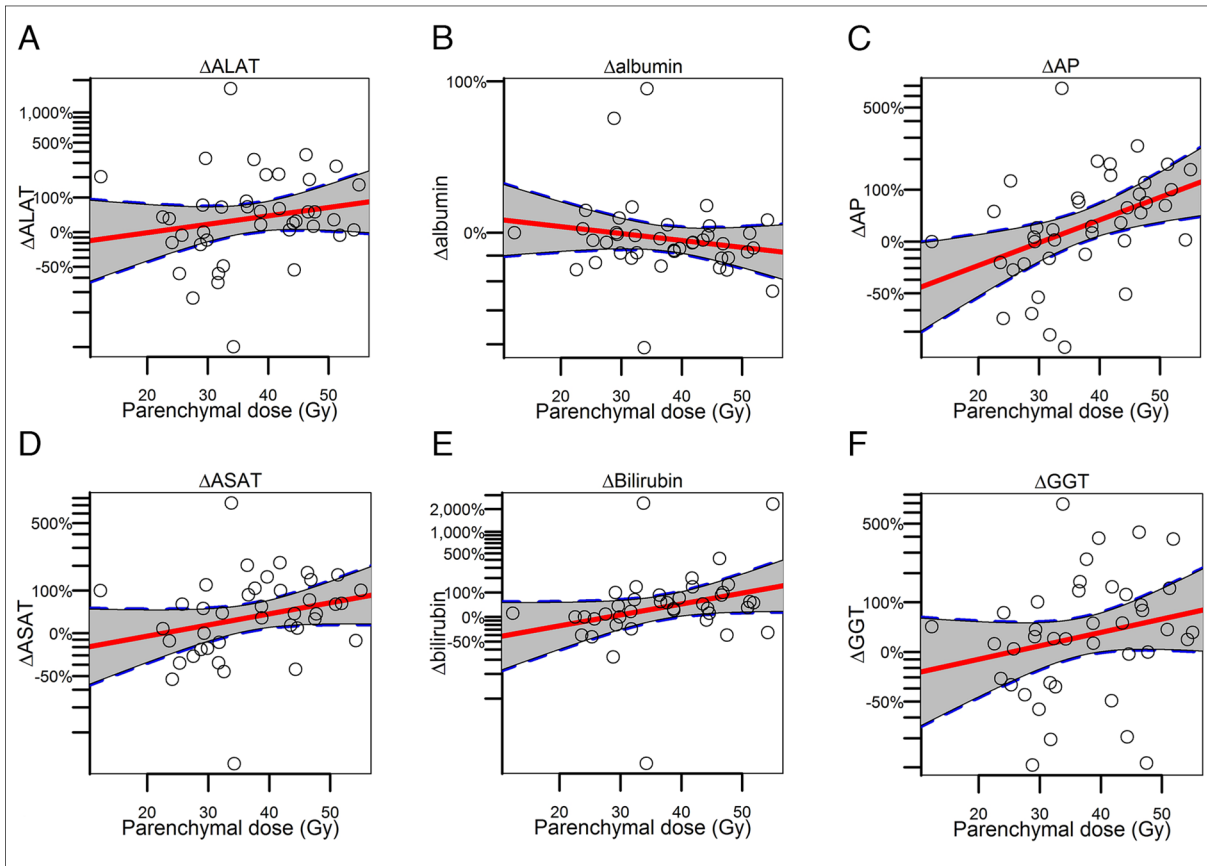


Figure 2a-f. Association between change in laboratory parameters and parenchymal absorbed dose. The red lines are the regression lines, with the 95% CIs indicated as the surrounding grey areas.

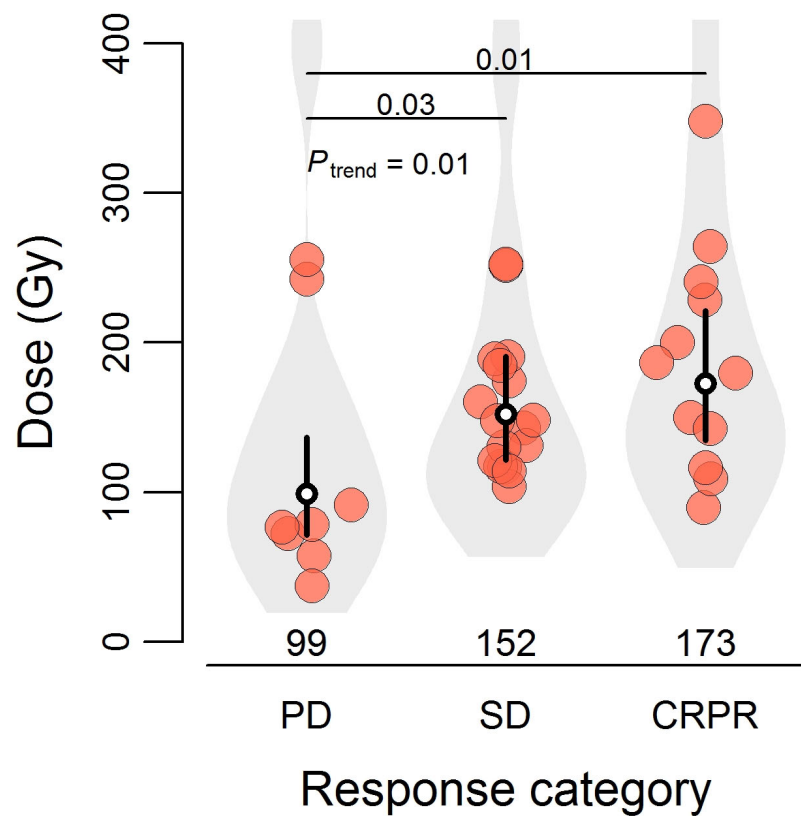


Figure 3. Relationship between mean tumor-absorbed dose per patient and metabolic response to treatment at three-months follow-up. The bullets show the mean tumor-absorbed dose per patient. Black vertical lines are the 95% CIs of the mean doses per response category, with the white dot in the middle indicating the mean tumor-absorbed dose per response category. This figure is based on the unadjusted linear mixed-effects regression model as described in Table 3.

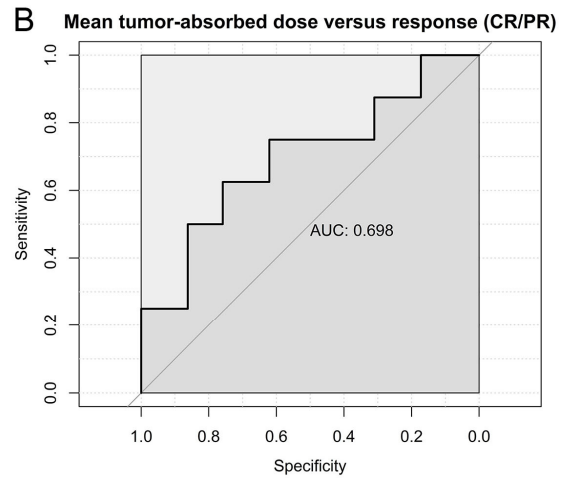
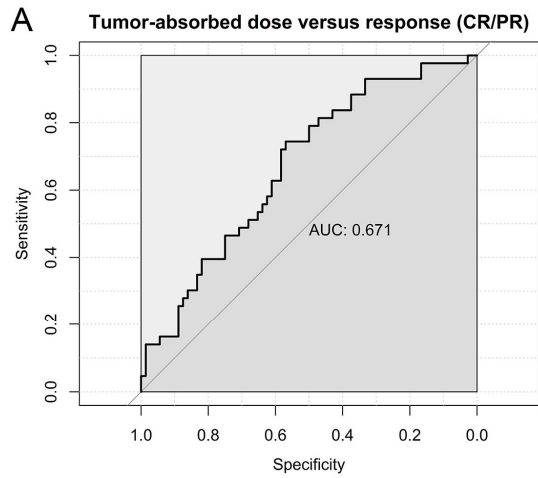
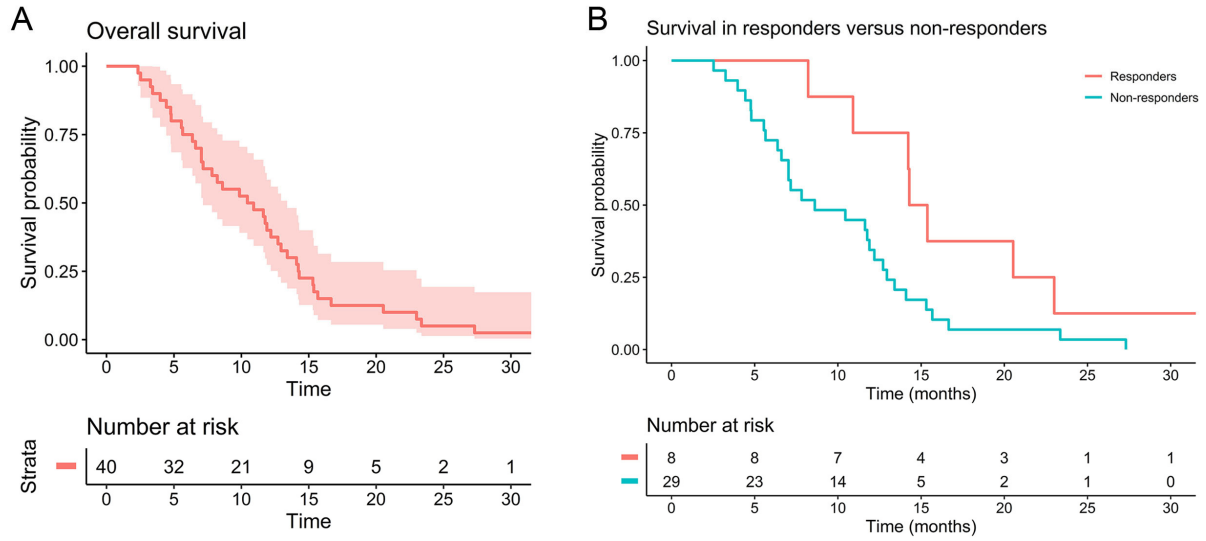


Figure 4a,b*. ROC-curve showing the discriminative value of tumor-absorbed dose for response (a) and ROC-curve showing the ability of mean tumor-absorbed dose per patient to discriminate between patients with CR/PR versus StD/PD (b). *The AUCs are based on a clustered data analysis, however, the ROC-curves are not.



Figures 5a,b. Overall survival curve (a). Survival curves for patients with and without a metabolic response (including the development of new lesions) at three months (b).

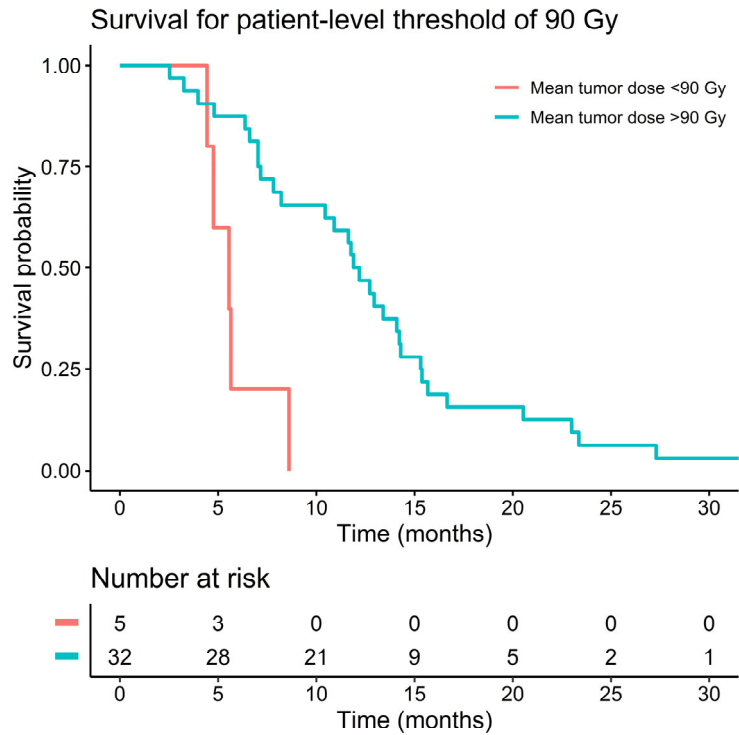


Figure 6. Survival curves for patients with a higher (>90 Gy) or lower (<90 Gy) mean tumor-absorbed dose.

Supplemental information on methods used for response estimation

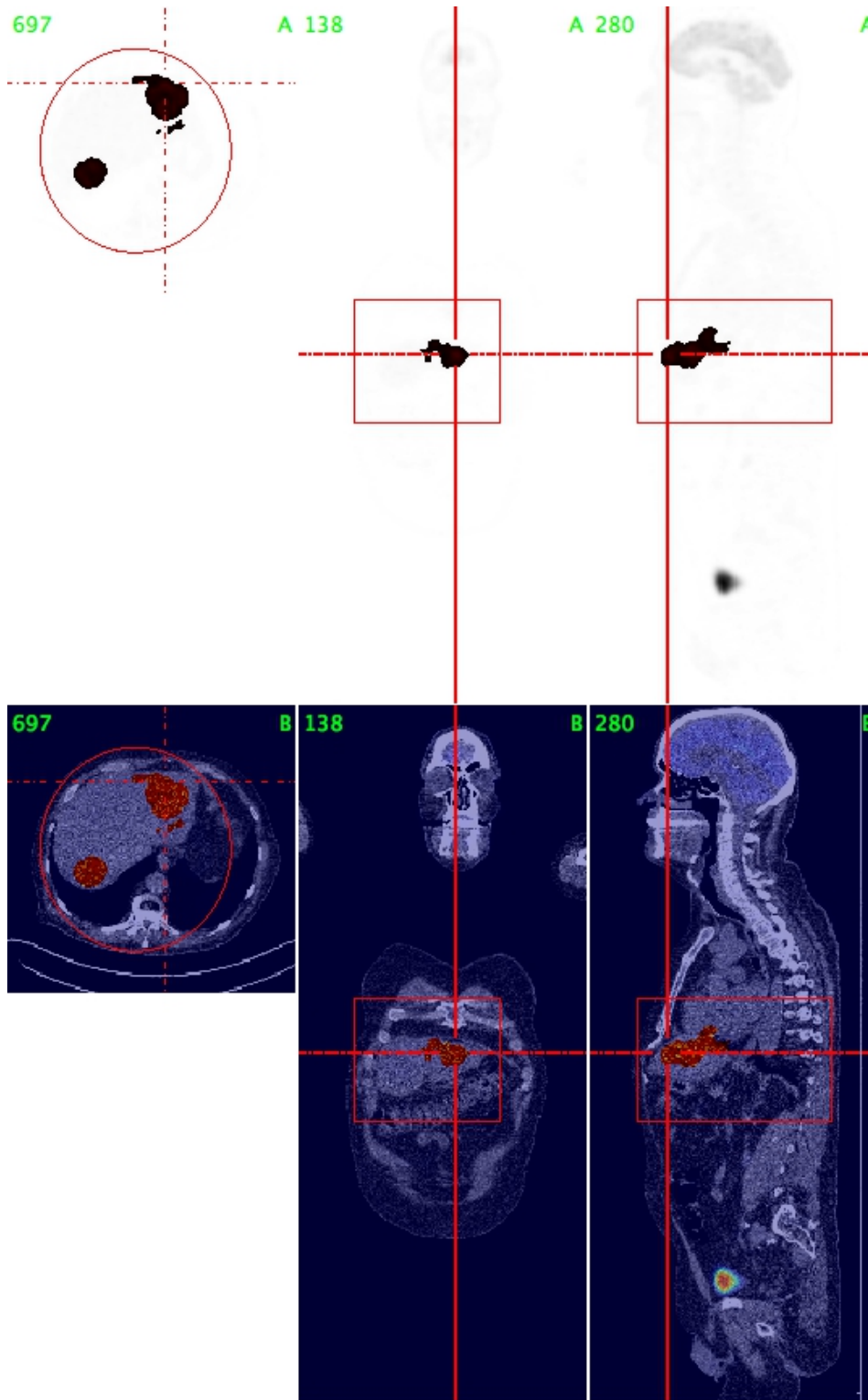
The threshold used for tumor delineation was defined per patient, based on twice the mean aortic blood pool SUV corrected for lean body mass. Using this patient-relative threshold and the volume restriction of 5 mL, tumors were automatically defined. This way, only regions with metabolic activity, significantly exceeding the background activity of the liver, were defined. The threshold used to delineate tumors at follow-up was defined again on the 3-month ^{18}F FDG PET/CT.

A rigid registration (using Elastix software (1)) of the CT scans of the PET and SPECT acquisitions was used to transfer the PET-based tumor- and liver contours to the corresponding ^{166}Ho -SPECT reconstructions. The previously manually contoured livers acted as a mask to focus registration on this region exclusively. All registration results were checked visually (RB). Minor manual adjustments were allowed, but only based on CT and never on nuclear imaging. A 1 cm dilation of the tumor and liver contours was used, to account for breathing movement, errors in registration and resolution differences. The counts in the dilated contours were used for activity calculation, but the volume of the non-dilated VOIs was used for absorbed dose calculation. The quantitative Monte Carlo-based SPECT reconstruction used in this study yields voxels that contain absolute activity (in units of MBq). The absorbed dose (Gy) in each voxel was subsequently calculated using the local deposition model, which posits that -at the resolution of SPECT- all dose is deposited within the voxel of origin. The average doses in parenchymal tissue and tumors was calculated using the transferred delineations, as described previously.

In case of fused lesions at follow-up, a volume-weighted average of the absorbed dose of the different components at baseline was calculated. Also, a weighted average, correcting for tumor volume, was calculated to obtain the mean tumor-absorbed dose per patient. The parenchymal-absorbed dose was determined using the activity in the entire (dilated) liver contour, with the activity in the (dilated) tumor regions subtracted.

References

1. Klein S, Staring M, Murphy K, Viergever MA. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196-205.



Example of tumor delineation using ROVER software (ABX, GmBH, Radeberg, Germany). Above, axial, coronal and transverse images of the PET/CT are shown, which are matched to the accompanying low-dose CTs (below). A mask is drawn around the liver: within the volume of the mask, using a threshold based on the aortic blood pool activity, the tumors were automatically defined.

Statistical analyses – elaborate version

Patient demographics and treatment characteristics were summarized using descriptive analyses. The strength of association between CTCAE toxicity grade and parenchymal-absorbed dose was assessed using linear regression models with CTCAE grade in categories as the dependent continuous variable and parenchymal absorbed dose as the independent continuous variable. For clinical significance, CTCAE grading of any clinical and laboratory toxicity was also dichotomized in the following categories: grade 0/I/II versus grade III/IV/V and analyzed using logistic regression with Firth's correction for small sample bias (2). The association between relative change in laboratory parameters (represented as Δ laboratory parameter) and healthy liver tissue dose was analyzed using simple linear regression models with percentage change as the dependent continuous variable and parenchymal absorbed dose as the independent continuous variable, after log-transformation of the dependent variable to fulfill model assumptions. All toxicity analyses were also adjusted for response to therapy (binary coded as response/non-response), previous treatment (defined as number of prior systemic treatment lines, categorical variable) and tumor load (defined as percentage involvement of the liver by tumors, continuous variable) as possible confounders, which were identified by making directed acyclic graphs.

The relationship between tumor-absorbed dose and response was analyzed using a linear mixed-effects regression model with tumor-absorbed dose as the dependent variable. This type of analysis was chosen to account for correlation of tumors within patients. To fulfill model assumptions, dose was log-transformed. Nested models were compared using Akaike's Information Criterion (AIC). The dose-effect relationship was best explained using a random intercept per patient without random slopes. A geometric mean of the tumor-absorbed dose per patient per response category was estimated because the anti-log of the arithmetic mean of log-transformed values is the geometric mean. A trend test was also done with response as a continuous variable in the model, to test the presence of an ordered

relationship across response categories. By including them as co-variables, analyses were adjusted for the following possible confounders: previous treatment (coded as factor with the following categories: yes/no previous treatment with anti-VEGF medication) and tumor load (continuous). An ROC analysis, according for clustered data, was done to determine the discriminatory power of tumor dose in response estimation (3). The 95% confidence interval of the area under the curve shows the boundaries of the likely discriminative ability of tumor dose for response in this cohort. Using a threshold of 100% sensitivity for response (CR/PR), the threshold for a minimal mean tumor-absorbed dose was determined and used in the survival analyses. The same threshold of 100% sensitivity for response was used to determine the threshold for a minimal tumor-absorbed dose (lesion-level).

The agreement between response according to PERCIST and according to RECIST was analyzed using Cohen's kappa, with disagreements weighted according to their squared distance from perfect agreement.

Overall survival was defined as the interval between treatment and death from any cause. Cox regression models were made using Firth's correction for small sample bias (2). Analyses were adjusted for the following possible confounders: tumor load, parenchymal dose and the presence of extrahepatic disease at baseline. Inspection of Schoenfeld residuals showed that the proportionality of the hazard assumption was not violated. Analyses were performed using R statistical software, version 3.6.2 for Windows. The following R libraries were used: readxl version 1.3.1, dplyr version 0.8.3, data.table version 1.12.8, lme4 version 1.1-21, nlme version 3.1-143, ggplot2 version 3.2.1, gdata version 2.18.0, gmodels version 2.18.1, ggpubr version 0.2.4, Hmisc version 4.3-0, lmerTest version 3.1.0, foreign version 0.8-72, ggfortify version 0.4.8, logistf version 1.23, grid version 3.6.2, car version 3.0-5, pROC version 1.15.3, ggeffects version 0.14.0, splines version 3.6.2, sjmisc version 2.8.3, rel version 1.4.1 and

rcompanion version 2.3.21. We report effect estimates with associated 95%CIs and corresponding two-sided p-values.

References

1. Le Tourneau C, Razak AR, Gan HK, et al. Heterogeneity in the definition of dose-limiting toxicity in phase I cancer clinical trials of molecularly targeted agents: a review of the literature. *Eur J Cancer*. 2011;47:1468-1475.
2. Heinze G, Dunkler D. Avoiding infinite estimates of time-dependent effects in small-sample survival studies. *Stat Med*. 2008;30:6455-6469.
3. Obuchowski NA. Nonparametric Analysis of Clustered ROC Curve Data. *Biometrics*. 1997;53:567-578.

Table S1. Relation between parenchymal dose (Gy) and clinical toxicity based on linear regression analyses with parenchymal dose as the dependent variable.			
		<i>CTCAE grade 0 -V</i>	
Independent variable	Number of patients with toxicity	Mean change in parenchymal dose (Gy) per step increase in CTCAE grade category (95% CI); p-value	
		<i>Unadjusted</i>	<i>Adjusted (for tumor dose, previous treatment and response)</i>
Any variable, highest grade	40	3.43 (-0.23 – 7.10); 0.065	6.56 (1.96 – 11.16); 0.0067
Abdominal pain	30	0.26 (-3.24 – 3.77); 0.88	0.99 (-3.06 – 5.04); 0.62
Nausea	26	3.09 (-0.46 – 6.65); 0.086	2.62 (-0.99 – 6.22); 0.15
Vomiting	14	1.70 (-2.87 – 6.26); 0.46	2.33 (-2.27 – 6.93); 0.31
Fatigue	33	3.78 (-0.25 – 7.81); 0.065	3.66 (-0.60 – 7.91); 0.090
Fever	9	3.61 (-1.32 – 8.55); 0.15	4.85 (-2.44 – 12.13); 0.18
Anorexia	15	1.13 (-3.46 – 5.71); 0.62	0.47 (-4.42 – 5.37); 0.85
Ascites	3	2.92 (-1.86 – 7.70); 0.22	6.86 (0.53 – 13.19); 0.035

The mean change indicates the average increase or decrease in parenchymal dose per step increase in

CTCAE grade toxicity. For example, for any clinical toxicity: a unit increase in toxicity results in an

increase in average parenchymal dose of 3.4 Gy (unadjusted analysis).

Table S2. Relation between parenchymal dose (Gy) and cumulative toxicity over three months, based on logistic regression analyses using Firth's correction with parenchymal dose (per 10 Gy) as the independent variable.		
	<i>CTCAE grade 0-II vs III-V</i>	
Dependent variable	Odds ratio for toxicity parameter per 10 Gy increase in parenchymal dose (95% CI); p-value	
	<i>Unadjusted model</i>	<i>Adjusted model (for tumor dose, previous treatment and response)</i>
Any variable <u>laboratory</u> toxicity, highest grade (n=33 vs n=7)	1.05 (0.46 – 2.48); 0.91	1.08 (0.39 – 3.12); 0.88
Any variable <u>clinical</u> toxicity, highest grade (n=32 vs n=8)	3.61 (1.37 – 13.10); 0.022	9.68 (2.18 – 124.20); 0.019

The odds ratio represents the odds for a CTCAE grade III-V toxicity for every 10 Gy increase in parenchymal absorbed dose.

Table S3a. Relation between parenchymal dose (Gy) and cumulative laboratory toxicity over three months, based on linear regression analyses with parenchymal dose as the dependent variable.

		<i>CTCAE grade 0-V</i>	
Independent variable	Number of patients with toxicity	Mean change (95% CI); p-value	
		<i>Unadjusted</i>	<i>Adjusted for tumor dose, previous treatment and response</i>
Any variable, highest grade	37	2.51 (-1.05 – 6.07); 0.16	3.47 (-0.82 – 7.75); 0.11
GGT	25	2.06 (-0.87 – 4.98); 0.16	2.07 (-1.19 – 5.33); 0.20
AP	20	3.64 (0.65 – 6.63); 0.018	3.34 (-0.09 – 6.78); 0.057
Albumin	13	2.75 (-2.07 – 7.56); 0.26	3.20 (-2.26 – 8.66); 0.24
Bilirubin	5	1.98 (-1.71 – 5.66); 0.29	4.98 (0.02 – 9.93); 0.049
ALAT	26	0.48 (-4.44 – 5.41); 0.84	0.84 (-5.58 – 7.27); 0.79
ASAT	30	3.30 (-2.22 – 8.81); 0.23	4.49 (-1.98 – 10.96); 0.17

The mean change indicates the average increase or decrease in parenchymal dose per unit increase in CTCAE grade toxicity. For example, for GGT: a unit increase in toxicity results in an increase in average parenchymal dose of 2.06 Gy (unadjusted analysis).

Table S3b. Relation between parenchymal dose (Gy) and change in laboratory parameters over three months, based on linear regression analyses with parenchymal dose (per 10 Gy) as the independent variable

Dependent variable	Mean percent change (95% CI); p-value	
	<i>Unadjusted</i>	<i>Adjusted for tumor dose, previous treatment and response</i>
ΔGGT	19.6% (-9.1 – 57.3); 0.17	34.1.2% (0.5 – 79.7), 0.043
ΔAP	34.5% (8.5 – 66.7); 0.0063	33.3% (6.5 – 66.6), 0.011
ΔAlbumin	-4.1% (-9.7 – 1.8); 0.28	-3.1% (-9.1 – 1.8), 0.71
ΔASAT	18.9% (-4.9 – 48.5); 0.11	14.5% (-9.1 – 44.1), 0.21
ΔALAT	17.3% (-12.7 – 57.7); 0.25	13.3% (-16.5 – 53.6), 0.37
ΔBilirubin	35.1% (-4.5 – 90.7); 0.077	46.6% (3.4 – 107.7), 0.029

The mean change indicates the increase or decrease in average toxicity per 10 Gy increase in

parenchymal dose. For example, for GGT: for every 10 Gy increase in parenchymal dose, there is a 19.6%

increase in GGT (unadjusted analysis).