ORIGINAL RESEARCH

Open Access

Measuring total liver function on sulfur colloid SPECT/CT for improved risk stratification and outcome prediction of hepatocellular carcinoma patients

Stephen R. Bowen[1,2*], Tobias R. Chapman[1], Joshua Borgman[3], Robert S. Miyaoka[2], Paul E. Kinahan[2], Iris W. Liou[4], George A. Sandison[1], Hubert J. Vesselle[2], Matthew J. Nyflot[1] and Smith Apisarnthanarax[1]

Abstract

Background: Assessment of liver function is critical in hepatocellular carcinoma (HCC) patient management. We evaluated parameters of [^{99m}Tc] sulfur colloid (SC) SPECT/CT liver uptake for association with clinical measures of liver function and outcome in HCC patients.

Methods: Thirty patients with HCC and variable Child-Turcotte-Pugh scores (CTP A5-C10) underwent [^{99m}Tc]SC SPECT/CT scans for radiotherapy planning. Gross tumor volume (GTV), anatomic liver volume (ALV), and spleen were contoured on CT. SC SPECT image parameters include threshold-based functional liver volumes (FLV) relative to ALV, mean liver-to-spleen uptake ratio (L/S$_{mean}$), and total liver function (TLF) ratio derived from the product of FLV and L/S$_{mean}$. Optimal SC uptake thresholds were determined by ROC analysis for maximizing CTP classification accuracy. Image metrics were tested for rank correlation to composite scores and clinical liver function parameters. Image parameters of liver function were tested for association to overall survival with Cox proportional hazard regression.

Results: Optimized thresholds on SC SPECT were 58 % of maximum uptake for FLV, 38 % for L/S$_{mean}$, and 58 % for TLF. TLF produced the highest CTP classification accuracy (AUC = 0.93) at threshold of 0.35 (sensitivity = 0.88, specificity = 0.86). Higher TLF was associated with lower CTP score: TLF$_A$ = 0.6 (0.4–0.8) versus TLF$_B$ = 0.2 (0.1–0.3), $p < 10^{-4}$. TLF was rank correlated to albumin and bilirubin ($|R| > 0.63$). Only TLF >0.30 was independently associated with overall survival when adjusting for CTP class (HR = 0.12, 95 % CI = 0.02–0.58, $p = 0.008$).

Conclusions: SC SPECT/CT liver uptake correlated with differential liver function. TLF was associated with improved overall survival and may aid in personalized oncologic management of HCC patients.

Keywords: SPECT/CT, Sulfur colloid, Child-Pugh, ROC, HCC, Liver function

* Correspondence: srbowen@uw.edu
[1]Department of Radiation Oncology, University of Washington School of Medicine, 1959 NE Pacific St, Box 356043, Seattle, WA 98195, USA
[2]Department of Radiology, University of Washington School of Medicine, 1959 NE Pacific St, Box 356043, Seattle, WA 98195, USA
Full list of author information is available at the end of the article

Bowen *et al. EJNMMI Research* (2016) 6:57

Page 2 of 10

Background

The accurate assessment of liver function is of critical importance in the management of patients with hepatocellular carcinoma (HCC) [1]. The majority of patients with HCC have associated chronic liver disease (viral hepatitis, alcohol injury, or steatosis) [2] and resultant cirrhosis that places patients at risk of morbidity and mortality from liver-directed therapies [3–7]. Traditionally, severity of chronic liver disease has been graded using the Child-Turcotte-Pugh (CTP) classification system [8, 9], originally designed over 50 years ago to estimate mortality in patients with cirrhosis and portal hypertension prior to undergoing surgical treatment of portal hypertension. The CTP system [9] assigns points to 3 continuous objective variables (serum bilirubin, albumin, and prothrombin time) and 2 subjective variables (ascites and hepatic encephalopathy). These points are added together to generate a composite score to assign patients into one of the three classes with increasing liver dysfunction: A, B, or C.

While the CTP classification system has been shown to be an independent prognostic factor, it has inherent limitations that can under- or overestimate the degree of liver dysfunction, such as empiric cutoff values, equal weighting of variables, and challenges with the subjective interpretation of ascites and encephalopathy [10]. More recently, other prognostic models of cirrhosis have been designed, including the model for end-stage liver disease (MELD) score [11, 12] and its derivatives [10], that are utilized for prioritization of liver transplantation. These systems, however, were not specifically designed for patients with HCC, unlike the recently published albumin-bilirubin (ALBI) grade [13]. Staging systems that incorporate both tumor and liver function parameters for HCC patients include the Barcelona Clinic Liver Cancer (BCLC), Cancer of the Liver Italian Program (CLIP), and Japan Integrated Scoring system (JIS) [14]. However, all of these classification systems and models (including CTP and MELD) still rely on traditional measures of liver function, only provide a *global* assessment of liver function, and/or do not take into account spatial variations in liver function [15]. Knowledge of this patient-specific liver function heterogeneity may be valuable in the local therapeutic management of HCC patients [16, 17].

Quantitative imaging of liver function holds promise as an alternative to traditional measures of liver function, if properly validated. Diagnostic imaging of liver function includes positron emission tomography (PET) with [18F]fluorodeoxygalactose [18, 19] and dynamic contrast-enhanced magnetic resonance imaging [20] with gadoxetic acid [21] or gadoxetate disodium [22, 23]. Single photon emission computed tomography (SPECT) of several radiotracers have been investigated as surrogates for liver function [18, 19, 21, 24], most notably [99mTc] hepatobiliary iminodiacetic acid [25], [99mTc]

galactosyl-human serum albumin [24], and [99mTc] sulfur colloid (SC). The latter is a well-established FDA-approved diagnostic tracer that images the reticuloendothelial (Kupffer) cells of the liver and has been shown to correlate with chronic liver disease severity and function [26–29]. However, the reliability of quantitative molecular imaging parameters as surrogates of global and regional liver function and their relationship to conventional liver function parameters remains unclear. Historical methods of qualitatively or semi-quantitatively assessing SC distribution with 2-D planar scintigraphy provided limited contrast and spatial resolution that relied on expert observer interpretation. Subjective visual scoring of planar scintigraphy was ineffective in detecting variations in liver function [30], which highlights the critical need for more sensitive tests with standardized quantitative parameters.

The evolution of combined SPECT/CT technology has improved regional contrast between functional and non-functional liver from tomographic image acquisition. It has also increased the accuracy in local SC uptake estimation from collimator-detector response modeling, scatter correction, and CT-based attenuation correction, resulting in distinct uptake patterns between patients with differential liver function (Fig. 1). In this era of modern SC SPECT/CT imaging, however, it is not known which imaging metrics are most closely associated with clinical liver function and patient outcomes. The purpose of this study, therefore, is to test measures of [99mTc]SC SPECT/CT liver uptake for statistical association with traditional clinical measures of liver function and survival in HCC patients. Automatic SPECT region-of-interest definition was first optimized to differentiate between CTP classes, and then, SC SPECT/CT metrics were independently correlated to clinical outcome as part of univariate and multivariate analyses. In the absence of absolute quantitative SPECT characterization, relative image intensity ratios were constructed to quantify both magnitude and volumetric extent of liver function.

Methods

Patient characteristics

After obtaining Institutional Review Board approval, 30 HCC patients (19 males, 11 females) with a median age of 67 years (range 43–84) were included in the study, from 2013 to 2015. Patients were selected for external beam radiotherapy after multi-disciplinary assessment for suitability and ineligibility for other liver-directed therapies. Twenty-eight patients had unifocal HCC lesions, and gross tumor volumes (GTV) ranged from 1–1573 cm^3 (median 15 cm^3). The majority of patients (71 %) had received prior liver-directed therapy (LDT) for HCC and presented with persistent disease, local recurrence, or new HCC lesions, as defined by either contrast enhanced CT or MRI. Prior LDTs included radiofrequency ablation
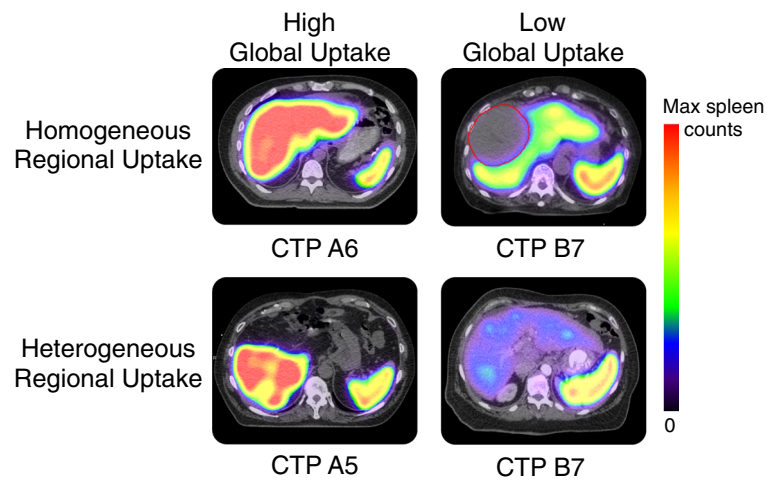
Fig. 1 Sulfur colloid SPECT/CT uptake patterns in hepatocellular carcinoma patients with different Child-Turcotte-Pugh score. Four patients are shown with varying degrees of global liver uptake relative to spleen and regional uptake variation, which may be a consequence of their clinical liver function status. Imaging can parameterize these uptake patterns and provide non-invasive longitudinal biomarkers of liver function. Color scales have been normalized to spleen uptake to allow inter-patient comparison of liver uptake. Gross tumor is outlined in *red* where indicated

($n = 7$), transarterial chemoembolization ($n = 15$), radio-embolization ($n = 3$), or bland embolization ($n = 2$), and exactly half of all patients received more than one LDT. The median time from the last LDT was 3 months (range 1–35 months). All patients had underlying clinically diagnosed cirrhosis, predominantly with either well-compensated or mildly decompensated liver function: CTP A ($n = 16$), CTP B ($n = 12$), and CTP C ($n = 2$), respectively (range A5–C10). Cirrhosis was related to either hepatitis C ($n = 17$), alcohol intake ($n = 8$), non-alcoholic fatty liver disease ($n = 5$), hepatitis B ($n = 6$), or a combination of these factors ($n = 7$). A minority of patients presented with hepatic encephalopathy ($n = 6$) or ascites ($n = 5$).

CTP class was calculated using history and physical examination findings, imaging results, and laboratory data [8, 9]. Patients presented with a range of baseline albumin (2.1–4.0 g/dL), bilirubin (0.3–2.4 mg/dL), and international normalized ratios (INR) (0.9–5.8). Portal hypertension was defined as clinical or radiographic evidence of portal venous system varices/splenomegaly or ongoing medical management for varices. Splenomegaly was defined when the greatest oblique superior to inferior dimension on sagittal CT scan was >15 cm.

SPECT/CT image acquisition, reconstruction, and registration

All patients with a history of cirrhosis and/or prior liver-directed therapy treated in our department from 2013 to 2015 underwent [99mTc]SC SPECT/CT scans prior to definitive radiotherapy as part of their clinical care and were reproducibly immobilized in treatment position. SPECT/CT images were acquired on a Precedence™

(Philips Healthcare, Andover, MA) scanner comprising a dual head gamma camera and 16 slice CT scanner. Following the injection of 7 mCi (259 MBq) [99mTc]SC, SPECT scans were acquired 15 min post-injection over a fixed time-averaged frame (64 views, 20 s/view, 180° arc). Emission images were corrected for scatter, collimator-detector response, and attenuation using a tidal breathing end-exhale position CT image. Reconstructions were performed with the Astonish™ (Philips Healthcare, Andover, MA) ordered subset expectation-maximization (OSEM) iterative algorithm over 2 iterations and 16 subsets that included a 10-mm Hanning filter and 4.7 mm isotropic voxels.

Liver anatomy from the end-exhale attenuation correction CT acquired with each SPECT scan was registered to the reference liver anatomy from the end-exhale respiratory phase of a radiotherapy planning CT acquired under free-breathing ($n = 13$), abdominal compression ($n = 10$), or active breathing control (ABC™, Elekta Inc., Stockholm, Sweden) breath-hold conditions ($n = 7$). Rigid registration between the planning CT and SPECT/CT was performed in MIM 6.4™ (MIM Software Inc., Cleveland, OH) using built-in mutual information methods. The resulting spatial transformations estimated from CT-to-CT registration were applied to the respective SPECT images. Deformable registration techniques between the radiotherapy planning CT and attenuation correction CT were initially evaluated but did not provide sufficiently improved liver registration accuracy, particularly in the context of end-exhale CT scans and low spatial resolution SPECT, to warrant their implementation for this study.

Bowen et al. EJNMMI Research (2016) 6:57

Page 4 of 10

SC SPECT metrics

Regions of interest such as the GTV, anatomic liver volume (ALV, defined as the liver minus GTV), and spleen were defined on the planning CT, with the assistance of contrast-enhanced CT scans as needed. On the spatially aligned SC SPECT image, two relative metrics of functional liver volume (FLV) and liver function magnitude were calculated: the ratio of an uptake threshold-defined functional volume to ALV and the ratio of mean liver counts within the ALV to mean spleen counts (L/S_{mean}). These parameters did not rely on absolute quantitative estimates of tracer activity concentration (MBq/mL). Ratios of uptake in the liver versus spleen have been found useful in prior reports [28] and assume that the majority of SC uptake is in the liver and spleen, even in the presence of shunting to the bone marrow in cirrhotic patients. The FLV ratio represents a *volumetric measure* of liver function, while L/S_{mean} represents a *global magnitude* of liver function. A macroparameter calculated from product of the FLV and L/S_{mean} ratios, termed the total liver function (TLF) ratio, is a measure of *integral* liver function. These SC SPECT parameters were calculated by the following equations:

$$FLV = fV/ALV \tag{1}$$

$$L/S_{mean} = I_{mean}^{liver}/I_{mean}^{spleen} \tag{2}$$

$$TLF = FLV \times L/S_{mean} \tag{3}$$

The functional volume (fV) was defined by an image intensity threshold as a percentage of maximum SC SPECT image intensity within the ALV. The mean image intensity (I_{mean}) was also defined by a fixed threshold as a percentage of maximum SC SPECT image intensity within the ALV or the spleen. A range of image intensity thresholds were tested, as described in the "Statistical analysis" section. The ensemble of SC SPECT uptake measures (FLV, L/S_{mean}, TLF) is analogous to the quantitative FDG PET parameters of metabolic tumor volume (MTV), mean tumor uptake (SUV_{mean}) or mean tumor-to-blood ratio (T/B_{mean}), and total lesion glycolysis (TLG).

Statistical analysis

Optimal image thresholds for SC SPECT parameter association with CTP classification (A vs. B/C class) were interrogated by receiver-operator characteristic (ROC) analysis. A range of image intensity thresholds to define functional liver regions of interest, calculated as a percentage of maximum SC SPECT liver uptake from 20 to 70 % in increments of 2 %, were included in ROC analysis of FLV ratio, L/S_{mean} ratio, and TLF ratio to maximize the ROC area under the curve (AUC). The maximum AUC image thresholds defined functional liver regions of interest and the SC SPECT metrics for

all remaining statistical analysis. Significant differences in Wilcoxon rank sum of ROC-optimized SC SPECT parameters were tested between CTP classes. Spearman rank correlations between SC SPECT parameters and clinical liver function parameters were tabulated for the following categories: (1) composite liver function scoring systems (CTP score, ALBI grade), (2) individual quantitative liver function components (albumin, bilirubin, INR), and (3) individual qualitative clinical liver function components (encephalopathy, portal hypertension, ascites, and splenomegaly). Qualitative components were treated as binary categorical variables (yes/no).

Imaging and clinical parameters of liver function were tested for association with overall survival. Continuous variables were first dichotomized above and below a threshold value that maximized accuracy of classification under ROC with balanced sensitivity and specificity. Univariate Cox proportional hazard regression was then performed on all continuous and categorical variables. Statistically significant variables that were not highly correlated were selected for multivariate Cox proportional hazard regression to maximize model log-likelihood. For this preliminary investigation in a limited patient cohort, one imaging variable and one clinical variable were selected. The simple multivariate model served to adjust for the effect size of established clinical liver function scoring systems (CTP, ALBI) with known association with overall survival. All statistical calculations were performed in OriginPro 9.1™ (OriginLab Corporation, Northampton, MA), and all hypothesis testing reported two-tailed *p* values.

Results

Figure 2 displays the ROC analysis for determination of optimal SC SPECT image uptake thresholds and image parameter for CTP class differentiation (A vs. B/C). As image thresholds of percentage maximum SC SPECT uptake were increased (Fig. 2a), the FLV ratio AUC increased monotonically before reaching a global maximum at 58 %. The L/S_{mean} ratio remained relatively constant in its classification accuracy of CTP outside of a global maximum at 38 %. The TLF ratio as a product of FLV and L/S_{mean} ratios maximized classification accuracy at a threshold of 58 %. The optimal thresholds for each parameter (58 % max SC SPECT uptake for FLV and TLF ratios, 38 % max SC SPECT uptake for L/S_{mean}) defined a set of ROC curves in Fig. 2b, which show equal increase in AUC from 0.87 to 0.93. Among these three image parameters, the TLF ratio at the optimal 58 % threshold was best at predicting CTP class with an AUC of 0.93. For an operating point that equalizes specificity and sensitivity, the optimal cutoff values were an FLV ratio of 0.40 (sensitivity = 0.75, specificity = 0.71), L/S_{mean} ratio of 0.88 (sensitivity = 0.88, specificity = 0.86), and TLF ratio of 0.35 (sensitivity = 0.88, specificity = 0.86).

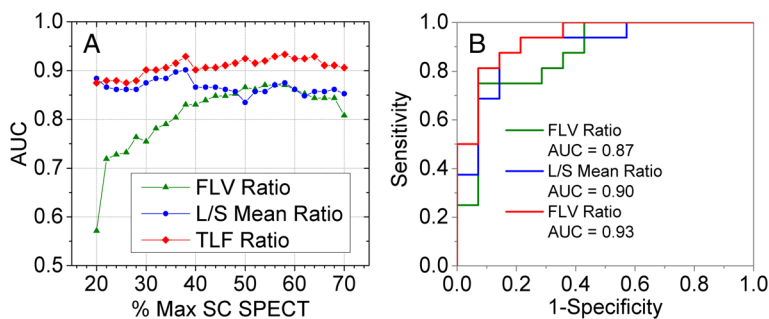Bowen et al. EJNMMI Research (2016) 6:57

Page 5 of 10

Fig. 2 Optimal sulfur colloid SPECT/CT uptake threshold region-of-interest definition for liver imaging parameter prediction of clinical liver function defined by Childs-Turcotte-Pugh A versus B/C class. **a** The receiver operator characteristic (ROC) area under the curve (AUC) is plotted versus the percentage maximum sulfur colloid SPECT uptake threshold used to define three imaging metrics: functional liver volume (FLV) ratio of total liver volume, mean liver-to-spleen (L/S) uptake ratio, and their product as the total liver function (TLF) ratio. **b** ROC curves at the optimal SC SPECT threshold for FLV ratio, L/S mean ratio, and TLF ratios. The product of liver function magnitude and volume yields the highest accuracy for CTP classification, with a TLF cutoff value of 0.35 having a sensitivity of 0.88 and specificity of 0.86

Table 1 lists the difference in the optimized image threshold parameters between CTP classes. The FLV ratio displayed the lowest sensitivity in median difference between CTP A and B/C patients, whereas the L/S_{mean} ratio showed the highest sensitivity in median difference but also the highest intra-class quartile range. The TLF ratio as their product demonstrated both high sensitivity in median difference but low variability in intra-class quartile range, reflected in the lowest probability of a type I error.

From the optimized image thresholds, the FLV ratio, L/S_{mean} ratio, and TLF ratio were rank correlated to clinical liver function parameters in Fig. 3. With decreasing clinical liver function (increasing CTP score), all three ratios also decreased. FLV ratio had low variability for patients with CTP A5-B7 but was highly variable for patients with CTP B8-C10. L/S_{mean} ratio had higher variability but decreased monotonically with increasing score from CTP A5-B9. TLF ratio had the most significantly negative rank correlation with CTP score ($R_{TLF} = -0.76$) compared to the other imaging parameters ($R_{FLV} = -0.66$, $R_{L/Smean} = -0.70$). Note that while the interquartile ranges in TLF ratio at each CTP score were narrow, the overall range of TLF ratios was especially high at the interface between CTP A6 and B7. This discordance between SC SPECT uptake measures and CTP score for certain patients may be addressed with improved liver function classification models.

Figure 4 illustrates absolute Spearman rank correlation coefficients as a function of several clinical liver function
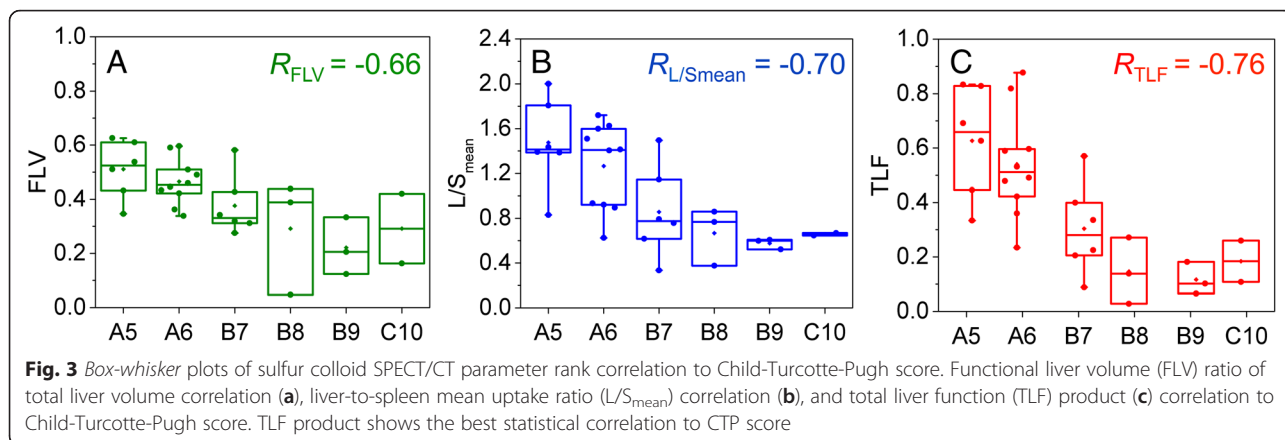
parameters, grouped by composite scoring systems (red), individual quantitative parameters (yellow), and individual qualitative parameters (gray). FLV ratio was highly correlated to both CTP and ALBI composite scoring systems, but only somewhat correlated to albumin, bilirubin, and encephalopathy. L/S_{mean} ratio was strongly correlated to individual quantitative components (albumin, bilirubin, INR) and portal hypertension but was less correlated to the ALBI grade. TLF ratio had the highest absolute rank correlation to the composite scoring systems ($|R_{TLF}| = 0.64–0.76$) and all components ($|R_{TLF}| > 0.4$) with the exception of splenomegaly.

The survival analysis was conducted on the patient cohort at a median follow-up post-treatment of 17 months (range 1–29 months), with 12 deaths considered as events and survivors right-censored at the last follow-up. Among survivors, the minimum follow-up interval was 9 months and median follow-up interval was 22 months. Patients who died presented with baseline median TLF = 0.22 (interquartile range = 0.09–0.48). ROC analysis for classification of overall survival yielded a cutoff value of TLF = 0.30 (sensitivity = 0.88, specificity = 0.75). Table 2 lists the variables with statistically significant association to overall survival (OS), including hazard ratios (HR, 95 % confidence intervals) and Cox proportional hazard p values. TLF as a dichotomous variable (TLF > 0.30) was most strongly associated to OS (HR = 0.08, $p < 0.001$). For comparison, TLF was dichotomized by a simple median cutoff (TLF > 0.38), under which statistical

Table 1 Liver function imaging parameter comparison between Child-Turcotte-Pugh A versus B/C patients

	CTP A ($n = 16$)	CTP B/C ($n = 14$)	Classification cutoff	Rank sum p
Functional liver volume ratio (of liver-GTV)	0.5 (0.4–0.6)	0.3 (0.2–0.4)	0.40	6×10^{-4}
Mean liver-to-spleen uptake ratio	1.4 (0.9–1.6)	0.7 (0.6–0.8)	0.88	2×10^{-4}
Total liver function (FLV * L/S mean)	0.6 (0.4–0.8)	0.2 (0.1–0.3)	0.35	6×10^{-5}

Median (interquartile range) values are reported for functional liver volume (FLV), mean liver-to-spleen uptake ratio, and total liver function for Child-Turcotte-Pugh class A and B/C groups. Statistically significant differences between CTP classes were estimated by Wilcoxon rank-sum testing ($a = 0.05$)

Bowen *et al. EJNMMI Research* (2016) 6:57

Page 6 of 10

Fig. 3 *Box-whisker* plots of sulfur colloid SPECT/CT parameter rank correlation to Child-Turcotte-Pugh score. Functional liver volume (FLV) ratio of total liver volume correlation (**a**), liver-to-spleen mean uptake ratio (L/S$_{mean}$) correlation (**b**), and total liver function (TLF) product (**c**) correlation to Child-Turcotte-Pugh score. TLF product shows the best statistical correlation to CTP score

significance of univariate Cox regression was maintained (HR = 0.20, *p* = 0.02).

Multivariate Cox regression was performed by selecting a single imaging parameter (TLF > 0.30) with the highest level of significance and greatest hazard ratio deviation from unity. This prevented inclusion of imaging variables with high cross-correlation. The single imaging variable was paired with a single clinical variable with significant and substantial association to overall survival, either CTP class or ALBI grade, which produced two equivalent multivariate models, listed in Table 3, with

log-likelihood *p* < 0.001. After adjusting for clinical scoring systems of liver function (CTP class, ALBI grade), only TLF ratio above/below a threshold of 0.30 was independently associated with overall survival (model A: HR = 0.12, 95 % CI = 0.02–0.58, *p* = 0.008; model B: HR = 0.10, 95 % CI = 0.02–0.42, *p* = 0.002).

Discussion
Parameters of sulfur colloid SPECT/CT have potential to measure liver function and further refine risk stratification models for HCC patients. This investigation
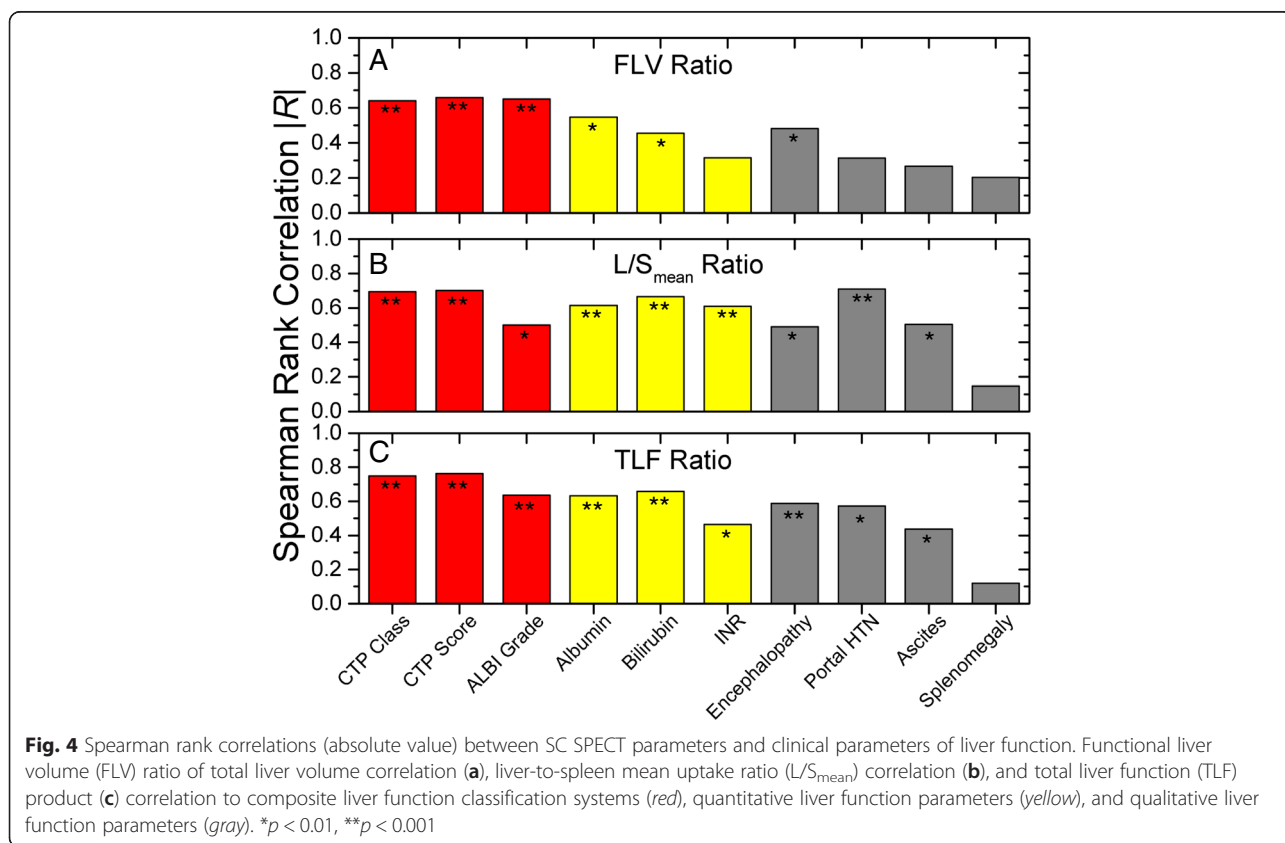

Fig. 4 Spearman rank correlations (absolute value) between SC SPECT parameters and clinical parameters of liver function. Functional liver volume (FLV) ratio of total liver volume correlation (**a**), liver-to-spleen mean uptake ratio (L/S$_{mean}$) correlation (**b**), and total liver function (TLF) product (**c**) correlation to composite liver function classification systems (*red*), quantitative liver function parameters (*yellow*), and qualitative liver function parameters (*gray*). *p < 0.01, **p < 0.001

Bowen *et al. EJNMMI Research* (2016) 6:57

Page 7 of 10

Table 2 Predictors of overall survival under univariate Cox proportional hazard regression

	HR	95 % CI	Cox *p*
FLV (continuous)	0.005	0–0.64	0.02
L/S$_{mean}$ (>0.81)	0.19	0.05–0.42	0.008
TLF (continuous)	0.03	0.001–0.60	0.02
TLF (>0.30)	0.08	0.02–0.34	<0.001
CTP class	6.21	1.63–23.59	0.007
CTP score	1.73	1.18–2.55	0.005
ALBI grade	4.24	1.28–14.04	0.02

Hazard ratios (HR) and 95 % confidence intervals (CI) are reported along with Cox test of significance. Continuous variables were dichotomized by a threshold value that maximized accuracy under ROC classification with balanced sensitivity/specificity

sought to establish SC SPECT/CT image parameters that best correlate to clinical liver function by optimizing the automatic thresholding of SC SPECT and correlating relative ratios to composite and individual clinical parameters of liver function. The total liver function (TLF) as a product of liver function magnitude (L/S$_{mean}$) and volume (FLV) was found to be the best imaging classifier of clinical liver function as a result of having the highest accuracy and lowest probability of a type I error in differentiating CTP class. TLF was most correlated with quantitative variables of liver function (albumin, bilirubin), including independent correlation to the recent ALBI composite scoring system. Correlation to qualitative clinical parameters such as portal hypertension was driven by the liver-to-spleen ratio rather than functional liver volume.

Survival analysis revealed that TLF was associated with overall survival after adjusting for CTP class or ALBI grade. Following independent validation of this preliminary finding, TLF may be considered as a candidate imaging biomarker that complements traditional liver function classification systems, particularly in cases of discordant findings. These results support further investigation on the incorporation of complementary volumetric functional imaging into HCC risk stratification models, which

Table 3 Independent predictors of overall survival under multivariate Cox proportional hazard regression

Model A	HR	95 % CI	Cox *p*
TLF (>0.30)	0.12	0.02–0.58	0.008
CTP class	2.80	0.58–13.56	0.20
Model B	HR	95 % CI	Cox *p*
TLF (>0.30)	0.10	0.02–0.42	0.002
ALBI grade	2.04	0.52–8.00	0.30

Statistically significant univariate predictors (one imaging variable, one clinical variable) were paired. Hazard ratios (HR) and 95 % confidence intervals (CI) are reported along with Cox test of significance. Two equivalent multivariate regression models were generated (model A: comparison to CTP class; model B: comparison to ALBI grade; log likelihood *p* < 0.001)

would enable quantification of global liver function and characterization of spatial heterogeneity in liver function. The latter is indicated in settings where prior liver-directed therapies may have compromised regions of liver function. Such liver function imaging parameters may become essential for targeted liver therapy planning of external beam radiotherapy, radioembolization, chemoembolization, and surgical resection.

TLF is analogous in construction to total lesion glycolysis from cancer imaging with FDG PET/CT, which in several disease sites has been shown to better predict clinical outcome and survival than imaging parameters of uptake magnitude (SUV$_{mean}$) and metabolic tumor volume (MTV) alone [31–37]. Past investigations have characterized liver function with SC image parameters of magnitude (liver-to-spleen ratio) [28, 29] or size/density (perfused hepatic mass) [26, 27], but the impact of liver dysfunction on oncology management of HCC patients and post-therapy toxicity may be best parameterized by integrating function with the liver volume. When assessing our patient cohort with L/S$_{mean}$ and FLV alone, eight patients (27 %) had image parameters that resulted in different CTP classification (i.e., CTP A using L/S$_{mean}$ and CTP B using FLV and vice versa), which reinforces a complex interplay between volume and magnitude of liver function. The TLF represents a functional liver mass that takes into account both the density and volumetric extent of liver function in patients. This concept of a functional image-defined hepatic mass is intuitively in agreement with surgical literature on the future liver remnant whose functional changes are closely tied to post-surgical outcomes [38–40]. TLF, when strategically paired with quantitative clinical parameters, may prove valuable for prediction of hepatic toxicity and long-term outcome following validation in a larger patient cohort.

Several limitations of this investigation should be noted: (1) sample size: the limited number of events and patients in this preliminary survival analysis prevented definitive multivariate Cox regression with cross-validation to power testing on independent patient cohorts but rather supported the identification of TLF as a candidate imaging biomarker; (2) variability in contouring of the normal liver on CT: despite a single radiation oncologist (S.A.) defining the anatomic liver volume region of interest, variation introduced by manual contouring methods can affect the denominator of the functional liver volume ratio; (3) standardization of imaging protocols: while a fixed threshold can be optimized for a particular imaging protocol at a single institution, any lack of protocol standardization that changes reconstruction algorithms or filters introduces a bias in the estimate of image parameters and propagates these errors across multiple institutions. Infrastructure and resources for establishing and standardizing quantitative FDG PET/CT parameters [41, 42] should also be

Bowen *et al. EJNMMI Research* (2016) 6:57

Page 8 of 10

adapted for quantitative imaging of liver function; (4) relative quantitation of SPECT/CT: uncertainty in modeling of photon collimator blurring, coupled with uncertainty in scatter and attenuation correction, have historically limited the absolute quantitative accuracy of SPECT imaging [43]. Advances in SPECT image reconstruction that incorporate improved physics modeling suggest an achievable quantitative accuracy with errors of less than 5 % [44]. This study bridges the gap between qualitative assessment and absolute quantitation of tracer uptake by SPECT imaging through definition of semi-quantitative relative metrics. Future investigations will explicitly calculate tracer uptake that has been standardized to provide absolute quantitative accuracy and precision characterization, which may improve quantitative imaging-based prediction of liver function; (5) the anatomic liver volume was defined on a radiotherapy planning CT acquired during a separate patient exam whereas functional liver parameters were defined on the sulfur colloid SPECT. Co-registration of these images can further introduce unwanted variation in liver morphology between the scans and bias the volume estimates. These sources of error in the FLV ratio may explain the lower degree of correlation to several clinical liver function parameters, although in general, the imaging parameters were defined as relative ratios to mitigate inter-patient variability in region-of-interest definition; and lastly, (6) SC SPECT/CT images Kupffer cells but not hepatocytes and may provide an indirect measure of liver function. Despite this possibility, prior large prospective studies have shown that imaging of Kupffer cells as measured by perfused hepatic mass independently predicts long-term clinical outcomes in patients with chronic liver disease [26] and is correlated with indocyanine green (ICG) clearance in plasma [28], a validated quantitative measure of global liver function [45].

Our data revealed several discordances between clinical parameters and imaging parameters of liver function. Of the individual clinical liver function parameters tested, splenomegaly had the poorest correlation to imaging parameters. The dependence of splenomegaly assessment on spleen size could in principle affect the liver-to-spleen ratio, but no statistical association was observed. One possible explanation for a lack of correlation is the non-standardized method for identifying clinical splenomegaly, which relies on either subjective evaluation or an arbitrary size cutoff, and is therefore prone to variability in visual perception. Such qualitative parameters have been removed from recent scoring systems like ALBI, which further reinforces the value in quantitative functional imaging.

Moreover, there was a subset of patients who had discordances between liver function as defined by TLF and clinically defined portal hypertension, a patient characteristic that is often used by surgeons as a relative contraindication for therapeutic surgical resection due to concerns of post-resection-related liver failure and morbidity. Of the patients with documented portal hypertension, 8 of 18 patients had TLF ratios above the cutoff value of 0.30 associated with improved overall survival, suggesting that up to 44 % of patients could be considered for surgical resection depending on the extent of planned resection [46]. This implication is hypothesis-generating and would need to be validated in a larger cohort of patients. However, the dissociation of portal hypertension and SC imaging as seen in our study is not entirely surprising as a prior study demonstrated that the degree of abnormality detected on SC SPECT scans was not necessarily directly due to portal hypertension but related to impaired SC extraction by the reticuloendothelial cells of the liver [47]. To provide additional clarity in the discordance between SC SPECT image parameters and clinical liver function parameters, future work in our group will include integrating other quantitative methods of assessing liver function such as ICG clearance.

The power of SC SPECT/CT imaging resides in the parameterization of liver function spatial heterogeneity. With improvements in SPECT/CT quantitative accuracy [48, 49], come opportunities to perform radiomics analysis [50] of image intensity, image texture [51], and image shape features to characterize functional heterogeneity. Standardization of image acquisition, reconstruction, and segmentation prior to radiomics analysis [52, 53] will be critical to improve reproducibility of radiomic feature estimation [54, 55]. Such high dimensional imaging feature sets will necessitate machine learning algorithms of training/(cross)-validation patient groups for feature selection and model parameter fitting, which may lead to robust risk stratification and outcome prediction of individual patients. It still remains unknown whether liver function radiomics will add clinical utility relative to simple imaging or tissue biomarkers of liver function. Nevertheless, prediction of both global magnitude and regional variation in liver function may critically inform on selection and guidance of liver-directed therapies within the scope of future investigations.

Conclusions

Sulfur colloid SPECT/CT demonstrated utility through correlation with clinical liver function and association with overall survival in our preliminary investigation of 30 HCC patients undergoing definitive radiotherapy. Following its validation as an imaging biomarker, the total liver function (TLF) ratio may complement existing risk classification schemes of HCC patients to promote precise selection of patients for liver-directed therapies *a priori*, including surgery, targeted intra-arterial embolizations, and even liver transplantation assessment, as well as modify radiotherapeutic strategies to reduce

Bowen et al. EJNMMI Research (2016) 6:57

Page 9 of 10

treatment-related liver morbidity and optimize tumor control [17]. Further investigation in a larger patient cohort spanning a wider range of clinical liver function is needed and underway to validate these initial findings.

Author details

[1]Department of Radiation Oncology, University of Washington School of Medicine, 1959 NE Pacific St, Box 356043, Seattle, WA 98195, USA. [2]Department of Radiology, University of Washington School of Medicine, 1959 NE Pacific St, Box 356043, Seattle, WA 98195, USA. [3]Department of Physics, University of Washington, 3910 15th Ave. NE, Seattle, WA 98195, USA. [4]Division of Gastroenterology, Department of General Internal Medicine, University of Washington School of Medicine, 1959 NE Pacific St, Seattle, WA 98195, USA.

References

1. El-Serag HB. Hepatocellular carcinoma. N Engl J Med. 2011;365(12):1118–27. doi:10.1056/NEJMra1001683.
2. Simonetti RG, Camma C, Fiorello F, Politi F, D'Amico G, Pagliaro L. Hepatocellular carcinoma. A worldwide problem and the major risk factors. Dig Dis Sci. 1991;36(7):962–72.
3. Cheng JC, Wu JK, Huang CM, Huang DY, Cheng SH, Lin YM, et al. Radiation-induced liver disease after radiotherapy for hepatocellular carcinoma: clinical manifestation and dosimetric description. Radiother Oncol. 2002;63(1):41–5.
4. Liang SX, Zhu XD, Xu ZY, Zhu J, Zhao JD, Lu HJ, et al. Radiation-induced liver disease in three-dimensional conformal radiation therapy for primary liver carcinoma: the risk factors and hepatic radiation tolerance. Int J Radiat Oncol Biol Phys. 2006;65(2):426–34. doi:10.1016/j.ijrobp.2005.12.031.
5. Hsin IF, Hsu CY, Huang HC, Huang YH, Lin HC, Lee RC, et al. Liver failure after transarterial chemoembolization for patients with hepatocellular carcinoma and ascites: incidence, risk factors, and prognostic prediction. J Clin Gastroenterol. 2011;45(6):556–62. doi:10.1097/MCG.0b013e318210ff17.
6. Salem R, Lewandowski RJ, Mulcahy MF, Riaz A, Ryu RK, Ibrahim S, et al. Radioembolization for hepatocellular carcinoma using Yttrium-90 microspheres: a comprehensive report of long-term outcomes. Gastroenterology. 2010;138(1):52–64. doi:10.1053/j.gastro.2009.09.006.
7. Teh SH, Nagorney DM, Stevens SR, Offord KP, Therneau TM, Plevak DJ, et al. Risk factors for mortality after surgery in patients with cirrhosis. Gastroenterology. 2007;132(4):1261–9. doi:10.1053/j.gastro.2007.01.040.
8. Child CG, Turcotte JG. Surgery and portal hypertension. Major Probl Clin Surg. 1964;1:1–85.
9. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg. 1973;60(8):646–9.
10. Durand F, Valla D. Assessment of prognosis of cirrhosis. Semin Liver Dis. 2008;28(1):110–22. doi:10.1055/s-2008-1040325.
11. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology. 2000;31(4):864–71. doi:10.1053/he.2000.5852.
12. Durand F, Valla D. Assessment of the prognosis of cirrhosis: Child-Pugh versus MELD. J Hepatol. 2005;42(1):S100–7. doi:10.1016/j.jhep.2004.11.015.
13. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. J Clin Oncol. 2015;33(6): 550–8. doi:10.1200/JCO.2014.57.9151.
14. Hsu CY, Hsia CY, Huang YH, Su CW, Lin HC, Lee PC, et al. Selecting an optimal staging system for hepatocellular carcinoma: comparison of 5 currently used prognostic models. Cancer. 2010;116(12):3006–14. doi:10.1002/cncr.25044.
15. Mergo PJ, Ros PR, Buetow PC, Buck JL. Diffuse disease of the liver: radiologic-pathologic correlation. Radiographics. 1994;14(6):1291–307. doi:10.1148/radiographics.14.6.7855342.
16. Garcea G, Ong SL, Maddern GJ. Predicting liver failure following major hepatectomy. Dig Liver Dis. 2009;41(11):798–806. doi:10.1016/j.dld.2009.01.015.
17. Bowen SR, Saini J, Chapman TR, Miyaoka RS, Kinahan PE, Sandison GA, et al. Differential hepatic avoidance radiation therapy: proof of concept in hepatocellular carcinoma patients. Radiother Oncol. 2015;115(2):203–10. doi:10.1016/j.radonc.2015.04.011.
18. Sorensen M, Frisch K, Bender D, Keiding S. The potential use of 2-[(1)(8)F]fluoro-2-deoxy-D-galactose as a PET/CT tracer for detection of hepatocellular carcinoma. Eur J Nucl Med Mol Imaging. 2011;38(9):1723–31. doi:10.1007/s00259-011-1831-z.
19. Sorensen M, Mikkelsen KS, Frisch K, Bass L, Bibby BM, Keiding S. Hepatic galactose metabolism quantified in humans using 2-18 F-fluoro-2-deoxy-D-galactose PET/CT. J Nucl Med. 2011;52(10):1566–72. doi:10.2967/jnumed.111.092924.
20. Cao Y, Wang H, Johnson TD, Pan C, Hussain H, Balter JM, et al. Prediction of liver function by using magnetic resonance-based portal venous perfusion imaging. Int J Radiat Oncol Biol Phys. 2013;85(1):258–63. doi:10.1016/j.ijrobp.2012.02.037.
21. Sirlin CB, Hussain HK, Jonas E, Kanematsu M, Min Lee J, Merkle EM, et al. Consensus report from the 6th International forum for liver MRI using gadoxetic acid. J Magn Reson Imaging. 2014;40(3):516–29. doi:10.1002/jmri.24419.
22. Ringe KI, Husarik DB, Sirlin CB, Merkle EM. Gadoxetate disodium-enhanced MRI of the liver: part 1, protocol optimization and lesion appearance in the noncirrhotic liver. AJR Am J Roentgenol. 2010;195(1):13–28. doi:10.2214/AJR.10.4392.
23. Cruite I, Schroeder M, Merkle EM, Sirlin CB. Gadoxetate disodium-enhanced MRI of the liver: part 2, protocol optimization and lesion appearance in the cirrhotic liver. AJR Am J Roentgenol. 2010;195(1):29–41. doi:10.2214/AJR.10.4538.
24. Beppu T, Hayashi H, Okabe H, Masuda T, Mima K, Otao R, et al. Liver functional volumetry for portal vein embolization using a newly developed 99mTc-galactosyl human serum albumin scintigraphy SPECT-computed tomography fusion system. J Gastroenterol. 2011;46(7):938–43. doi:10.1007/s00535-011-0406-x.
25. Wang H, Cao Y. Spatially resolved assessment of hepatic function using 99mTc-IDA SPECT. Med Phys. 2013;40(9):092501. doi:10.1118/1.4816655.
26. Everson GT, Shiffman ML, Hoefs JC, Morgan TR, Sterling RK, Wagner DA, et al. Quantitative liver function tests improve the prediction of clinical outcomes in chronic hepatitis C: results from the Hepatitis C Antiviral Long-term Treatment Against Cirrhosis Trial. Hepatology. 2012;55(4):1019–29. doi:10.1002/hep.24752.
27. Hoefs JC, Wang F, Kanel G. Functional measurement of nonfibrotic hepatic mass in cirrhotic patients. Am J Gastroenterol. 1997;92(11):2054–8.
28. Zuckerman E, Slobodin G, Sabo E, Yeshurun D, Naschitz JE, Groshar D. Quantitative liver-spleen scan using single photon emission computerized

Bowen *et al. EJNMMI Research* (2016) 6:57

Page 10 of 10

tomography (SPECT) for assessment of hepatic function in cirrhotic patients. J Hepatol. 2003;39(3):326–32.

29. Groshar D, Slobodin G, Zuckerman E. Quantitation of liver and spleen uptake of (99 m)Tc-phytate colloid using SPECT: detection of liver cirrhosis. J Nucl Med. 2002;43(3):312–7.

30. Esmaili J, Gholamrezanezhad A, Ebizadeh A. Correlation of liver-spleen scan findings with modified Child-Pugh classification. Rev Esp Med Nucl. 2008;27(2):99–102.

31. Satoh Y, Nambu A, Ichikawa T, Onishi H. Whole-body total lesion glycolysis measured on fluorodeoxyglucose positron emission tomography/computed tomography as a prognostic variable in metastatic breast cancer. BMC Cancer. 2014;14:525. doi:10.1186/1471-2407-14-525.

32. Kim TM, Paeng JC, Chun IK, Keam B, Jeon YK, Lee SH, et al. Total lesion glycolysis in positron emission tomography is a better predictor of outcome than the international prognostic index for patients with diffuse large B cell lymphoma. Cancer. 2013;119(6):1195–202. doi:10.1002/cncr.27855.

33. Keam B, Lee SJ, Kim TM, Paeng JC, Lee SH, Kim DW, et al. Total lesion glycolysis in positron emission tomography can predict gefitinib outcomes in non-small-cell lung cancer with activating EGFR mutation. J Thorac Oncol. 2015;10(8):1189–94. doi:10.1097/JTO.0000000000000569.

34. Chen HH, Chiu NT, Su WC, Guo HR, Lee BF. Prognostic value of whole-body total lesion glycolysis at pretreatment FDG PET/CT in non-small cell lung cancer. Radiology. 2012;264(2):559–66. doi:10.1148/radiol.12111148.

35. Kim CY, Hong CM, Kim DH, Son SH, Jeong SY, Lee SW, et al. Prognostic value of whole-body metabolic tumour volume and total lesion glycolysis measured on (1)(8)F-FDG PET/CT in patients with extranodal NK/T-cell lymphoma. Eur J Nucl Med Mol Imaging. 2013;40(9):1321–9. doi:10.1007/s00259-013-2443-6.

36. Van de Wiele C, Kruse V, Smeets P, Sathekge M, Maes A. Predictive and prognostic value of metabolic tumour volume and total lesion glycolysis in solid tumours. Eur J Nucl Med Mol Imaging. 2013;40(2):290–301. doi:10.1007/s00259-012-2280-z.

37. Lim R, Eaton A, Lee NY, Setton J, Ohri N, Rao S, et al. 18F-FDG PET/CT metabolic tumor volume and total lesion glycolysis predict outcome in oropharyngeal squamous cell carcinoma. J Nucl Med. 2012;53(10):1506–13. doi:10.2967/jnumed.111.101402.

38. Tanaka K, Endo I. ALPPS: short-term outcome and functional changes in the future liver remnant. Ann Surg. 2015;262(2):e88–9. doi:10.1097/SLA.0000000000000665.

39. Tanaka K, Matsuo K, Murakami T, Kawaguchi D, Hiroshima Y, Koda K, et al. Associating liver partition and portal vein ligation for staged hepatectomy (ALPPS): short-term outcome, functional changes in the future liver remnant, and tumor growth activity. Eur J Surg Oncol. 2015;41(4):506–12. doi:10.1016/j.ejso.2015.01.031.

40. Yokoyama Y, Nishio H, Ebata T, Igami T, Sugawara G, Nagino M. Value of indocyanine green clearance of the future liver remnant in predicting outcome after resection for biliary cancer. Br J Surg. 2010;97(8):1260–8. doi:10.1002/bjs.7084.

41. Buckler AJ, Boellaard R. Standardization of quantitative imaging: the time is right, and 18F-FDG PET/CT is a good place to start. J Nucl Med. 2011;52(2):171–2. doi:10.2967/jnumed.110.081224.

42. Buckler AJ, Bresolin L, Dunnick NR, Sullivan DC, Aerts HJ, Bendriem B, et al. Quantitative imaging test approval and biomarker qualification: interrelated but distinct activities. Radiology. 2011;259(3):875–84. doi:10.1148/radiol.10100800.

43. Shcherbinin S, Celler A, Belhocine T, Vanderwerf R, Driedger A. Accuracy of quantitative reconstructions in SPECT/CT imaging. Phys Med Biol. 2008;53(17):4595–604. doi:10.1088/0031-9155/53/17/009.

44. Zeintl J, Vija AH, Yahil A, Hornegger J, Kuwert T. Quantitative accuracy of clinical 99mTc SPECT/CT using ordered-subset expectation maximization with 3-dimensional resolution recovery, attenuation, and scatter correction. J Nucl Med. 2010;51(6):921–8. doi:10.2967/jnumed.109.071571.

45. Gottlieb ME, Stratton HH, Newell JC, Shah DM. Indocyanine green. Its use as an early indicator of hepatic dysfunction following injury in man. Arch Surg. 1984;119(3):264–8.

46. Benson 3rd AB, Abrams TA, Ben-Josef E, Bloomston PM, Botha JF, Clary BM, et al. NCCN clinical practice guidelines in oncology: hepatobiliary cancers. J Natl Compr Canc Netw. 2009;7(4):350–91.

47. Horisawa M, Goldstein G, Waxman A, Reynolds T. The abnormal hepatic scan of chronic liver disease: its relationship to hepatic hemodynamics and colloid extraction. Gastroenterology. 1976;71(2):210–3.

48. Bailey DL, Willowson KP. An evidence-based review of quantitative SPECT imaging and potential clinical applications. J Nucl Med. 2013;54(1):83–9. doi:10.2967/jnumed.112.111476.

49. Bailey DL, Willowson KP. Quantitative SPECT/CT: SPECT joins PET as a quantitative imaging modality. Eur J Nucl Med Mol Imaging. 2014;41 Suppl 1:S17–25. doi:10.1007/s00259-013-2542-4.

50. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6. doi:10.1016/j.ejca.2011.11.036.

51. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present… any future? Eur J Nucl Med Mol Imaging. 2016. doi:10.1007/s00259-016-3427-0.

52. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. PLoS One. 2014;9(7):e102107. doi:10.1371/journal.pone.0102107.

53. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. Sci Rep. 2015;5:11075. doi:10.1038/srep11075.

54. Leijenaar RT, Carvalho S, Velazquez ER, van Elmpt WJ, Parmar C, Hoekstra OS, et al. Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. Acta Oncol. 2013;52(7):1391–7. doi:10.3109/0284186X.2013.812798.

55. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. J Med Imaging (Bellingham). 2015;2(4):041002. doi:10.1117/1.JMI.2.4.041002.