ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

Biomedical Signal Processing and Control 26 (2016) 109–116

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Automatic detection of misleading blood flow values in CT perfusion studies of lung cancer

Alessandro Bevilacqua[a,b,*], Domenico Barone[c], Silvia Malavasi[b,d], Giampaolo Gavelli[c]

[a] Dept. of Computer Science and Engineering (DISI), University of Bologna, Italy
[b] Advanced Research Centre for Electronic Systems (ARCES), University of Bologna, Via Toffano 2/2, 40125 Bologna, Italy
[c] Radiology Unit, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, via Piero Maroncelli 40, 47014 Meldola, FC, Italy
[d] Dept. of Electrical, Electronic and Information Engineering "Guglielmo Marconi" (DEI), University of Bologna, Italy

ARTICLE INFO

Article history:
Received 29 July 2015
Received in revised form
19 December 2015
Accepted 10 January 2016

Keywords:
Quantitative imaging
Error analysis
Image processing
Imaging artefacts
Cancer

ABSTRACT

In the oncology field, the anti-angiogenetic therapies aim at inhibiting tumour vascularization, that is the development of new capillary blood vessels in tumours, that allows them to grow and spread and, potentially, to metastasi. Computed tomography perfusion (CTp) is a dynamic contrast-enhanced technique that has emerged in the last few years as a promising approach for earlier assessment of such therapies, and of tumour response, in general, since functional changes precede morphological changes, that take more time to become evident. However several issues, such as patient motion and several types of artefacts, jeopardize quantitative measurements, this preventing CTp to be used in standard clinics. This paper presents an original automatic approach, based on the voxel-based analysis of the time–concentration curves (TCCs), that allows emphasizing those physiological structures, such as vessels, bronchi or artefacts, that could affect the final computation of blood flow perfusion values in CTp studies of lung cancer. The automatic exclusion of these misleading values represents a step towards a quantitative CTp, hence its routine use in clinics.

1. Introduction

In the last years the computed tomography perfusion (CTp) has aroused lively interest because of its capability of providing morphologically detailed functional maps that could find application in monitoring functional activity of tumours at their different stages [1], predicting treatment outcome or early therapeutic response of anti-angiogenetic therapies [2–6], before morphological changes become visible [7]. This widely available and non-invasive technique relies on the estimation of tissue contrast agent delivery, and on the corresponding haemodynamic parameters, which can be derived from the analysis of the tissue time-concentration curves (TCCs) signals generated by the contrast agent before, during, and after reaching the tumour lesion. As such, they can be employed to detect changes in its vascular structure, hinting at possible anomalies in blood supply (i.e., tumour angiogenesis [2,8]). The main obstacles preventing the use of CTp in the standard clinical practice is represented by the difficulty to measure its reproducibility and, in the last analysis, the reliability of perfusion values [9]. In its turn, this arises from the difficulty to achieve reliable TCCs because of motion artefacts due to breathing [10] or a high concentration of contrast agent that may also yield streaking [11] or simulate partial volume effects in the repeated acquisitions of a CTp study. In addition, streaks or dark bands may also originate from beam hardening (for instance, induced by the presence in the lesion's neighbourhood of a main vessel flooded with contrast agent). A detailed description of possible artefacts in CT examinations, also holding for CTp examinations, can be found in [12,13]. Besides artefacts [14], also vessels [15] and bronchi [16] are usually manually excluded from the perfusion analysis not to jeopardize the visual analysis of perfusion colour maps nor the automatic computation of local or global statistical indexes regarding perfusion values.

In general, the unreliable perfusion values in colour maps are simply considered as those being out of range of physiological parameters (for instance, due to vessels [17] and are detected and excluded by manually adjusting an appropriate window level [9]. So far, the TCC fitting errors and goodness of fit indexes have been mainly used to evaluate the reliability of given simulated model fitting, from a theoretical point of view, in lung CT [18] or liver MRI [19] perfusion studies, rather than to assess the voxel-based reliability of perfusion values. Just recently, some

* Corresponding author at: Dept. of Computer Science and Engineering (DISI), University of Bologna, Italy. Tel.: +39 0512095409; fax: +39 0512095410.
E-mail address: alessandro.bevilacqua@unibo.it (A. Bevilacqua).

http://dx.doi.org/10.1016/j.bspc.2016.01.004
1746-8094/

works have appeared to try addressing explicitly the reliability of TCCs in relation to the outcome of perfusion studies, suggesting methods to detect where the error in fitting a TCC according to a given pharmacokinetic model could prejudice the computation of correct perfusion values [20] or even to improve the way a TCC is built, so as to reduce the number of possible flawed TCCs [21]. However, they do not associate fitting errors to their causes, if not to generic motion artefacts and, not at all, to anatomical structures.

This methodological work presents a novel quantitative and automatic approach to detect those anatomical structures (mainly vessels and bronchi) and those regions undergoing CTp reconstruction and acquisition artefacts, that could compromise the correct interpretation of a CTp colour map and, ultimately, the clinical outcome. The approach is based on the computation of a statistical error index connected to the quality (i.e., goodness of fit) of TCCs. The ability of our method to automatically remove the "misleading" regions is assessed and compared with the performance of two 25-year experienced radiologists who detected, and manually bounded using a graphic tablet, the anatomical structures and the regions undergoing artefacts. The main errors made when operating manually, and their consequences on the definite perfusion maps, are also discussed. Moreover, we analyze how mean of perfusion values and, above, their standard deviation and coefficient of variation (CV), change before and after removing automatically segmented regions. Finally, we also discuss some meaningful comparisons between definite colour maps achieved by using our approach and the manual thresholding on BF values commonly used by readers.

The paper is organized as follows. Section 2 describes the materials and the methods employed from image acquisition to building of the final perfusion maps, including the automatic histogram-based image segmentation approach developed by exploiting goodness-of-fit errors. Section 3 presents the choices performed to evaluate the experimental results which are subsequently analysed in Section 4 and discussed in Section 5. Finally, drawing conclusions are reported in Section 6.

2. Materials and methods

46 patients with non-small cell lung cancer (NSCLC) underwent CTp. As the inclusion criteria, patients over eighteen, with lesions having maximum transverse diameter greater than 2.5 cm, and area wider than 3.14 cm^2 were considered. Consequently, the remaining 22 patients (age range, 36–81 years; median, 66 years; mean, 64.8 years; 17 men, 5 women) were enrolled for the study, for a total amount of thirty-four examinations, with as many lesions. The data set utilized in this study was selected from a prospective CTp study approved by the institutional review board.

2.1. Perfusion model

To achieve a perfusion examination, a region of interest (ROI) has to be repeatedly sampled. The kinetics of a contrast agent in tissue directly depends on the tissue properties themselves, so it can be used to estimate them. In particular, the well-known Mullani–Gould model [22], grounding on the simplified assumption of no venous washout [23], exploits the tissue's blood flow (BF) to estimate perfusion by measuring the instantaneous concentration of the contrast agent over time, limitedly to the first pass phase, where the amount of contrast agent diffusing from the intra-vascular to the extra-cellular compartment is negligible. Accordingly, the Hill's equation [24], commonly used in pharmacodynamic models to describe non-linear drug dose–response

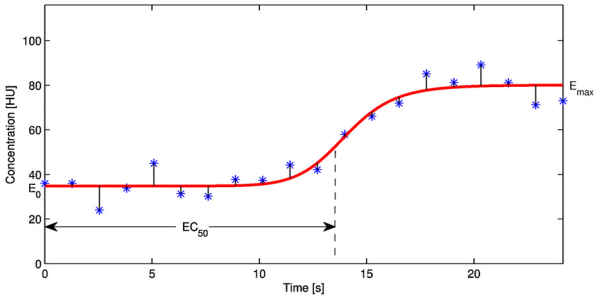

Fig. 1. The ideal TCC $\hat{f}(t)$ according to the Hill model. A time-series of sampled HU values (blue stars) and their representative TCC (in red), with indication of the different parameters. α mostly affects the slope of the TCC and is omitted since it has not any geometrical representation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

relationships, can be also employed to model the pharmacokinetic of the contrast agent [25], thus representing a TCC:

$$\hat{f}(t) = E_0 + (E_{MAX} - E_0)\frac{t^\alpha}{(EC_{50})^\alpha + t^\alpha} \quad (1)$$

where E_0 and E_{MAX} represent the baseline and the saturation expressed in Hounsfield Units (HU) values, respectively, EC_{50} is the time instant of half-maximum response concentration of the curve and α is the non-linear parameter mostly affecting the slope of the curve. Fig. 1 points out an ideal TCC and a time-series of HU values. The BF values are then calculated according to the *maximum slope method* [26] during the first-pass phase:

$$BF = \frac{\frac{d\hat{f}_t(t)}{dt}|_{\max}}{\hat{f}_a(t)|_{\max}} \quad (2)$$

that represents the ratio between the maximum slope of the tissue TCC $\hat{f}_t(t)$ and the peak density reached by the TCC of the arterial input $\hat{f}_a(t)$. For the sake of clarity, hereinafter the tissue TCC $\hat{f}_t(t)$ will be indicated as $\hat{f}(t)$.

2.2. CTp protocol and image acquisition

CTp examinations were performed using a 256-slice CT system (Brilliance iCT, Philips Medical Systems, Best, The Netherlands). Patients were instructed for breath-hold and laid in the supine (feet first) position. An initial full-body, unenhanced, low-dose axial CT scan was performed to identify the target lesion at baseline condition. Then, an intravenous 50-mL bolus of contrast agent (Iomeron, Bracco, Milan, Italy) was injected at 5 mL/s. 5 s after, axial CTp scan was performed at fixed tube voltage (80 kV) and current (250 mA), with an exposure of 100 mAs. A single acquisition of 25-second duration consisted of $M = 20$ scans centred on the target lesion, yielding volumes of 55 mm of z-coverage with $N = 11$ slices (5-mm thickness each, 0.4-second rotation time). Finally, image data were then reconstructed to 220 cine images (512 × 512 pixel, 350 mm × 350 mm, 5-mm slice spacing, 1.25-second temporal resolution).

2.3. Building perfusion maps

Before computing voxel-based perfusion values, the sequences of slices have been registered according to the 3D rigid procedure described in [20]. The target lesions and the arterial input (aorta) were selected in agreement by two 25-year experienced radiologists, on a *reference slice*, where two ROIs for the aorta and the lesion section (hereafter, *reference ROI*), respectively, were manually outlined. Let $f(x, y, z, t_m)$ be the generic voxel, sampled at the time instant t_m, $m \in \{1, \ldots, M\}$, with (x, y) being the pixel

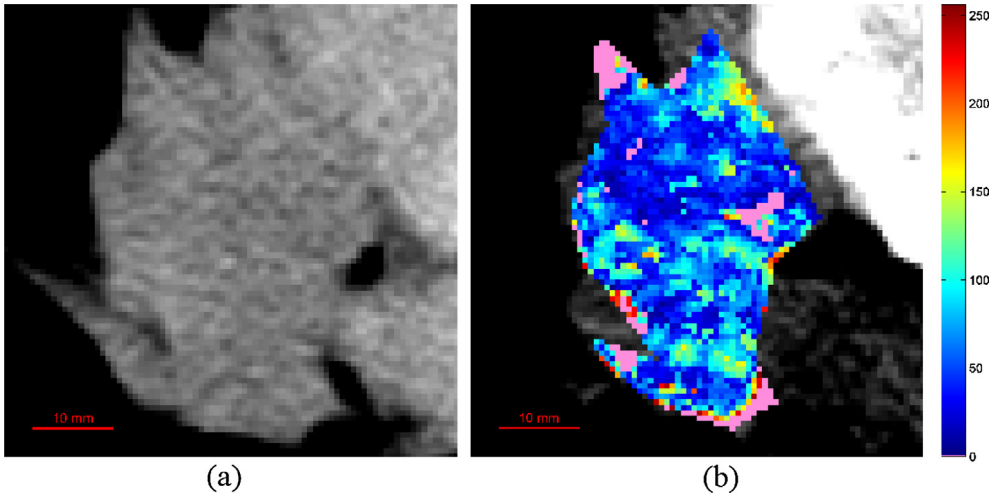

Fig. 2. ROI of lesion ID8 (left), and related BF colour map (right). (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

coordinates referring to the slice z, $z \in \{1, \ldots, N\}$, and $f(x, y, z, t)$ a generic sampled TCC of the resulting *reference sequence*, at ROI's voxel location (x, y, z) (hereinafter, $f(t)$ for the sake of readability). Then, all the sampled TCCs are fitted according to the model described in Eq. (1) by using the Levenberg–Marquardt least-squares minimization routine available in MATLAB (MathWorks, Natick, MA), and the resulting $\hat{f}(t)$ utilized to calculate the functional BF maps. Fig. 2 shows the ROI referring to lesion ID8, with related BF colour map (in red, the higher values).

BF values strictly lower than 1 mL/min/100 g were considered as being unlikely compliant with physiological values and rather ascribable to numerical errors (since the computing method forces BF to have positive values only), hence marked as unreliable and highlighted in the colour map with the "pink" colour.

2.4. Automatic image segmentation

In order to detect the presence of vessels, bronchi, and artefacts, we used a histogram-based method, referring to a voxel-oriented index of *goodness-of-fit* of the fitted curve in correspondence of the TCC samples. In particular, for each TCC we first considered, for each sample time instant m, the absolute value of residual, ϵ_m, defined as:

$$\epsilon_m = |f(t_m) - \hat{f}(t_m)| \tag{3}$$

that is, the absolute difference (expressed in HU) between the concentration value *measured* in that voxel and the *expected* value derived from the fitting curve at the same time instant. Then, the temporal mean of these M absolute values, μ_ϵ, defined as:

$$\mu_\epsilon = \frac{1}{M}\sum_{m=1}^{M}\epsilon_m \tag{4}$$

represents the error index for a given voxel. The μ_ϵ errors can be also represented with a colour map, as shown in Fig. 3(a). Of course, we expect that low μ_ϵ values point out voxels with reliable BF values. On the opposite, a high μ_ϵ value for a TCC would suggest to exclude its corresponding voxel from the perfusion analysis. In order to find out, for each examination, a proper threshold T capable of discriminating between "low" and "high" error values, we considered the distribution (or better, the histogram, Fig. 3(b)) of these μ_ϵ errors so as to take advantage of statistical indexes, such as the mean $E[\mu_\epsilon]$ and the standard deviation σ (although the correct notation should be σ_{μ_ϵ}, for the sake of readability we always use σ). Accordingly, as the first approach we considered the 2-σ rule [20], implying that all voxels whose $\mu_\epsilon > E[\mu_\epsilon] + 2\sigma$ are highlighted and discarded. In that case, the threshold is defined as $T_{2\sigma} = E[\mu_\epsilon] + 2\sigma$. Histograms have positive values only and long right tails and, subsequently, the 2-σ rule is quite conservative. For this reason, we proposed a second well-established and more aggressive approach (also known as "the triangle method") conceived for "removing" tails in unimodal distributions [27]. Hereafter, we refer to its threshold as T_T. In particular, given a straight line joining the modal and the last bin, T_T is defined as the index of the bin having the maximum perpendicular distance to that line. As a matter of fact, this geometric approach is less conservative than the 2-σ rule, hence $T_T < T_{2\sigma}$. Fig. 3(b) shows an example of thresholded histogram (referring to lesion ID8), in cyan, with the blue colour representing errors such that $\mu_\epsilon > T_{2\sigma}$ and the red colour referring to $T_T < \mu_\epsilon \leq T_{2\sigma}$. Fig. 3(c) shows the corresponding thresholded colour mask, which emphasizes how pixels with same colours (i.e., undergoing same thresholding) appear as structural aggregates, with spatial coherence, meaning that same errors are shared by quite connected regions. Also, usually red colours extend and surround blue regions, this pointing out a transition from high-error to reliable regions. As a matter of fact, while histograms are used to automatically define a proper threshold (mainly for the triangle method), the thresholded error colour masks are utilized for the subsequent comparison with the original reference slices manually annotated.

2.5. Manual annotation

For each examination, the initial sequence of 20 scans is looked through to detect and annotate the main causes affecting the analysis of perfusion maps, represented by artefacts, and anatomical structures such as vessels and bronchi. The purpose of this manual annotation stage was twofold. On the one hand, we wanted to achieve the best possible performance, and to this purpose the first radiologist (hereafter, Reader A) was given no time limit for his analysis. On the other hand, we aimed to obtain realistic results, and for this reason the second radiologist (hereafter, Reader B) had to complete his analysis in a time compliant with a "routine" CTp study, under the same conditions as those in a clinical environment. The analyses were accomplished in a blind fashion.

First of all, the two Readers analysed the whole set of images in cine-mode and detected the presence of vessels and bronchi inside the lesion. After that, they manually outlined in the reference slice the anatomical structures also visible in the reference sequence (Fig. 4(a) and (b)) using a graphic device (Intuos®Pro, Wacom,

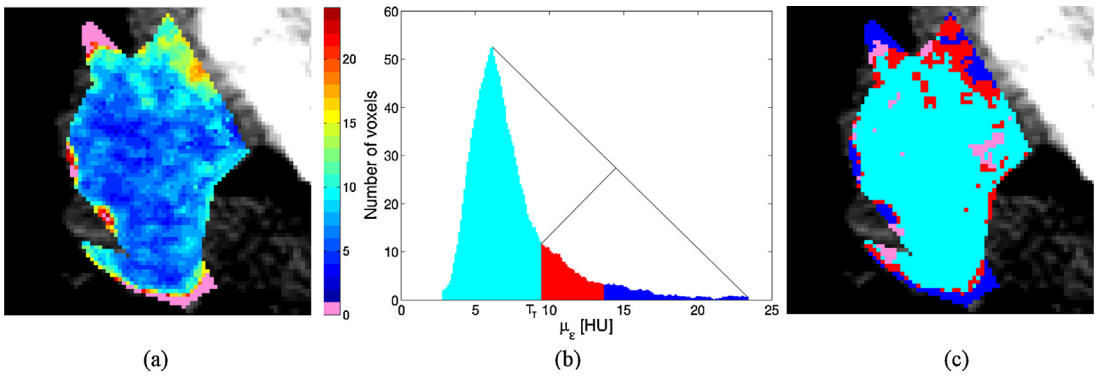

Fig. 3. An example of the μ_ϵ error colour map (a). Its thresholded histogram (b) and the corresponding thresholded colour mask (c) are shown, referring to lesion ID8, where the blue and the red regions correspond to pixels whose fitting error was such that $\mu_\epsilon > T_{2\sigma}$ and $T_T < \mu_\epsilon \leq T_{2\sigma}$, respectively. Pink pixels come from zero-value perfusion voxels, hence marked as unreliable, while the cyan region is composed by the "survived" pixels, whose perfusion values can be correctly used. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Krefeld, Germany). Then, the radiologists identified, and manually annotated in the reference slice, the lesion's regions undergoing the different type of artefacts, although mainly arising from partial volume effects induced by residual motion and from beam hardening (Fig. 4(c) and (d)). As one can see in Fig. 4, while vessels and bronchi are usually well identifiable, detecting artefacts is much harder and, even when succeeding, both boundaries and extent cannot be detected with accuracy.

3. Assessment of results

One of the purposes of this work is to determine and "quantify" how skilled two very expert readers are to detect and circumscribe the possible causes of errors in perfusion values, stemming from the CTp sequence. This task was carried out by comparing the ROIs manually drawn and the thresholded error colour mask automatically computed, using statistical indexes. In addition, some perfusion maps are compared where errors are removed through our approach versus manual thresholding on perfusion values.

Finally, mean, standard deviation and CV (computed as the ratio of the standard deviation to the mean) of the BF values in the original examination, and of the remaining BF values after thresholding using 2-σ rule and triangle method are computed and compared.

3.1. Comparison between annotated slices and thresholded error masks

The way the anatomical structures and artefacts were detected and outlined, mainly the uncertainty in delineating artefacts, drove our choice regarding the approach to compare the regions manually outlined with the outcome of our automatic error detection approach. For this reason, we have considered the *number* of numerical structures found (or missed) and the *presence* of artefacts in a given region, since their extent was not visually assessable. Nevertheless, different types of artefacts, in the same lesion are considered separately. Accordingly, there are four different outcomes from the matching procedure: "hit" or true positive (TP), false positive (FP), true negative (TN), "miss" or false negative (FN), all arranged into three 2×2 contingency tables, for vessels, bronchi, and artefacts, respectively [28]. Also, it is worth recalling that we are interested in detecting only those acquisition and reconstruction artefacts, vessels, and bronchi that can hamper the perfusion values, rather than artefacts or anatomical structures in themselves. Therefore, as the reference (i.e., the "ground truth") we chose the thresholded error masks and the regions manually outlined by radiologists were considered as the test condition. For instance, when a Reader detects an item that has no correspondence on the thresholded error mask (this representing a FP), this does not necessarily hints a mistake, rather most probably what detected does not hamper the computation of perfusion values. On the other side, in case that a Reader does not outline any item in correspondence of an error structure present in the thresholded error map, this is considered a FN only after that a radiologist confirmed the nature of that error structure, that is vessel, bronchus or artefact. The simple case is when no error structures are detected, nor visually neither automatically, when the TN number in each contingency table is increased by one. As far as the TP are concerned, at the beginning we hypothesized to segment the thresholded error maps into connected ROIs and perform an automatic matching between manual and computed ROIs. Then, several known strategies could have been considered to decree a match, ranging from inclusion criteria to overlapping thresholds, even weighted to allow for possible different sizes [29]. As a matter of fact, this could work for anatomical structures, but would fail for artefacts and using

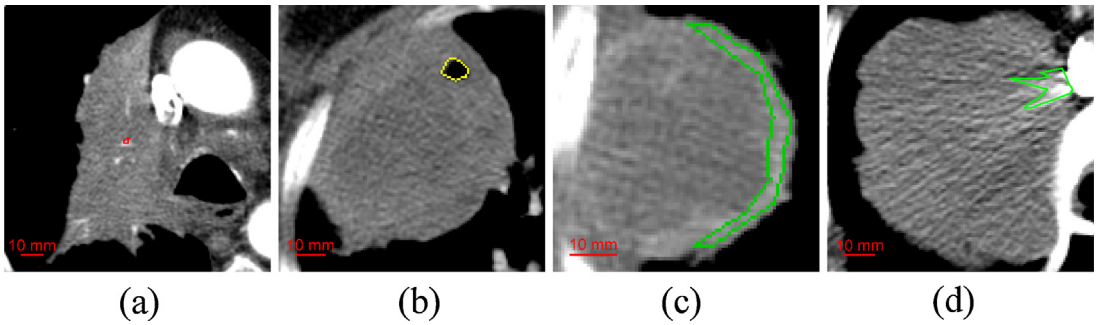

Fig. 4. From left to right, lesions ID15, ID34, ID14, ID17. A small vessel (a) and a bronchus (b) manually contoured by Readers A and B, respectively; artefact regions induced by motion (c) and beam hardening (d), graphically annotated by Readers A and B, respectively.

Table 1
Table summarizing contingency tables and statistical indexes relative to the analysis conducted by the two radiologists on the presence of vessels, bronchi, and artefacts.

Type	P	N	Reader A						Reader B					
			TP	TN	FP	FN	FPR (%)	FNR (%)	TP	TN	FP	FN	FPR (%)	FNR(%)
Vessels	18	26	16	26	0	2	0	11	12	24	2	6	8	33
Bronchi	8	27	7	27	0	1	0	13	7	26	1	1	4	13
Artefacts	41	9	26	7	2	15	22	37	20	3	6	21	67	51

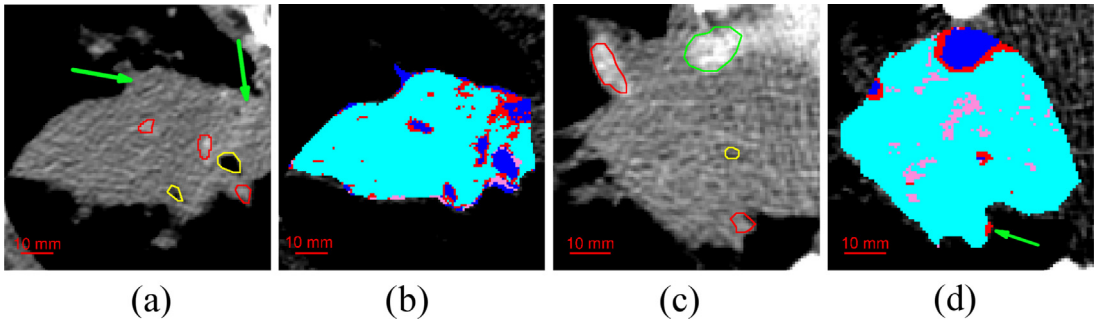

Fig. 5. Lesions ID29 ((a) and (b)), and ID19 ((c) and (d)). Three vessels (in red) and two bronchi (in yellow), manually highlighted by Reader A, together with two missed artefact regions (green arrows) (a), with corresponding thresholded error masks (b); two vessels (in red) detected by Reader A only, one bronchus (in yellow), and one artefact from beam hardening (in green), manually contoured by Reader A (c), with corresponding thresholded error mask (d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

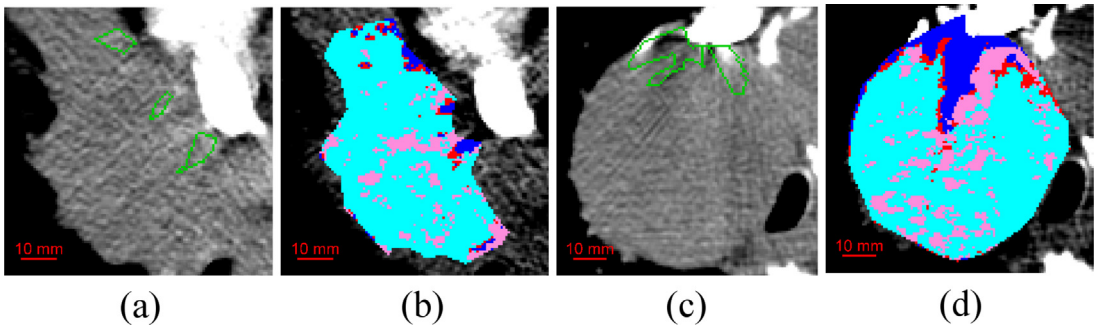

Fig. 6. Lesions ID12 ((a) and (b)), and ID33 ((c) and (d)). A beam-hardening artefact manually detected with a high accuracy by Reader B (a) and the corresponding thresholded error mask (b); a well-defined shape of an artefact drawn by Reader B (c), that however fails in detecting the wide extent of its effects, highlighted in the error mask (d).

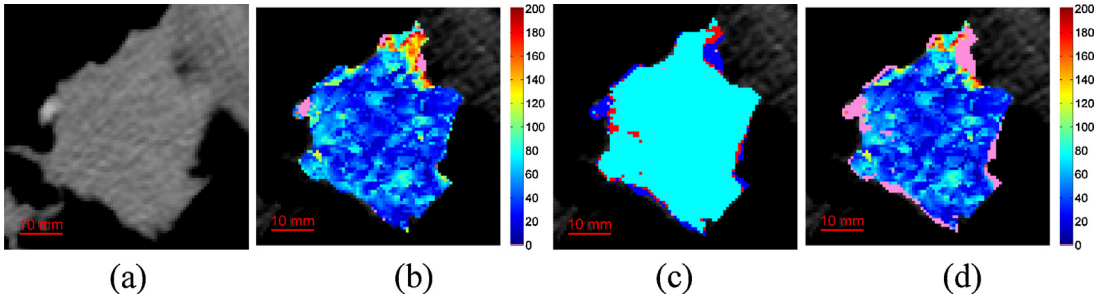

Fig. 7. Lesions ID32, where no misleading regions are detected (a), with its colour BF map manually thresholded (b); the thresholded error mask (c) and the final BF map achieved after excluding the error regions and a subsequent visual thresholding by using the same value as before (d).

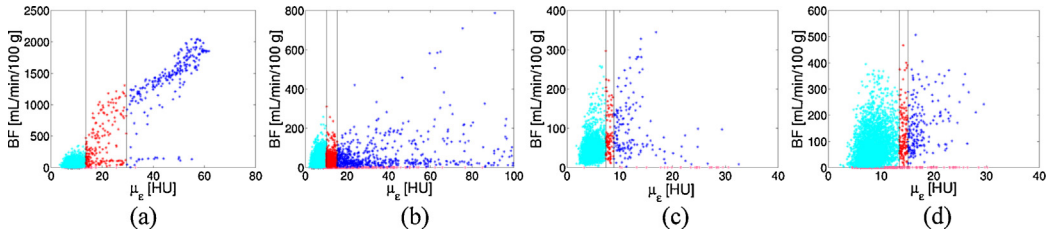

Fig. 8. Scatter plots to explore the relationship between BF values and μ_ϵ errors, referring to cases ID19 (Fig. 5(c)) (a), ID12 (Fig. 6(a)) (b), ID32 (Fig. 7(a)) (c), ID33 (Fig. 6(c)) (d). As one can see, errors also affect physiological BF values. Blue dots are detected by the 2-σ rule, while the red ones are detected by the triangle method only. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

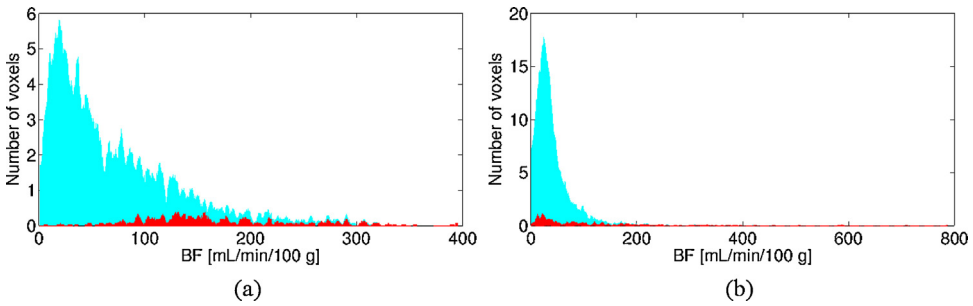

Fig. 9. BF values of the whole lesion (cyan colour) and those removed by the triangle method (red colour), referring to lesion ID14 (a) and ID33 (b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

different approaches would not have been fair. Besides that, in any case determining the FNs would require the intervention of radiologists. In addition, if our purpose is to quantify the readers errors (with the meaning defined at the beginning of this paragraph), then it is enough to find out how many error causes, in terms of structures and artefacts, are missed visually. For all these reasons, we decided that a visual matching performed by the Readers would be appropriate for our purposes.

3.2. Statistical and data analysis

To assess the Readers' performance we measured how their errors, in terms of FP and FN, impact on the total number of negative (N =FP + TN) and positive (P = TP + FN) cases, respectively. This is given by the frequency of their errors over the total number of cases, namely, the FP rate (FPR = FP/N) and the FN rate (FNR = FN/P), computed on each of the three contingency tables. FPR and FNR are also known as Type I and Type II statistical errors, respectively. Also, in order to better understand the implications of these errors, it could be useful to think of them as functions of specificity (SP) and sensitivity (SE), these being entities more used in clinics, where FPR = 1-SP and FNR = 1-SE.

The paired two-sided Wilcoxon signed rank test was used to compare the mean BF values before and after thresholding, while the one-tail *F* test and *Z* test were computed to assess the reduction of standard deviation and CV, respectively, after thresholding. All tests are implemented in R software (version 3.2.1, The R Foundation for Statistical Computing). *P* values $\leq$.01 were considered as being statistically significant.

4. Experimental results

Table 1 shows the outcome of the comparison between manual annotations and computed thresholded error maps, related to the 34 examinations. The first consideration concerns the possible causes of alteration of perfusion data (column 'P'), that in this study were artefacts (41), vessels (18, with mean area of about 16 mm^2), and bronchi (8, mean area around 23 mm^2), in descending frequency order. Also, FPR and FNR columns highlight that Reader A always performed better than, or at most as the same as, Reader B. In particular, Reader A had a specificity of 100% regarding the detection of bronchi and vessels, although Reader B also showed quite good performance in terms of FPR for these physiological structures. As an example, Fig. 5(a) shows a manual detection by Reader A of bronchi and vessels, all of them detected automatically also by the $T_{2\sigma}$ threshold (Fig. 5(b), blue regions). However, both Readers' performance decrease when it does not come down to missing structures. For instance, Fig. 5(c) shows two vessels detected by Reader A only, and highlighted by our approach (one of which, pointed out by the green arrow in Fig. 5(d), detected through T_T only).

Things change as Readers face artefacts. Although manifest artefacts can be detected by both Readers (e.g., the beam-hardening artefact of Fig. 5(c), in green), apparently more subtle artefacts, like the two shown in Fig. 5(a), originating from partial volume effects (left green arrow) and beam hardening (right arrow), are missed by both Readers (this representing two FNs), whereas they are correctly detected by our automatic method Fig. 6(b). Or else, even when artefacts are correctly detected, with a surprisingly high precision (the three green sharpened ROIs in Fig. 6(a)), their ramifications (mostly highlighted by pink pixels in Fig. 6(b)) are left out of consideration – for the sake of honesty, they are almost impossible to be assessed to the naked eye. Nonetheless, in the lesion shown in Fig. 6(c), Reader B strives to argue the extent of this beam-hardening artefact induced by the high concentration of the contrast agent into vena cava, during the initial phase of CTp acquisition. However, although the shape was not far wrong, again the extent is heavily underestimated, as shown by the outcome of the automatic method in Fig. 6(d). Moreover, this is an example where any clinical consideration regarding this case could be severely misleading, due to more than one third of the lesion (2537 out of 7299 voxels) being unreliable in a scatter manner, and the examination is strongly suggested to be definitely excluded.

The great benefits of using an automatic method to exclude unreliable pixels can be also appreciated when considering the most spread alternative, that is manual thresholding on perfusion colour maps. Fig. 7(a) shows an interesting example of lesion, coming from quite stable CTp sequences, where neither physiological structures nor artefacts where detected. Accordingly, the few pink pixels in the corresponding perfusion colour map (Fig. 7(b)) only arise from the voxels with too high perfusion values, removed by Readers through agreed manual thresholding. The resulting BF map is plausible and shows a perfusion peak, on the top right border, whose BF values are around 200 mL/min/100 g. However, the thresholded error mask in Fig. 7(c) clearly highlights a high-error region, just in correspondence of the perfused area, correctly removed as shown by the presence of the pink pixels in the BF map in Fig. 7(d). This is a meaningful example where a simple manual thresholding fails, since the error affects voxels with apparently acceptable perfusion values. The scatter plots in Fig. 8 resume this concept, with colour convention as the same as Fig. 3(b). Although, as expected, there are cases where the errors are almost exclusively associated to non-plausible BF values (Fig. 8(a), referring to lesion of Fig. 5(c)), most of times errors are shared between high and low perfusion values, as shown in Fig. 8(b) (referring to lesion of Fig. 6(a)) and Fig. 8(c) (pertaining to Fig. 7(a)). Nonetheless, in case of wide extending artefacts is not rare to find lesions where errors affect almost exclusively the voxels whose BF values fall in the physiological range, as shown in Fig. 8(d), referring to the lesion shown in Fig. 6(c).

The results presented above can be better acknowledged when looking at the meaningful BF histograms of Fig. 9, referring to

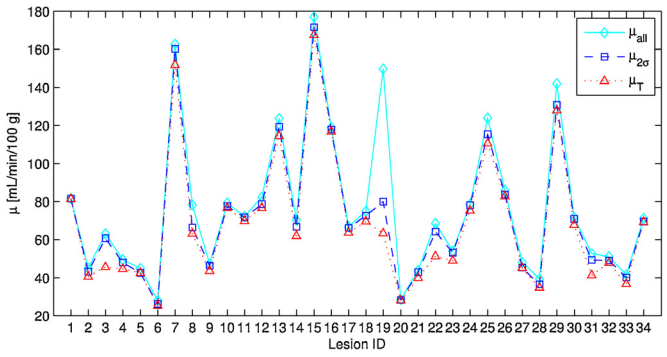

Fig. 10. Mean of BF values (μ) in the original examinations (solid cyan line, diamonds), after 2-σ (dashed blue line, squares) and triangle (dotted red line, triangles) thresholding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

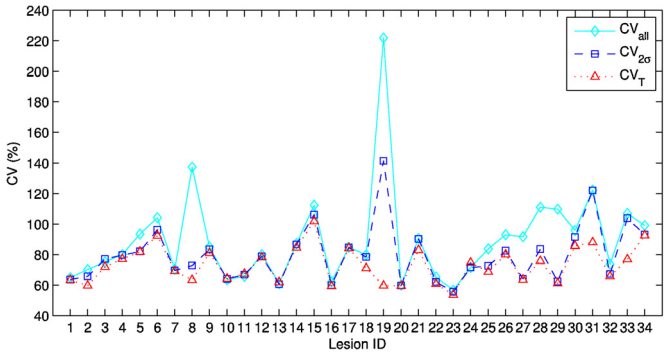

Fig. 12. CVs related to the BF values in the original examinations (solid cyan line, diamonds), after 2-σ (dashed blue line, squares) and triangle (dotted red line, triangles) thresholding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lesions ID14 (a) and ID33 (b), showing which BF values are removed by thresholding (for a better readability, only the triangle method is shown, in the red colour). As one can see, there are cases (Fig. 9(a)) where non-physiological BF values mostly lie on tail as well as there are lesions (Fig. 9(b)) where removed BF values are spread throughout the whole range, almost proportionally, and the average of removed BF values is almost the same as that of the original distribution. As a consequence, as shown in Fig. 10 mean values basically remain unchanged for 2-σ thresholding (6 cases out of 34), almost half of them (18 out of 34) with triangle thresholding reduce ($P < .005$). As regards standard deviations, they all reduce (Fig. 11), most of them with statistical significance ($P < .005$, 20 cases with 2-σ thresholding, and even 29 using triangle). In particular, with the 2-σ rule three out of the 20 cases have P values ranging from .003 and 10^{-4}, and the remaining ones are far below 10^{-5}, while with triangle all the 29 cases show $P < 10^{-5}$. These reductions partly reflect in CVs, reported in Fig. 12, which diminish in 12 cases with 2-σ and 16 cases with triangle thresholding ($P < .005$).

5. Discussion

Quantitative imaging has aroused an increasing interest in these last years at the same rate as the need of personalized therapies [30]. The uncertainty on reliability of measured perfusion values represents one of the most relevant causes preventing the diffusion of perfusion CT in clinical oncology, to assess the outcome of anti-angiogenetic treatments. Several strategies have been adopted in these last years to improve the quality of the image sequences in the CTp studies, ranging from the adoption of advanced CT technologies to more and more advanced image registration approaches for

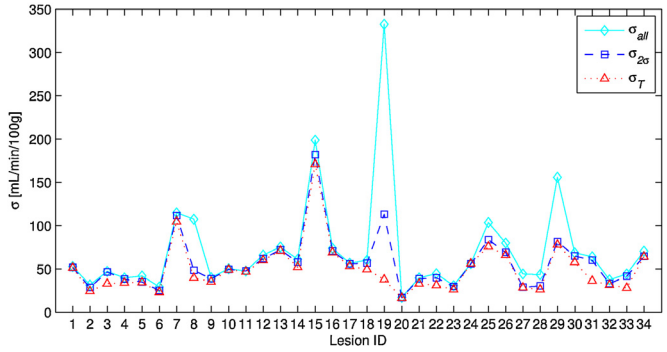

Fig. 11. Standard deviations (σ) of BF values in the original examinations (solid cyan line, diamonds), after 2-σ (dashed blue line, squares) and triangle (dotted red line, triangles) thresholding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

effective motion compensation. However, the presence of errors in the perfusion maps is far from being avoided, also because most of them inhere the CT examination itself, which can induce several types of artefacts. Rather, the early issue is admitting that CTp maps contain errors and then finding out where they are.

As a matter of fact, it is a common practice to visually analyze the CTp sequence to exclude from the study, *via* manual or assisted segmentation, the anatomical structures (mainly vessels and bronchi). However, dealing with artefacts is much more challenging because, if they could limitedly affect common visual CT morphological analysis, they definitely prejudice the quantitative analysis that should stand behind CTp studies.

In the first place, this research work shows that the highest values of our error index derived from goodness of fit are related just to those regions corresponding to anatomical structures or affected by diverse types of artefacts, altering the computation of the perfusion colour map. In particular, the results presented confirm that removing unreliable BF values could yield significant changes (commonly, reductions) in the mean BF value of a lesion, often used as a significant statistical parameter in several clinical studies [31,32], where global ROIs are usually considered, this contributing to carry out more precise clinical assessments. Also, what this study proves is that visual-based annotation is undoubtedly inadequate to discover the source of uncertainty in a CTp sequence. At most, devoting more time to this task reduces the number of FP and FN, always improving FPR and FNR (except for bronchi), but the improvement is quite relevant limitedly to the anatomical structures. In fact, the best results point out that 37% of times artefacts are missed and the remaining times their extent is widely underestimated, up to drawing clinical consideration on a perfusion map that should have been discarded because mostly unreliable. On the other hand, this study highlights that the different types of artefacts represent the most spread causes possibly hampering perfusion parameter values, much more than physiological structures, perhaps this not being surprisingly and in line with what most probably every radiologists expect. However, what seems to be heavily underestimated is the extent of the negative effects of artefacts, this representing the highest risk if relying only on visual assessment of CTp perfusion sequences and maps.

In the second place, it has not to be forgotten that perfusion maps are usually provided “as is”, without any pixel-wise indication on their reliability and what radiologists usually do through visual analysis is, at most, just detecting perfusion peaks not compliant with physiological values and exclude them. This behaviour is prone to either neglect local perfusion peaks or including all local errors if compatible with physiological outcome. This work shows that in most cases errors are present also in the whole range of physiological perfusion values, mostly due to the presence of

artefacts, highlighting at least one examination where an apparent perfusion peak would have been kept when clipping the colour scale manually.

6. Conclusion

The error detection strategy presented constitutes a gain in knowledge and fixes all problems just highlighted. In addition, an automated segmentation method offers several natural advantages over the classical manual approach, besides simple automation of the detection stage. First, it allows detecting the physical structures (i.e., vessels and bronchi) and the artefacts not perceived by radiologists, though affecting the reliability of computed perfusion values. Second, and most important, the automated error detection permits to highlight the correct extent of the different types of artefacts and exclude them from computation of perfusion values. Although missing structures, and sometimes artefacts, could not hamper the overall *qualitative* assessment of perfusion maps, they represent a relevant obstacle for *quantitative* assessment and, ultimately, for CTp standardization [33], and the TCC-based error analysis presented in this work should be carried out in every CTp study. The automatic error detection represents a methodological approach towards a more and more quantitative CTp imaging, this constituting an ineluctable way to achieve a routine clinical use of CTp.

Funding

The work has been jointly funded by the University of Bologna and IRST.

Acknowledgements

The authors would like to thank the whole staff of the Radiology Unit of IRST.

References

[1] G. Petralia, L. Preda, G. D'Andrea, S. Viotti, L. Bonello, R. De Filippi, M. Bellomi, CT perfusion in solid-body tumours. Part I: technical issues, Radiol. Med. 115 (2010) 843–857.
[2] A. Kambadakone, S.S. Yoon, T.-M. Kim, D.L. Karl, D.G. Duda, T.F. DeLaney, D.V. Sahani, CT perfusion as an imaging biomarker in monitoring response to neoadjuvant bevacizumab and radiation in soft-tissue sarcomas: comparison with tumor morphology, circulating and tumor biomarkers, and gene expression, Am. J. Roentgenol. 204 (1) (2015) W11–W18.
[3] H. Sun, Y. Xu, Q. Yang, W. Wang, Assessment of tumor grade and angiogenesis in colorectal cancer: whole-volume perfusion CT, Acad. Radiol. 21 (6) (2014) 750–757.
[4] K. Hayano, T. Fujishiro, D.V. Sahani, A. Satoh, T. Aoyagi, G. Ohira, T. Tochigi, H. Matsubara, K. Shuto, Computed tomography perfusion imaging as a potential imaging biomarker of colorectal cancer, World J. Gastroenterol. 20 (46) (2014) 17345–17351.
[5] N. Tacelli, T. Santangelo, A. Scherpereel, A. Duhamel, V. Deken, E. Klotz, A. Cortot, J.-J. Lafitte, F. Wallyn, J. Remy, M. Remy-Jardin, Perfusion CT allows prediction of therapy response in non-small cell lung cancer treated with conventional and anti-angiogenic chemotherapy, Eur. Radiol. 23 (8) (2013) 2127–2136.
[6] F. Fraioli, M. Anzidei, G. Serra, S. Liberali, A. Fiorelli, F. Zaccagna, et al., Whole-tumour CT-perfusion of unresectable lung cancer for the monitoring of anti-angiogenetic chemotherapy effects, Br. J. Radiol. 86 (1029) (2013) 20120174.
[7] T.D. Nguyen-Kim, T. Frauenfelder, K. Strobel, P. Veit-Haibach, M.W. Huellner, Assessment of bronchial and pulmonary blood supply in non-small cell lung cancer subtypes using computed tomography perfusion, Invest. Radiol. (2015).
[8] J.-B. Tylcza, K. El Alaoui-Lasmaili, E.-H. Djermoune, N. Thomas, B. Faivre, T. Bastogne, Data-driven modeling and characterization of anti-angiogenic molecule effects on tumoral vascular density, Biomed. Signal Proces. 20 (2015) 52–60.
[9] K.A. Miles, T.Y. Lee, V. Goh, E. Klotz, C. Cuenod, S. Bisdas, A.M. Groves, M.P. Hayball, R. Alonzi, T. Brunner, Current status and guidelines for the assessment of tumour vascular support with dynamic contrast-enhanced computed tomography, Eur. Radiol. 22 (2012) 1430–1441.
[10] T. Koh, C. Thng, S. Hartono, P. Lee, S. Choo, D. Poon, H. Toh, S. Bisdas, Dynamic contrast-enhanced CT imaging of hepatocellular carcinoma in cirrhosis: feasibility of a prolonged dual-phase imaging protocol with tracer kinetics modeling, Eur. Radiol. 19 (2009) 1184–1196.
[11] R. García-Figueiras, V.J. Goh, A.R. Padhani, S. Baleato-González, M. Garrido, L. León, A. Gómez-Caama no, CT perfusion in oncologic imaging: a useful tool? Am. J. Roentgenol. 200 (1) (2013) 8–19.
[12] J.F. Barret, N. Keat, Artefacts in CT: recognition and avoidance, Radiographics 24 (6) (2004) 1679–1691.
[13] L.W. Goldman, Principles of CT and CT technology, J. Nucl. Med. Technol. 35 (2007) 115–128.
[14] M.K. Fuld, A.F. Halaweish, S.E. Haynes, A.A. Divekar, J. Guo, E.A. Hoffman, Pulmonary perfused blood volume with dual-energy CT as surrogate for pulmonary perfusion assessed, Radiology 267 (3) (2013) 747–756.
[15] C.S. Ng, A.G. Chandler, W. Wei, D.H. Herron, E.F. Anderson, R. Kurzrock, C. Charnsangavej, Reproducibility of CT perfusion parameters in liver tumors and normal liver, Radiology 260 (2011) 762–770.
[16] M.W. Huellner, T.D. Collen, P. Gut, R. Winterhalder, C. Pauli, J. Diebold, B. Seifert, K. Strobel, P. Veit-Haibach, Multiparametric PET/CT-perfusion does not add significant additional information for initial staging in lung cancer compared with standard PET/CT, EJNMMI Research 4 (6) (2014) 1–13.
[17] Y.-W. Chen, H.-B. Pan, H.-H. Tseng, Y.-T. Hung, J.-S. Huang, C.-P. Chou, Assessment of blood flow in hepatocellular carcinoma: correlations of computed tomography perfusion imaging and circulating angiogenic factors, Int. J. Mol. Sci. 14 (9) (2013) 17536–17552.
[18] T.S. Koh, W. Shi, C.H. Thng, J.T. Ho, J.B. Khoo, D.L. Cheong, T.C. Lim, Assessment of tumor blood flow distribution by dynamic contrast-enhanced CT, IEEE Trans. Med. Imaging 32 (8) (2013) 1504–1514.
[19] H. Wang, Y. Cao, Correction of arterial input function in dynamic contrast-enhanced MRI of the liver, J. Magnetic Res. Imaging 36 (2) (2012) 411–421.
[20] A. Bevilacqua, D. Barone, S. Malavasi, G. Gavelli, Quantitative assessment of effects of motion compensation for liver and lung tumors in CT perfusion, Acad. Rad. 21 (11) (2014) 1416–1426.
[21] A. Gibaldi, D. Barone, G. Gavelli, S. Malavasi, A. Bevilacqua, Effects of guided random sampling of TCCs on blood flow values in CT perfusion studies of lung tumors, Acad. Rad. 22 (1) (2015) 58–69.
[22] N. Mullani, K. Gould, First-pass measurements of regional blood flow with external detectors, J. Nucl. Med. 24 (7) (1983) 577–581.
[23] J. Naish, D. McGrath, L. Bains, K. Passera, C. Roberts, Y. Watson, S. Cheung, M. Taylor, J. Logue, D. Buckley, J. Tessier, H. Young, J. Waterton, G. Parker, Comparison of dynamic contrast-enhanced MRI and dynamic contrast-enhanced CT biomarkers in bladder cancer, Magn. Reson. Med. 66 (1) (2011) 219–226.
[24] S. Goutelle, M. Maurin, F. Rougier, X. Barbaut, L. Bourguignon, M. Ducher, P. Maire, The Hill equation: a review of its capabilities in pharmacological modelling, Fundam. Clin. Pharmacol. 22 (6) (2008) 633–648.
[25] S. Wu, R. Ogden, J. Mann, R. Parsey, Optimal metabolite curve fitting for kinetic modeling of ^{11}C-WAY-100635, J. Nucl. Med. 48 (6) (2007) 926–931.
[26] D. Ippolito, C. Capraro, A. Casiraghi, C. Cestari, S. Sironi, Quantitative assessment of tumour associated neovascularisation in patients with liver cirrhosis and hepatocellular carcinoma: role of dynamic-CT perfusion imaging, Eur. Radiol 22 (4) (2012) 803–811.
[27] G.W. Zack, W.E. Rogers, S.A. Latt, Automatic measurement of sister chromatid exchange frequency, J. Histochem. Cytochem. 25 (7) (1977) 741–753.
[28] J. Zar, Biostatistical Analysis, 5th ed., Pearson Prentice-Hall, Upper Saddle River, NJ, 2010.
[29] A. Gherardi, S. Bravaccini, A. Bevilacqua, Computer assisted detection of regions of interest in histopathology using a hybrid supervised and unsupervised approach, in: Proc. SPIE: Digital Pathology Workshop in SPIE Medical Imaging, February 9–14, Lake Buena Vista, FL, USA, 8676:86760W-1-8, 2013.
[30] R.M. Subramaniam, A defining moment (Guest Editorial), Acad. Rad. 22 (1) (2015) 1–2.
[31] T.D. Nguyen-Kim, T. Frauenfelder, K. Strobel, P. Veit-Haibach, M.W. Huellner, Assessment of bronchial and pulmonary blood supply in non-small cell lung cancer subtypes using computed tomography perfusion, Invest. Radiol. 50 (3) (2015) 179–186.
[32] A. Kambadakone, S.S. Yoon, T.M. Kim, D.L. Karl, D.G. Duda, T.F. DeLaney, D.V. Sahani, CT perfusion as an imaging biomarker in monitoring response to neoadjuvant bevacizumab and radiation in soft-tissue sarcomas: comparison with tumor morphology, circulating and tumor biomarkers, and gene expression, Am. J. Roentgenol. 204 (1) (2015) W11–W18.
[33] R.G. Abramson, K.R. Burton, J.-P.J. Yu, et al., Methods and challenges in quantitative imaging biomarker development, Acad. Rad. 22 (1) (2015) 25–32.