



Influence of different measurement methods of arterial input function on quantitative dynamic contrast-enhanced MRI parameters in head and neck cancer

Wanxin Dong, Andreas Volk, Meriem Djaroum, Charly Girot, Corinne Balleyguier, Vincent Lebon, Gabriel Garcia, Samy Ammari, Stéphane Temam, Philippe Gorphe, et al.

► To cite this version:

Wanxin Dong, Andreas Volk, Meriem Djaroum, Charly Girot, Corinne Balleyguier, et al.. Influence of different measurement methods of arterial input function on quantitative dynamic contrast-enhanced MRI parameters in head and neck cancer. *Journal of Magnetic Resonance Imaging*, 2022, 2022, 10.1002/jmri.28486 . hal-03875359

HAL Id: hal-03875359

<https://hal.science/hal-03875359>

Submitted on 28 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

J. MAGN. RESON. IMAGING 2022.

View this article online at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com).

DOI: 10.1002/jmri.28486 Received Jul 12, 2022, Accepted for publication Oct 6, 2022

TITLE

Influence of different measurement methods of arterial input function on quantitative dynamic contrast enhanced MRI parameters in head and neck cancer.

*Wanxin Dong, MSc,1*Andreas Volk, PhD,1Meriem Djaroum, MS,1Charly Girot, PhD,1Corinne Balleyguier, MD, PhD,1,2Vincent Lebon, MD, PhD,1Gabriel Garcia, MD,2Samy Ammari, MD, PhD,1,2Stéphane Temam, MD, PhD,3Philippe Gorphe, MD,3Lecong Wei, MS,1Stéphanie Pitre-Champagnat, PhD,1Nathalie Lassau, MD, PhD,1,2andFrançois Bidault, MD, PhD1,2*

ABSTRACT

Background: Head and neck cancer (HNC) is the sixth most prevalent cancer worldwide. DCE-MRI helps in diagnosis and prognosis. Quantitative DCE-MRI requires an arterial input function (AIF), which affects the values of pharmacokinetic parameters (PKP).

Purpose: To evaluate influence of four individual AIF measurement methods on quantitative DCE-MRI parameters values (K^{trans} , v_e , k_{ep} , v_p), for HNC and muscle.

Study Type: Prospective

Population: 34 HNC patients (23 males, 11 females, age range 24-91)

Field Strength/Sequence: 3T; 3D SPGR gradient echo sequence with partial saturation of inflowing spins

Assessment: Four AIF methods were applied: automatic AIF (AIFa) with up to 50 voxels selected from the whole FOV, manual AIF (AIFm) with 4 voxels selected from the internal carotid artery, both conditions without (Mc-) or with (Mc+) motion correction. Comparison endpoints were peak AIF values, PKP values in tumor and muscle, and tumor/muscle PKP ratios.

Statistical Tests: Non-parametric Friedman test for multiple comparisons. Non-parametric Wilcoxon test, without and with Benjamini Hochberg correction, for pairwise comparison of AIF peak values and PKP values for tumor, muscle and tumor/muscle ratio, $p\text{-value} \leq 0.05$ was considered statistically significant.

Results: Peak AIF values differed significantly for all AIF methods, with mean AIFmMc+ peaks being up to 66.4% higher than those for AIFaMc+. Almost all PKP values were significantly higher for AIFa in both, tumor and muscle, up to 76% for mean K^{trans} values. Motion correction effect was smaller. Considering tumor/muscle parameter ratios, most differences were not significant ($0.068 \leq \text{Wilcoxon } p \text{ value} \leq 0.8$).

Data Conclusion: We observed important differences in PKP values when using either AIFa or AIFm, consequently choice of a standardized AIF method is mandatory for DCE-MRI on HNC. From the study findings, AIFm and inflow compensation are recommended. Use of the tumor/muscle PKP ratio should be of interest for multicenter studies.

Keywords: DCE-MRI, head and neck cancer, arterial input function (AIF), pharmacokinetic parameters (PKP), AIF measurement methods

INTRODUCTION

Head and neck cancer (HNC) is the sixth most prevalent cancer worldwide with 890,000 new cases and 450,000 deaths in 2018¹. The major histological type of HNC tumors is squamous cell carcinoma (HNSCC). Based on Global Cancer Observatory (GLOBOCAN) 2018's global cancer incidence and mortality estimating report, Johnson et al found that HNC continues to rise and is anticipated to increase by 30%, that is, 1.08 million new cases annually, by 2030¹.

Medical imaging complements the physical examination and plays an important role in tumor staging and treatment selection (NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Head and Neck Cancers Version 2.2022, https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf). Several functional imaging modalities have been described to predict and assess response to treatment, Dynamic Contrast Enhanced MRI (DCE-MRI) being one such method^{2,3}.

DCE-MRI is a minimally invasive functional imaging technique based on high temporal resolution dynamic acquisition of T1 weighed images to follow signal change during the passage of the contrast agent (CA) through the vasculature and into the tissues,

typically a gadolinium chelate injected intravenously. DCE-MRI allows the characterization of tissue micro-vasculature. Analysis of the signal time course using pharmacokinetic modeling, typically the extended Tofts model, provides quantitative maps of pharmacokinetic parameters (PKP): volume transfer constant between blood plasma and extravascular extracellular space (EES) K^{trans} (min^{-1}), EES fractional volume v_e (dimensionless or %), rate constant between EES and blood plasma k_{ep} (min^{-1}), and fractional blood plasma volume v_p (dimensionless or %)⁴⁻⁶.

Quantitative DCE-MRI requires an arterial input function (AIF) which describes the concentration of CA in the vascular space as a function of time, and which affects the PKP values. There are several ways to measure an AIF, each with its pros and cons. An individual AIF is an AIF extracted from each patient's images, whereas a population AIF is an average AIF generated from individual AIF measured on a group of patients and applied to the individual patient. The Quantitative Imaging Biomarkers Alliance (QIBA, <https://www.rsna.org/research/quantitative-imaging-biomarkers-alliance>) recommends that AIF should be measured individually whenever possible. Measuring an individual AIF is not trivial, it depends on the acquisition sequence, on post-processing of the images, also on location and number of voxels used for AIF measurement⁶.

Individual AIF measurements can be automatic or manual⁶. Previous studies in the field of oncologic DCE-MRI have focused on comparisons between different AIF types and their effects on PKP, mostly in the field of prostate DCE-MRI⁷⁻⁹. For example, Ziayee et al compared manual individual AIF, automatic individual AIF, and population-averaged AIF methods for prostate DCE-MRI⁷. Sanz-Requena et al compared automatic principal component analysis (PCA)-based AIF, manual AIF, and population-averaged AIF for prostate DCE-MRI examinations which included tumors

and controls⁸. Meng et al compared individual manual AIF, most commonly used population-average AIF, and double exponential population average AIF for prostate cancer⁹. Furthermore, in the field of peripheral artery disease, Li et al compared AIF resulting from different automatic and manual methods in human calf muscle, however without considering resulting PKP¹⁰.

In the field of head and neck cancer, the AIF is often measured manually in nearby major blood vessels within the vicinity of the tumor, for instance, the carotid arteries are commonly used^{6,11}. Examinations are typically performed on axial slices, and inflow effects may result in underestimation of AIF amplitudes^{12–16}. Authors have proposed solutions to reduce inflow effect. Yuan et al proposed vein concentration-time-curve methods to help to compensate for arterial in-flow effect and reduce kinetic parameter estimation error and inconsistency for DCE-MRI in head and neck¹⁶. Han et al proposed to reduce the inflow effect by using pre-saturation of blood flow located upstream of the imaged anatomical region¹⁷.

In the field of head and neck studies, researchers have compared individual AIF methods and population AIF method: Onxley et al found significant differences between individual side-specific AIF and population-averaged AIF¹⁸. Koopman et al compared PKP obtained with AIF from different cerebral arteries¹¹. In addition, muscular PKP were compared using manual AIF and population averaged AIF.

Manual AIF selection is time consuming and potentially prone to interobserver variability. Commercial software packages also comprise algorithms for fully automated AIF voxel selections. However, according to QIBA: "these possibly need some adjustments", depending on organ and sequence (QIBA Profile: DCE MRI Quantification (DCEMRI-Q), https://qibawiki.rsna.org/images/1/12/DCE-MRI_Quantification_Profile_v1.0.pdf).

The aim of this study was to investigate and quantify, using a commercial software package, the influence of the AIF measurement method (fully automatic or manual, with or without motion correction) on the values of quantitative DCE-MRI PKP, for HNSCC as well as for facial muscles as an ubiquitous potential reference tissue.

MATERIALS AND METHODS

Patients And Clinical Protocol

The study conformed to the Ethical Guidelines of the Declaration of Helsinki, all subjects provided written informed consent. It was approved by the national review board for biomedical research on human beings. Inclusion criteria were: adult presenting HNSCC with indication to perform a ^{18}F -fluorodeoxyglucose positron emission tomography (FDG-PET), intend for nodal surgery. Main exclusion criteria were: contraindication to FDG-PET or MRI examination.

This work included 34 patients (23 males (age: 63 ± 14), 11 females (age: 67 ± 12)) who underwent MRI between September 2018 and June 2020. All head and neck malignancies had histological evidence of squamous cell carcinoma. Patients had no surgery and treatment prior to imaging.

MRI Protocol

Examinations were performed on a 3 Tesla PET/MRI system (SignaTM, GE Healthcare, Milwaukee, WI, USA, antenna Head+Neck 40). The DCE-MRI acquisition was an integral part of a multimodal imaging protocol comprising also morphological imaging, diffusion weighted imaging, and FDG-PET.

The DCE examination was performed with injection of gadoteric acid (DOTAREM, Guerbet, Villepinte, France) at a dose of 0.1 mmol/kg 25 sec after acquisition start, and

with a flow rate of 2 mL/sec. The DCE-MRI acquisition was based on a 3-dimensional gradient echo sequence of the spoiled gradient recalled (SPGR) type (GE protocol name: Liver Acquisition with Volume Acquisition (LAVA)). The main characteristics of the DCE-MRI sequence were: axial slab plane; Echo Time (TE) = 1.4 ms; Repetition Time (TR) = 5.2 ms; flip angle = 30°; number of slices per phase = 8; duration of one phase = 3 sec; slice thickness = 5 mm; FOV (field of view) = 26x26 cm; matrix = 160x160 pixels; total duration = 390 seconds (corresponding to 130 phases), auto-calibrating reconstruction for cartesian sampling (ARC) acceleration method. A 6 cm thick saturation slab was positioned parallel and adjacent to the inferior side of the imaging slab, to compensate for the inflow effect.

Image Processing

Analysis of the DCE-MRI data was carried out using Olea Sphere® software (version. 3.0.18, Olea Medical Solutions, La Ciotat, France), specifically, the "MR Permeability" module. This software used the extended Tofts model to calculate the parametric maps K^{trans} , v_e , k_{ep} and v_p . Haematocrits were set to: large vessel hematocrit $hct_{LV} = 0.45$, small vessel hematocrit $hct_{SV} = 0.25$. For analysis, the T1 pre-contrast values at 3T were set to $T1_{blood} = 1650$ ms, $T1_{tissue} = 1400$ ms, and DOTAREM $r1$ relaxivity = 3.50 ($mM^{-1} \cdot s^{-1}$)^{19–21}.

Four different conditions of AIF selection have been performed: 1) Automatic selection of AIF voxels without motion correction, AIFaMc-, 2) Automatic selection of AIF voxels with motion correction, AIFaMc+, 3) Manual selection of AIF voxels without motion correction, AIFmMc-, 4) Manual selection of AIF voxels with motion correction, AIFmMc+.

The automatic AIF (AIFa) mode implemented in Olea Sphere[®] was based on the method described by Mouridsen et al and selected up to 50 voxels from the entire imaged volume (3D FOV) without user intervention²². For the manual AIF (AIFm) method, 4 voxels were selected on a single slice in an internal carotid artery (ICA) by a radiologist (FB) with 19 years head and neck MRI experience. AIFm voxel location was double checked by a second radiologist (SA) with 14 years' experience. The motion correction method implemented in Olea Sphere[®] software corresponded to a pairwise in-plane (acquisition plane) rigid co-registration of all DCE images. The tissue VOIs were manually outlined for the tumor and for a masseter muscle on images acquired by DCE-MRI. The VOIs were saved and applied to the calculated parametric maps in order to measure values of the quantitative parameters for both tissues (Fig. 1).

AIF was characterized by its peak amplitude corresponding to the concentration peak at the first passage of the CA (Fig. 2).

Statistics

Exclusion criteria were applied to patient data in order to carry out analysis on a population as homogenous as possible. Artifacts (swallowing during acquisition, dental artifacts), absence of ICA within the 3D FOV, tumor not visible within the 3D FOV, or absence of masseter muscle within the 3D FOV led to exclusion. Two head and neck MRI experienced radiologists (FB 19 years, SA 14 years) reached consensus on these exclusion criteria.

Statistical analysis was performed using R version 4.0.4. Results were presented as mean values with standard deviation (table), as median values, first and third quartiles (boxplots). For the 4 AIF measurement methods, we first performed a non-parametric

Friedman test for multiple comparisons of dependent samples. Then we performed non-parametric Wilcoxon paired-samples tests (AlFaMc- vs. AlFaMc+, AlFmMc- vs. AlFmMc+, AlFaMc- vs. AlFmMc-, AlFaMc+ vs. AlFmMc+), without and with Benjamini Hochberg (BH) correction for multiple tests, to compare AIF peak values and quantitative parameter values (K^{trans} , v_e , k_{ep} , v_p) for tumor, muscle, and tumor/muscle ratios. “Crossed comparisons” (AlFaMc+ vs. AlFmMc-; AlFaMc- vs. AlFmMc+) were not under consideration in our pairwise comparison study plan as these conditions combined the variation of two experimental conditions. P-value ≤ 0.05 was considered the threshold for significant difference.

RESULTS

Artifacts or tumor location not visible in the 3D FOV led to exclusion of 8 patients. More precisely, 3 tumors in oral cavity were not visible on MRI, 3 tumors in oropharynx (uvula) had swallowing artifacts, and 2 tumors in oral cavity had dental artifacts.

Among the remaining 26 patients, voxels for AIFm could be defined inside the internal carotid artery for 24 patients. For two patients, AIFm voxels could be only defined inside the common carotid artery, as the tumor was located in the inferior part of the head and neck region. These patients were excluded from the final analysis.

Furthermore, among the remaining 24 patients, muscle VOIs could be defined inside a masseter muscle for 22 patients. For the other 2 patients, the muscle VOIs could only be defined inside a levator scapulae muscle, and these patients were excluded from the final analysis.

After these exclusions, the same analysis protocol was performed on a total of 22 patients.

Variability Of AIF Peaks

The peaks for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following significant differences: AIFaMc- vs. AIFaMc+, AIFmMc- vs. AIFmMc+, AIFaMc- vs. AIFmMc-, and AIFaMc+ vs. AIFmMc+ (Fig. 2 and Fig. 3). Differences remained significant after BH correction. The mean peak values of AIFaMc-, AIFaMc+, AIFmMc-, AIFmMc+, were respectively (3.09 ± 0.90) mM, (2.62 ± 0.82) mM, (4.53 ± 0.93) mM, (4.36 ± 0.86) mM. AIFa versus AIFm presented the largest difference in peak mean value. The mean peak value of AIFmMc+ was 66.4% higher than that of AIFaMc+.

Variability Of PKP

All mean PKP values ($\pm$ standard deviation) for the 22 patients are summarized in Table 1.

Two-by-two comparisons according to the AIF method, for parameters K^{trans} and v_e , for the two tissues (tumor and muscle), as well as for the parameter ratios (tumor/muscle) are presented in Fig. 4 to Fig. 9. The corresponding figures for parameters k_{ep} and v_p are available as supplementary data. Significance levels are indicated on the figures as follows: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***), $p > 0.05$ (ns).

K^{trans} Values

For tumor, the K^{trans} values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided significant differences under the different AIF conditions, except for the AIFmMc- vs. AIFmMc+ comparison ($p = 0.059$) (Fig. 4). Significance status remained the same after BH correction. Mean value ratios

for significant differences were: $K^{\text{trans}}(\text{AIFaMc+})/K^{\text{trans}}(\text{AIFaMc-}) = 1.17$; $K^{\text{trans}}(\text{AIFaMc-})/K^{\text{trans}}(\text{AIFmMc-}) = 1.55$; $K^{\text{trans}}(\text{AIFaMc+})/K^{\text{trans}}(\text{AIFmMc+}) = 1.76$.

For muscle, the K^{trans} values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test showed significant differences only for automatic versus manual AIF, AIFaMc- vs. AIFmMc- and AIFaMc+ vs. AIFmMc+. Differences were not significant for AIFaMc- vs. AIFaMc+ ($p = 0.14$) and for AIFmMc- vs. AIFmMc+ ($p = 0.063$) (Fig. 5). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $K^{\text{trans}}(\text{AIFaMc-})/K^{\text{trans}}(\text{AIFmMc-}) = 1.53$; $K^{\text{trans}}(\text{AIFaMc+})/K^{\text{trans}}(\text{AIFmMc+}) = 1.52$.

For tumor/muscle ratio, the K^{trans} values for the four AIF methods were not significantly different from each other according to Friedman test ($p = 0.074$). Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.61$), AIFmMc- vs. AIFmMc+ ($p = 0.16$), AIFaMc+ vs. AIFmMc+ ($p = 0.13$) (Fig. 6). Tumor/muscle ratio K^{trans} showed significant difference only for AIFaMc- vs. AIFmMc-. After BH correction, this difference was no longer significant ($p = 0.082$). Mean value ratio for significant difference was: $K^{\text{trans}}(\text{AIFaMc-})/K^{\text{trans}}(\text{AIFmMc-}) = 1.16$.

v_e Values

For tumor, the v_e values for all four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.092$), AIFmMc- vs. AIFmMc+ ($p = 0.48$), AIFaMc- vs. AIFmMc- ($p = 0.054$), AIFaMc+ vs. AIFmMc+ ($p = 0.054$) (Fig. 7). However, we noticed that Wilcoxon test provided a significant difference for the “crossed” comparison of AIFaMc+ vs. AIFmMc-, which was not considered in this study (cf. Materials and Methods).

For muscle, the v_e values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.074$), and AIFmMc- vs. AIFmMc+ ($p = 0.48$) (Fig. 8). Muscle v_e showed significant differences only for automatic versus manual AIF, AIFaMc- vs. AIFmMc- and AIFaMc+ vs. AIFmMc+. Significance status remained the same after BH correction. Mean value ratios for significant differences were: $v_e(\text{AIFaMc-})/v_e(\text{AIFmMc-}) = 1.20$; $v_e(\text{AIFaMc+})/v_e(\text{AIFmMc+}) = 1.23$.

For tumor/muscle ratio, the v_e values for the four AIF methods were not significantly different from each other according to Friedman test ($p = 0.26$). Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.14$), AIFmMc- vs. AIFmMc+ ($p = 0.75$), AIFaMc- vs. AIFmMc- ($p = 0.074$), AIFaMc+ vs. AIFmMc+ ($p = 0.13$) (Fig. 9).

k_{ep} Values

For tumor, the k_{ep} values for all four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.18$), AIFmMc- vs. AIFmMc+ ($p = 0.11$). Tumor k_{ep} showed significant differences only for automatic versus manual AIF, AIFaMc- vs. AIFmMc- and AIFaMc+ vs. AIFmMc+ (Supplementary Material Fig. 1). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $k_{ep}(\text{AIFaMc-})/k_{ep}(\text{AIFmMc-}) = 1.37$; $k_{ep}(\text{AIFaMc+})/k_{ep}(\text{AIFmMc+}) = 1.47$.

For muscle, the k_{ep} values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.52$), AIFmMc- vs. AIFmMc+ ($p = 0.85$).

Muscle k_{ep} showed significant differences only for automatic versus manual AIF, AIFaMc- vs. AIFmMc- and AIFaMc+ vs. AIFmMc+ (Supplementary Material Fig. 2). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $k_{ep}(\text{AIFaMc-})/k_{ep}(\text{AIFmMc-}) = 1.25$; $k_{ep}(\text{AIFaMc+})/k_{ep}(\text{AIFmMc+}) = 1.29$.

For tumor/muscle ratio, the k_{ep} values for the four AIF methods were not significantly different from each other according to Friedman test ($p = 0.055$). Wilcoxon test provided the following non-significant differences: AIFaMc- vs. AIFaMc+ ($p = 0.19$), AIFaMc+ vs. AIFmMc+ ($p = 0.068$). Tumor/muscle ratio k_{ep} showed significant difference for AIFmMc- vs. AIFmMc+ and for AIFaMc- vs. AIFmMc comparisons (Supplementary Material Fig. 3). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $k_{ep}(\text{AIFmMc+})/k_{ep}(\text{AIFmMc-}) = 1.05$; $k_{ep}(\text{AIFaMc-})/k_{ep}(\text{AIFmMc-}) = 1.13$.

v_p Values

For tumor, the v_p values for all four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following significant differences: AIFaMc- vs. AIFaMc+, AIFaMc- vs. AIFmMc-, and AIFaMc+ vs. AIFmMc+. Tumor v_p showed non-significant difference only for AIFmMc- vs. AIFmMc+ comparison ($p = 0.44$) (Supplementary Material Fig. 4). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $v_p(\text{AIFaMc+})/v_p(\text{AIFaMc-}) = 1.35$; $v_p(\text{AIFaMc-})/v_p(\text{AIFmMc-}) = 1.55$; $v_p(\text{AIFaMc+})/v_p(\text{AIFmMc+}) = 2.09$.

For muscle, the v_p values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following significant differences: AIFaMc- vs. AIFaMc+, AIFaMc- vs. AIFmMc-, and AIFaMc+ vs. AIFmMc+. Muscle v_p

showed non-significant difference only for the AIFmMc- vs. AIFmMc+ comparison ($p = 0.054$) (Supplementary Material Fig. 5). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $v_p(\text{AIFaMc+})/v_p(\text{AIFaMc-}) = 1.50$; $v_p(\text{AIFaMc-})/v_p(\text{AIFmMc-}) = 2.00$; $v_p(\text{AIFaMc+})/v_p(\text{AIFmMc+}) = 3.00$.

For tumor/muscle ratio, the v_p values for the four AIF methods showed significant differences according to Friedman test. Wilcoxon test provided the following significant differences: AIFmMc- vs. AIFmMc+, AIFaMc- vs. AIFmMc-, and AIFaMc+ vs. AIFmMc+. Tumor/muscle ratio v_p showed non-significant difference only for AIFaMc- vs. AIFaMc+ comparison ($p = 0.8$) (Supplementary Material Fig. 6). Significance status remained the same after BH correction. Mean value ratios for significant differences were: $v_p(\text{AIFmMc+})/v_p(\text{AIFmMc-}) = 1.16$; $v_p(\text{AIFaMc-})/v_p(\text{AIFmMc-}) = 0.67$; $v_p(\text{AIFaMc+})/v_p(\text{AIFmMc+}) = 0.53$.

DISCUSSION

This study compared four methods of measuring AIF and their effects on the PKP K^{trans} , v_e , k_{ep} , v_p in a clinical setting of characterization of head and neck cancers by DCE-MRI. We noticed that v_p relative standard errors were very high in muscle tissues, probably due to low mean values. Furthermore, muscle PKP reported in the literature are limited to K^{trans} , v_e , and k_{ep} ^{11,23–25}. Consequently, this discussion only refers to K^{trans} , v_e , and k_{ep} .

Comparing a fully automatically generated AIF (AIFa) with an AIF generated after manual voxel selection (AIFm) in the internal carotid artery, we observed significantly higher mean AIF peak values for AIFm compared with AIFa. Consequently, most of the PKP measured in tumor or muscle regions using AIFa were significantly higher

than those measured using AIFm. The relationship between AIF peak variations and PKP values, that is underestimation of AIF peak leading to overestimation of parameter values, is expected from theory, and has also been demonstrated by simulations²⁶.

The significant difference in AIF peaks between different methods in our study may be explained by the method of AIF voxel selection and the correction of the inflow effect. For manual voxel selection, a small number of voxels were placed in the center of a large-caliber cervical carotid artery. Thus, the voxels can be considered to represent the arterial intravascular space with a minimum of partial volume. In addition, potential underestimation of the AIF peak due to inflow effects is partly compensated by the use of the presaturation pulse. The automatic voxel selection method implemented in Olea Sphere[®] selects up to 50 AIF voxels from the entire imaged volume, and the final AIFa is the average of the signals from all voxels. On visual inspection of the location of the voxels we observed that not all are located in the center of large-caliber cervical arteries, which could lead to a partial volume phenomenon on the voxels and a lower mean value. A semiautomatic AIF measurement mode with operator restriction of the AIFa voxel search to regions containing large caliber cervical arteries could be a compromise between AIFa and AIFm, but would also need additional operator time. Indeed, studies on prostate by Ziayee et al and Sanz-Requena et al showed similar AIFa and AIFm, probably due to the fact that, at that location, AIFa and AIFm voxels are all located inside large caliber vessels^{7,8}.

We also studied the effect of motion correction on AIF. Significant differences were observed for the AIF peaks (AIFMc- vs. AIFMc+), but they were much smaller than for AIFa vs. AIFm. Corresponding PKP were not significantly different except for tumor $K^{trans}(AIFaMc+)$ vs. $K^{trans}(AIFaMc-)$. Thus, AIFm seems to be less influenced by motion than AIFa and should be preferred.

A general limitation for DCE-MRI derived PKP is the absence of a gold standard to compare with.

DCE-MRI derived PKP have been studied previously in the head and neck region and other skeletal muscles, however comparison of absolute values is made difficult due to non-standard acquisition and analysis protocols.

Koopman et al studied head and neck muscles, with individual AIF measured in different cerebral arteries¹¹. They found higher muscle PKP values than ours when using an AIF from the internal carotid artery ($v_e = 0.74$, $K^{trans} = 0.21$). Explanations may be the absence of inflow effect compensation in their study leading to lower AIF peak values, the use of the standard Tofts model, and the use of a higher blood T1 baseline value (1932 ms). However, as in our study, they found lower parameter values for higher AIF peaks.

Chen et al studied head and neck tumors and muscles (longus capitis muscle) using the standard Tofts model²³. Their values for muscle PKP were: $K^{trans} = (0.160 \pm 0.088)$ min^{-1} , $v_e = (0.229 \pm 0.147)$. Values are in the same range as ours, however, in their study they mixed up measurements with manual AIF in a carotid artery and those with population AIF. Furthermore, compensation of inflow effects on AIF was not mentioned in this study. As in this study, they observed much higher parameter values in tumors than in muscle tissue.

Padhani et al studied various pelvic muscles, using a standard synthetic input function and the standard Tofts model²⁴. Their K^{trans} and v_e values were close to ours, despite the differences in anatomic locations and measurement methods: K^{trans} range from 0.083 to 0.245 min^{-1} (with standard deviations range from 55% to 84%), v_e range from 0.10 to 0.14 (with standard deviations range from 19% to 30%).

Tumor PKP values are even more difficult to be compared because of the heterogeneity of the tumors. Nevertheless, it can be noticed that mean values for primary head and neck tumors published by Chen et al, $K^{trans} = (0.251 \pm 0.066) \text{ min}^{-1}$, $v_e = (0.344 \pm 0.081)$, are close to our values with AIFm.

In our study, the effect of AIF type was similar for tumors and for muscle tissue. Indeed, when considering the tumor/muscle ratios for K^{trans} , v_e , and k_{ep} no significant differences were detected by the Friedman test. Pairwise Wilcoxon tests detected significant differences only for k_{ep} , for AIFmMc- vs. AIFmMc+ and AIFaMc- vs. AIFmMc-, with 13% maximum difference. PKP ratios from tumor and normal tissue are not frequently used to analyze tumors in DCE-MRI. Sureka et al successfully used K^{trans} and k_{ep} ratios in tumor and normal prostate tissue to distinguish prostate cancer from chronic prostatitis²⁷.

Results of our study suggest that important variations of AIF measurement leading to similar variations of tissue parameter values may be at least partly compensated by considering parameter ratios from tumor and muscle tissue. As facial muscle tissue is always present in a typical HNC FOV, this relatively simple procedure may be useful for multicenter retrospective or prospective studies of HNC cancer when DCE-MRI data acquisition protocols have not been or cannot be standardized (e.g., protocols with or without inflow effect compensation, different time resolutions), or when erroneous AIF is suspected (patient motion, noise). A prerequisite for the procedure is that facial muscle perfusion and microvascular characteristics are stable enough to capture information provided by tumor PKP values. El Rafei et al reported high variability of PKP in muscles around the hip joint, and they cautioned about using muscle tissue as a reference tissue²⁵. However, besides the different muscle location, their DCE-MRI acquisition had relatively low temporal resolution (13.5 sec), and they

did not report any inflow effect compensation procedure, which may both contribute to AIF and tissue parameter variability. According to their findings, the use of free-form VOIs matching the muscle region reduced variability, as we did in our study. It is of notice that coefficients of variation (CV) observed in our study in masseter muscles with AIFmMc+ ($CV(K^{trans}) = 72\%$, $CV(v_e) = 30\%$, $CV(k_{ep}) = 51\%$) were about half of mean CV reported in their study ($CV(K^{trans}) = 128\%$, $CV(v_e) = 61\%$, $CV(k_{ep}) = 90\%$). Further studies will be necessary to assess the variability and repeatability of facial muscle PKP, and to explore the usefulness of parameter ratios for HNC characterization and therapy follow-up.

Limitations

Only one radiologist selected the AIFm voxels and drew muscle and tumor VOI, though double-checked by a second radiologist. Furthermore, we used a single software for post-processing of the data and generation of PKP maps. Future studies should address inter-observer variability for the manual AIF definition, as well as compare different software for automatic AIF definition in the head and neck region, in order to define an optimized standardized method for AIF measurement. The performance of an optimized protocol should then be assessed on a bigger patient population by multicenter studies with different MRI systems. In addition, repeatability of the measurements on muscle tissue should be performed on the same patient in order to assess intra-patient variability of facial muscle PKP.

We used fixed predefined precontrast T1 values for blood and tissue, instead of values from T1 mapping. As T1 mapping, typically performed using variable flip angle SPGR sequences, may introduce additional variability, we preferred in this study, focused on effects of AIF variability, the use of fixed precontrast T1 values matching published or

measured values of the target tissues. Nevertheless, an optimized DCE-MRI protocol should comprise precontrast T1 mapping with an optimized and standardized mapping sequence^{28,29}.

Conclusion

This work identifies the importance of the choice of the AIF in DCE-MRI, demonstrating its influence on the calculation of PKP, and stressing the need for standardized data acquisition and processing protocols for head and neck cancer DCE-MRI. The findings may support the use of manual AIF in internal carotid artery, inflow effect compensation, and the use of tumor/muscle ratios as an additional parameter less sensitive to AIF errors, particularly for multicenter studies or longitudinal follow-up of patients. Future studies should allow to further optimize this protocol.

REFERENCES

1. Johnson DE, Burtneess B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Primer 2020;6(1):92.

2. Mui AWL, Lee AWM, Lee VHF, et al. Prognostic and therapeutic evaluation of nasopharyngeal carcinoma by dynamic contrast-enhanced (DCE), diffusion-weighted (DW) magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS). *Magn Reson Imaging* 2021;83:50–56.
3. Gaddikeri S, Gaddikeri RS, Tailor T, Anzai Y. Dynamic Contrast-Enhanced MR Imaging in Head and Neck Cancer: Techniques and Clinical Applications. *Am J Neuroradiol* 2016;37(4):588–595.
4. Cuenod CA, Balvay D. Perfusion and vascular permeability: Basic concepts and measurement in DCE-CT and DCE-MRI. *Diagn Interv Imaging* 2013;94(12):1187–1204.
5. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced t1-weighted MRI of a diffusable tracer: Standardized quantities and symbols. *J Magn Reson Imaging* 1999;10(3):223–232.
6. Shukla-Dave A, Obuchowski NA, Chenevert TL, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging* 2019;49(7):e101–e121.
7. Ziayee F, Müller-Lutz A, Gross J, et al. Influence of arterial input function (AIF) on quantitative prostate dynamic contrast-enhanced (DCE) MRI and zonal prostate anatomy. *Magn Reson Imaging* 2018;53:28–33.
8. Sanz-Requena R, Prats-Montalbán JM, Martí-Bonmatí L, et al. Automatic individual arterial input functions calculated from PCA outperform manual and population-averaged approaches for the pharmacokinetic modeling of DCE-MR images: PCA-Based Automatic Individual AIF. *J Magn Reson Imaging* 2015;42(2):477–487.
9. Meng R, Chang SD, Jones EC, Goldenberg SL, Kozlowski P. Comparison between Population Average and Experimentally Measured Arterial Input Function in Predicting Biopsy Results in Prostate Cancer. *Acad Radiol* 2010;17(4):520–525.
10. Li X, Conlin CC, Decker ST, et al. Sampling arterial input function (AIF) from peripheral arteries: Comparison of a temporospatial-feature based method against conventional manual method. *Magn Reson Imaging* 2019;57:118–123.
11. Koopman T, Martens RM, Lavini C, et al. Repeatability of arterial input functions and kinetic parameters in muscle obtained by dynamic contrast enhanced MR imaging of the head and neck. *Magn Reson Imaging* 2020;68:1–8.
12. Nazarpour M. Effect of concentration of contrast agent on the inflow effect for measuring absolute perfusion by use of inversion recovery T1-weighted TurboFLASH images. *Radiol Phys Technol* 2012;5(1):86–91.
13. Peeters F, Annet L, Hermoye L, Van Beers BE. Inflow correction of hepatic perfusion measurements using T1-weighted, fast gradient-echo, contrast-enhanced MRI. *Magn Reson Med* 2004;51(4):710–717.

14. Ivancevic MK, Zimine I, Montet X, et al. Inflow effect correction in fast gradient-echo perfusion imaging: Inflow Effect Correction. *Magn Reson Med* 2003;50(5):885–891.
15. Keil VC, Mädler B, Gieseke J, et al. Effects of arterial input function selection on kinetic parameters in brain dynamic contrast-enhanced MRI. *Magn Reson Imaging* 2017;40:83–90.
16. Yuan J, Chow SKK, Zhang Q, Yeung DKW, Ahuja AT, King AD. The Use of Dynamic Tracer Concentration in Veins for Quantitative DCE-MRI Kinetic Analysis in Head and Neck. *PLoS ONE* 2013;8(3):e59885.
17. Han M, Hargreaves BA. Reduction of flow artifacts by using partial saturation in RF-spoiled gradient-echo imaging: Flow Artifact Reduction by Partial Saturation. *Magn Reson Med* 2011;65(5):1326–1334.
18. Onxley JD, Yoo DS, Muradyan N, MacFall JR, Brizel DM, Craciunescu OI. Comprehensive Population-Averaged Arterial Input Function for Dynamic Contrast-Enhanced vMagnetic Resonance Imaging of Head and Neck Cancer. *Int J Radiat Oncol* 2014;89(3):658–665.
19. Lu H, Clingman C, Golay X, Zijl PCM van. Determining the longitudinal relaxation time (T₁) of blood at 3.0 Tesla. *Magn Reson Med* 2004;52(3):679–682.
20. Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3 T? *Magn Reson Imaging* 2017;35:69–80.
21. Shen Y, Goerner FL, Snyder C, et al. T₁ Relaxivities of Gadolinium-Based Magnetic Resonance Contrast Agents in Human Whole Blood at 1.5, 3, and 7 T: *Invest Radiol* 2015;50(5):330–338.
22. Mouridsen K, Christensen S, Gyldensted L, Østergaard L. Automatic selection of arterial input function using cluster analysis. *Magn Reson Med* 2006;55(3):524–531.
23. Chen L, Ye Y, Chen H, et al. Dynamic Contrast-Enhanced Magnetic Resonance Imaging for Differentiating Between Primary Tumor, Metastatic Node and Normal Tissue in Head and Neck Cancer. *Curr Med Imaging Rev* 2018;14(3):416–421.
24. Padhani AR, Hayes C, Landau S, Leach MO. Reproducibility of quantitative dynamic MRI of normal human tissues. *NMR Biomed* 2002;15(2):143–153.
25. El Rafei M, Teixeira P, Norberciak L, Badr S, Cotten A, Budzik J-F. Dynamic contrast-enhanced MRI perfusion of normal muscle in adult hips: Variation of permeability and semi-quantitative parameters. *Eur J Radiol* 2018;108:92–98.
26. Lavini C. Simulating the effect of input errors on the accuracy of Tofts' pharmacokinetic model parameters. *Magn Reson Imaging* 2015;33(2):222–235.
27. Sureka B, Elhence P, Khera PS, et al. Quantitative contrast-enhanced perfusion kinetics in multiparametric MRI in differentiating prostate cancer from chronic prostatitis: results from a pilot study. *Br J Radiol* 2019;92(1100):20190181.

28. Stikov N, Boudreau M, Levesque IR, Tardif CL, Barral JK, Pike GB. On the accuracy of T_1 mapping: Searching for common ground: Accuracy of T_1 Mapping. *Magn Reson Med* 2015;73(2):514–522.
29. Heye T, Boll DT, Reiner CS, Bashir MR, Dale BM, Merkle EM. Impact of precontrast T_{10} relaxation times on dynamic contrast-enhanced MRI pharmacokinetic parameters: T_{10} mapping versus a fixed T_{10} reference value: Impact of T_{10} on DCE-MRI. *J Magn Reson Imaging* 2014;39(5):1136–1145.

TABLE 1. Pharmacokinetic parameters for 4 AIF Methods

	K^{trans} (min^{-1})	v_e (dimensionless)	k_{ep} (min^{-1})	v_p (dimensionless)
Tumor				
AlFaMc-	0.589 ± 0.336	0.443 ± 0.168	1.497 ± 0.872	0.034 ± 0.019
AlFaMc+	0.688 ± 0.452	0.464 ± 0.164	1.631 ± 1.104	0.046 ± 0.025
AlFmMc-	0.380 ± 0.190	0.391 ± 0.146	1.090 ± 0.406	0.022 ± 0.012
AlFmMc+	0.392 ± 0.188	0.392 ± 0.152	1.112 ± 0.337	0.022 ± 0.013

Muscle

AlFaMc-	0.158 ± 0.154	0.208 ± 0.085	0.774 ± 0.473	0.004 ± 0.004
AlFaMc+	0.161 ± 0.134	0.216 ± 0.071	0.794 ± 0.518	0.006 ± 0.006
AlFmMc-	0.103 ± 0.071	0.174 ± 0.050	0.620 ± 0.288	0.002 ± 0.003
AlFmMc+	0.106 ± 0.080	0.176 ± 0.053	0.614 ± 0.320	0.002 ± 0.003
Tumor/Muscle	(Ratio)			
AlFa Mc-	5.665 ± 4.611	2.258 ± 0.654	2.283 ± 1.337	16.990 ± 15.788
AlFa Mc+	5.904 ± 5.091	2.240 ± 0.683	2.450 ± 1.644	15.683 ± 14.833
AlFm Mc-	4.871 ± 3.243	2.299 ± 0.656	2.027 ± 1.001	25.477 ± 23.396
AlFm Mc+	4.867 ± 2.967	2.279 ± 0.673	2.132 ± 1.022	29.666 ± 28.036

Data are reported as mean ± standard deviation.

AlFaMc- = automatic arterial input function no motion correction.

AlFaMc+ = automatic arterial input function with motion correction.

AlFmMc- = manual arterial input function no motion correction.

AlFmMc+ = manual arterial input function with motion correction.

FIGURE CAPTIONS

FIGURE 1: Example of VOI for a patient. Tumor VOI (yellow, volume = 5.85 cc, location: base of the tongue), muscle VOI (red, volume = 12.24 cc, location: masseter). Images

represent, from left to right and from top to bottom, the 8 slices acquired by DCE-MRI, at the post-injection time 54.24 sec.

FIGURE 2: Example of the AIF time-courses from four AIF measurement methods for the same patient. Note the differences in peak values. Peak values for the four AIF methods at time 54 sec are: peak AIFaMc- = 4.12 mM, peak AIFaMc+ = 2.54 mM, peak AIFmMc- = 5.50 mM, peak AIFmMc+ = 4.94 mM.

FIGURE 3: Boxplots of AIF peak values obtained with four methods of AIF measurement for all patients. The corresponding (mean $\pm$ standard deviation) values are given in the main text.

FIGURE 4: Boxplots of tumor K^{trans} values and statistical comparisons for the four methods of AIF measurements.

FIGURE 5: Boxplots of muscle K^{trans} and statistical comparisons for the four methods of AIF measurements.

FIGURE 6: Boxplots of tumor/muscle ratio K^{trans} values and statistical comparisons for the four methods of AIF measurements.

FIGURE 7: Boxplots of tumor v_e values and statistical comparisons for the four methods of AIF measurements.

FIGURE 8: Boxplots of muscle v_e values and statistical comparisons for the four methods of AIF measurements.

FIGURE 9: Boxplots of tumor/muscle ratio v_e values and statistical comparisons for the four methods of AIF measurements.

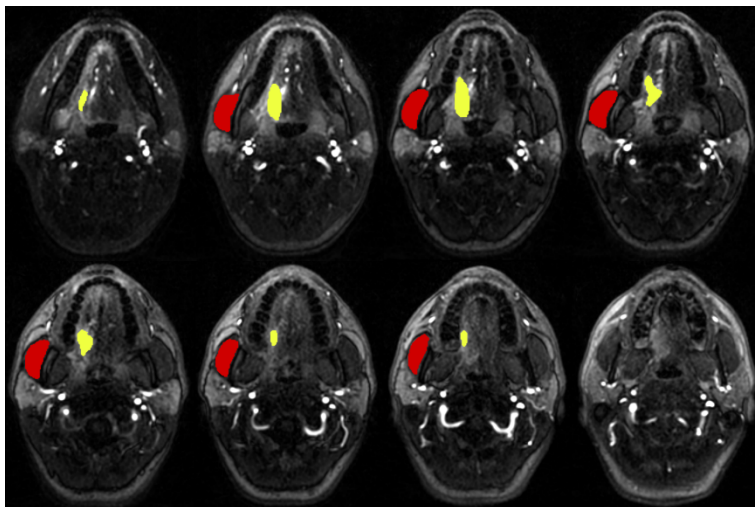


FIGURE 1: Example of VOI for a patient. Tumor VOI (yellow, volume = 5.85 cc, location: base of the tongue), muscle VOI (red, volume = 12.24 cc, location: masseter). Images represent, from left to right and from top to bottom, the 8 slices acquired by DCE-MRI, at the post-injection time 54.24 sec.

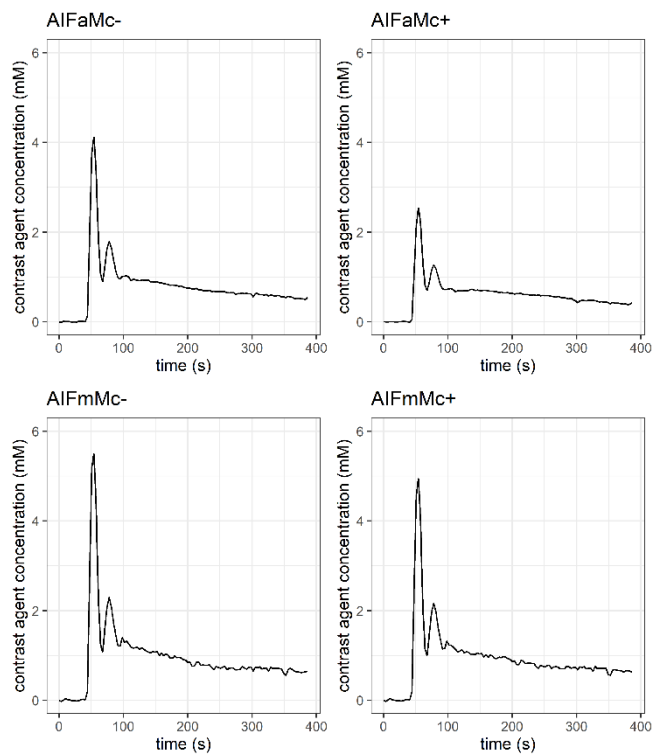


FIGURE 2: Example of the AIF time-courses from four AIF measurement methods for the same patient. Note the differences in peak values. Peak values for the four AIF methods at time 54 sec are: peak AIFaMc- = 4.12 mM, peak AIFaMc+ = 2.54 mM, peak AIFmMc- = 5.50 mM, peak AIFmMc+ = 4.94 mM.

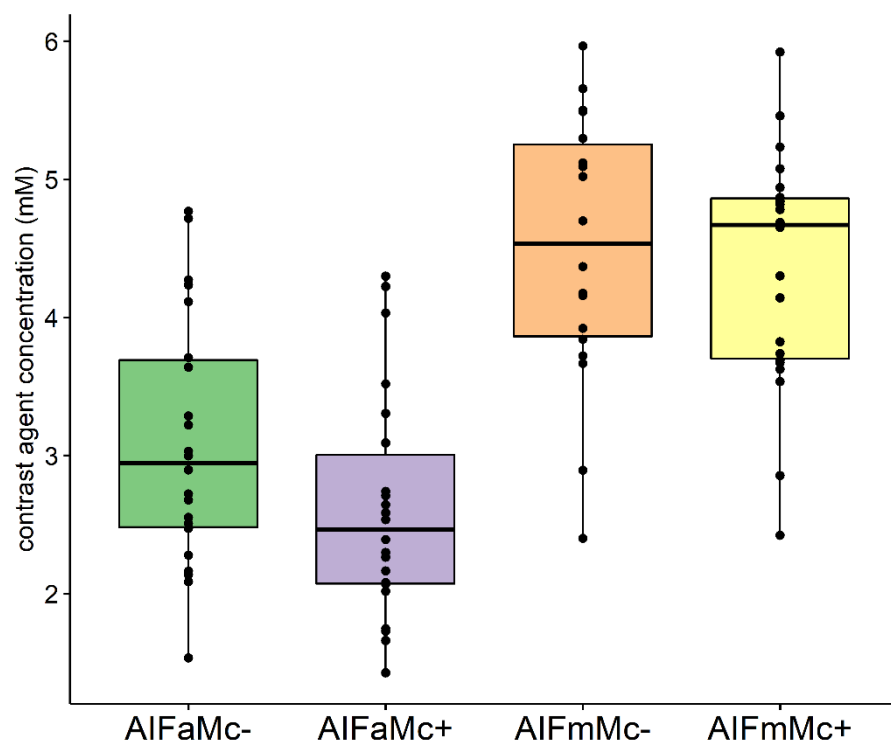


FIGURE 3: Boxplots of AIF peak values obtained with four methods of AIF measurement for all patients. The corresponding (mean $\pm$ standard deviation) values are given in the main text.

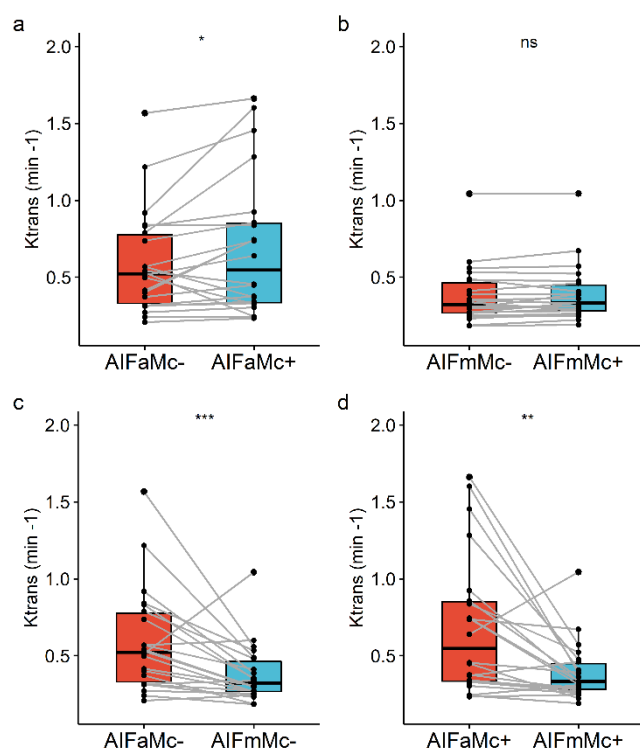


FIGURE 4: Boxplots of tumor K^{trans} values and statistical comparisons for the four methods of AIF measurements.

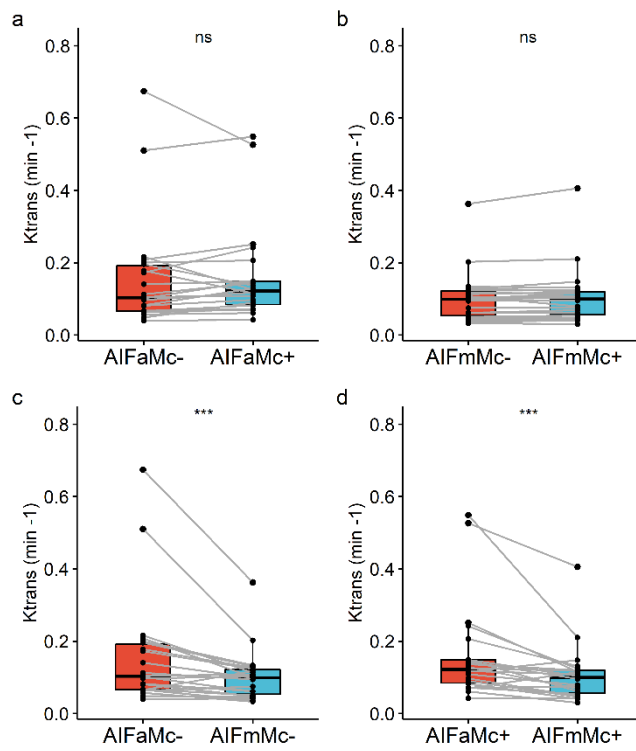


FIGURE 5: Boxplots of muscle K^{trans} and statistical comparisons for the four methods of AIF measurements.

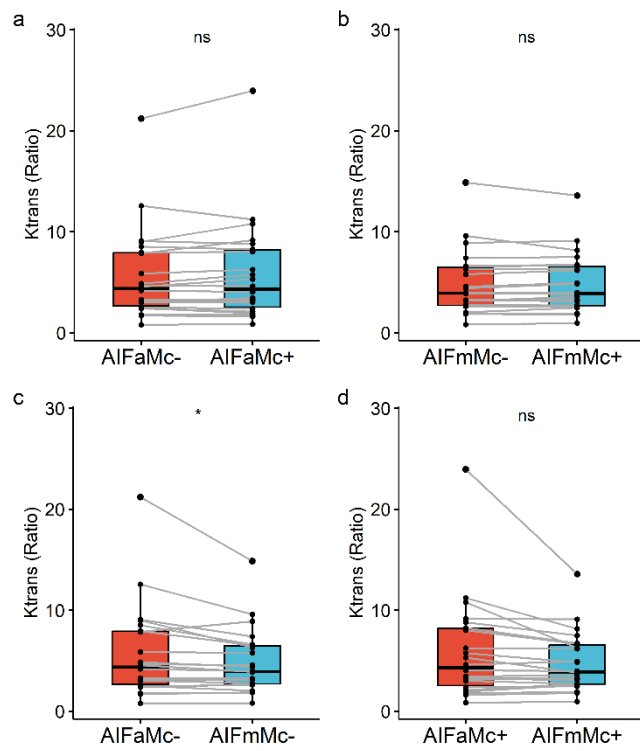


FIGURE 6: Boxplots of tumor/muscle ratio K^{trans} values and statistical comparisons for the four methods of AIF measurements.

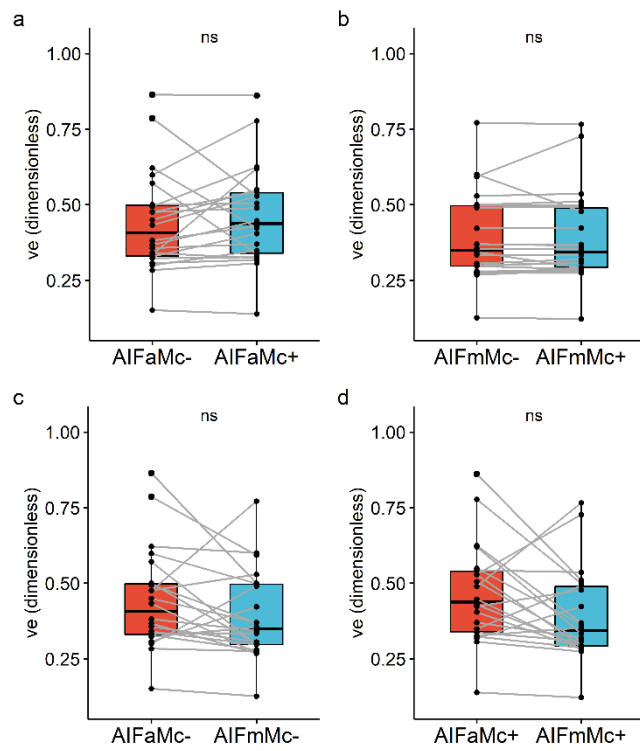


FIGURE 7: Boxplots of tumor v_e values and statistical comparisons for the four methods of AIF measurements.

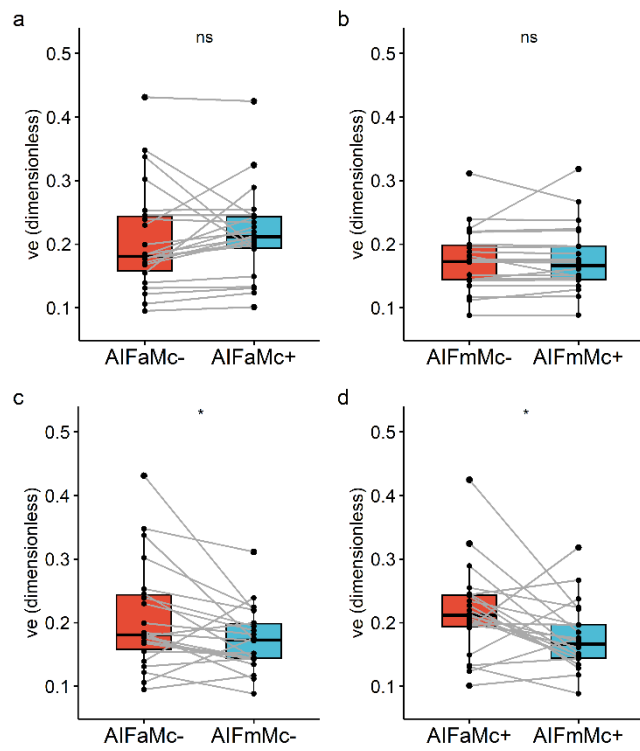


FIGURE 8: Boxplots of muscle v_e values and statistical comparisons for the four methods of AIF measurements.

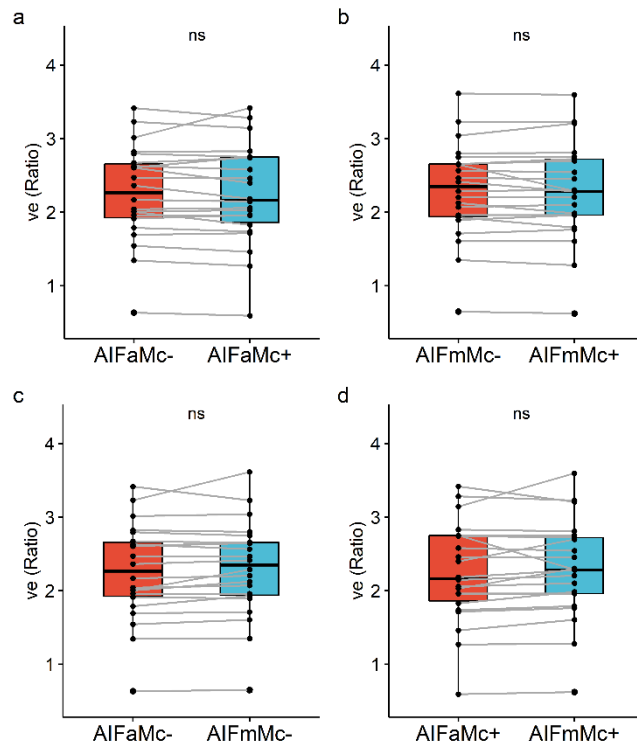


FIGURE 9: Boxplots of tumor/muscle ratio v_e values and statistical comparisons for the four methods of AIF measurements.