

RESEARCH ARTICLE

# Characterizing Heterogeneity within Head and Neck Lesions Using Cluster Analysis of Multi-Parametric MRI Data

Marco Borri<sup>1\*</sup>, Maria A. Schmidt<sup>1</sup>, Ceri Powell<sup>2</sup>, Dow-Mu Koh<sup>1,3</sup>, Angela M. Riddell<sup>3</sup>, Mike Partridge<sup>4</sup>, Shreerang A. Bhide<sup>2</sup>, Christopher M. Nutting<sup>2</sup>, Kevin J. Harrington<sup>2</sup>, Katie L. Newbold<sup>2</sup>, Martin O. Leach<sup>1</sup>

**1** CR-UK Cancer Imaging Centre, The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research, London, United Kingdom, **2** Head and Neck Unit, The Royal Marsden NHS Foundation Trust, London, United Kingdom, **3** Radiology Department, The Royal Marsden NHS Foundation Trust, London, United Kingdom, **4** CRUK/MRC Oxford Institute for Radiation Oncology, University of Oxford, Oxford, United Kingdom

\* [marco.borri@icr.ac.uk](mailto:marco.borri@icr.ac.uk)



OPEN ACCESS

**Citation:** Borri M, Schmidt MA, Powell C, Koh D-M, Riddell AM, Partridge M, et al. (2015) Characterizing Heterogeneity within Head and Neck Lesions Using Cluster Analysis of Multi-Parametric MRI Data. PLoS ONE 10(9): e0138545. doi:10.1371/journal.pone.0138545

**Editor:** Joseph Najbauer, University of Pécs Medical School, HUNGARY

**Received:** April 15, 2015

**Accepted:** September 1, 2015

**Published:** September 23, 2015

**Copyright:** © 2015 Borri et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** The authors acknowledge the support of the Cancer Research UK (CR-UK) and the Engineering and Physical Sciences Research Council (EPSRC) Cancer Imaging Centre in association with the Medical Research Council (MRC) and Department of Health (England) (grants C1060/A10334, C1090/A16464); the National Institute for Health Research (NIHR) funding to the Clinical Research Facility in Imaging and to the Biomedical Research Centre at the Royal Marsden

## Abstract

### Purpose

To describe a methodology, based on cluster analysis, to partition multi-parametric functional imaging data into groups (or clusters) of similar functional characteristics, with the aim of characterizing functional heterogeneity within head and neck tumour volumes. To evaluate the performance of the proposed approach on a set of longitudinal MRI data, analysing the evolution of the obtained sub-sets with treatment.

### Material and Methods

The cluster analysis workflow was applied to a combination of dynamic contrast-enhanced and diffusion-weighted imaging MRI data from a cohort of squamous cell carcinoma of the head and neck patients. Cumulative distributions of voxels, containing pre and post-treatment data and including both primary tumours and lymph nodes, were partitioned into  $k$  clusters ( $k = 2, 3$  or  $4$ ). Principal component analysis and cluster validation were employed to investigate data composition and to independently determine the optimal number of clusters. The evolution of the resulting sub-regions with induction chemotherapy treatment was assessed relative to the number of clusters.

### Results

The clustering algorithm was able to separate clusters which significantly reduced in voxel number following induction chemotherapy from clusters with a non-significant reduction. Partitioning with the optimal number of clusters ( $k = 4$ ), determined with cluster validation, produced the best separation between reducing and non-reducing clusters.

NHS Foundation Trust and Institute of Cancer Research; the MRC award for the 3T Clinical Research Infrastructure Facility (G0701533); the CR-UK programme grant C7224/A13407. MOL is an NIHR Senior Investigator. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Conclusion

The proposed methodology was able to identify tumour sub-regions with distinct functional properties, independently separating clusters which were affected differently by treatment. This work demonstrates that unsupervised cluster analysis, with no prior knowledge of the data, can be employed to provide a multi-parametric characterization of functional heterogeneity within tumour volumes.

## Introduction

Magnetic resonance imaging (MRI) is a highly flexible functional imaging (FI) technique enabling evaluation of multiple aspects of tumour biology in a single examination without radiation exposure [1,2], in contrast to modalities such as positron emission tomography (PET) or contrast-enhanced computer tomography (CT) [1,3]. Dynamic contrast-enhanced (DCE)-MRI provides information on tumour microcirculation, vascularity, blood volume and vessel permeability [4]; diffusion-weighted imaging (DWI) measures differences in tissue microstructure related to tumour cellularity, tissue disorganization, and increased extracellular space tortuosity [5]; intrinsic susceptibility imaging (IS) provides information relating to the oxygenation of tissues [6]. Assessment of changes in FI parameters during treatment could lead to a better understanding of the mechanism of treatment resistance or sensitivity [3]. A combination of FI parameters, each contributing complementary information, is likely to offer increased specificity for differentiation of cancerous tissue [7].

There is a pressing need to reliably describe functional or biological heterogeneity within a tumour. This is particularly relevant in squamous cell carcinoma of the head and neck (SCCHN), a disease which has a worldwide incidence of approximately 500,000 cases per annum [8]. One of the main challenges in advancing the management of SCCHN is the inability to predict treatment outcome due to tumour heterogeneity [2]. Multi-parametric FI may identify potentially treatment resistant or sensitive tumour sub-volumes [3,9] and thereby influence treatment selection such as radiation dose boosts, use of hypoxic cell radiosensitisers [10,11], treatment de-escalation to reduce normal tissue toxicity [12,13], or a change of treatment modality.

When multi-parametric imaging is available, unsupervised clustering techniques can segment different tissue types using a combination of FI parameters [14]. These methodologies are frequently employed to extract anatomical features within the image or to separate the tumour from the surrounding tissue, and their reliability is often assessed against visual inspection. The possibility of applying similar clustering techniques to characterize functional heterogeneity within the tumour volume alone remains underexplored. In this situation, prior knowledge of the data or anatomical references are limited or absent, and therefore testing the partitioning for robustness and consistency is advisable prior to attempting an interpretation of the results. Specifically, the number of clusters, a parameter which is normally pre-set arbitrarily, requires optimization in order for the structures revealed in the data to be meaningful [15].

In this paper we present a methodology, based on cluster analysis, to partition multi-parametric FI data into groups (or clusters) of similar functional characteristics, combining information from both DCE and DWI parameters, with the aim of describing functional heterogeneity within tumour volumes. This approach employs cluster validation techniques and principal component analysis to both understand the structure of the data and to assess the

consistency of the partitioning. This workflow was applied to a cohort of SCCHN patients receiving induction chemotherapy (IC) and undergoing MRI as part of a previously reported feasibility study [3]. The relationship between partitioned data and response to treatment was investigated in order to identify those parameters more likely to characterize treatment-resistant disease.

## Materials and Methods

### Clinical Examinations

As part of a prospective pilot study, described by Powell *et al* in [3], 9 SCCHN patients received two cycles of IC (cisplatin (75 mg/m<sup>2</sup>) day 1 and 5-FU (1000 mg/m<sup>2</sup>) days 1–4) followed by radical chemoradiotherapy (cisplatin (100 mg/m<sup>2</sup>) days 1 and 29); macroscopic and microscopic disease received 65 and 54 Gray (Gy) in 30 fractions, respectively, using intensity modulated radiotherapy (IMRT) with a simultaneous integrated boost technique). All patients underwent MRI at the following time-points: baseline, following IC, after 40 Gy of IMRT, three and six months post-treatment. This study has been approved by the UK National Research Ethics Service (EC 08/H0801/132), and written informed consent has been obtained from all the participants. This cohort included 6 stage IVA, 1 stage IVB and 3 stage III patients, with primary sites located as follows: 5 in the palatine tonsil, 2 in the pyriform fossa, 1 in the nasopharynx, and 1 in the oropharynx [3].

MRI was performed at 1.5T (Philips Intera, Philips Medical Systems, Best, Netherlands), with patients immobilized in a 5-point thermoplastic mask and lying on a flat rigid bed. Two flexible surface coils were centred over the tumour [16] and the following functional imaging sequences were acquired: DCE, following gadolinium injection (0.2 mg/kg, Dotarem<sup>®</sup>, Guerbet, France), with a transaxial 3D T1-weighted spoiled gradient-echo pulse sequence (TE = 1 ms, TR = 4 ms, FOV = 256x256 mm<sup>2</sup>, 2x2x5 mm<sup>3</sup> voxel, 50 mm coverage in the superior/inferior direction, 1.5 s temporal resolution, with a parallel imaging factor of 1.7 in the anterior/posterior direction); DWI (multi-slice EPI, b = 0, 100, 500, 1000 s/mm<sup>2</sup>, TE = 91 ms, TR = 2000 ms, FOV = 256x256 mm<sup>2</sup>, 2x2x5 mm<sup>3</sup> voxel, 20 slices, with SPAIR fat suppression). The DCE protocol acquired proton density-weighted images (FA = 4°), later used for signal conversion to contrast agent concentration, followed by 100 dynamic acquisitions (FA = 20°). DCE and DWI image volumes were aligned and had the same voxel size.

### Data Analysis

For each slice, regions of interest (ROI) were delineated on DCE data by a radiation oncologist together with an expert radiologist around all primary tumours (PT) and target lymph nodes (LN), employing the software package MRIW (Institute of Cancer Research, London, UK), which allows analysis of DCE-MRI data sets using standard pharmacokinetic models [17]. The outline of PTs and LNs was defined as the visible enhancing area immediately following carotid enhancement [18]. Once the onset of enhancement was identified for each voxel, the initial area under the gadolinium curve (IAUGC60—integration over 60 s from the onset) was computed inside the defined ROIs. Spatial registration of DWI and DCE images was performed using a rigid body mutual information algorithm [19] implemented in IDL (version 8.2, Exelis Visual Information Solutions, Boulder, Colorado, USA). After applying the resulting integer voxel shift to DWI data, apparent diffusion coefficient (ADC) maps were produced employing a mono-exponential model to fit all 4 b values, and the ROIs delineated on DCE data were transferred to the computed maps. For DWI data, an independent set of manually delineated ROIs was also produced, encompassing the area of impeded diffusion on the b = 1000 images.

## Cluster Analysis

**Clustering algorithm.** Cluster analysis was performed using in-house software coded in IDL. The K-means algorithm was employed to partition the ROIs into sub-regions of similar characteristics, defined in the two-dimensional feature space formed by the two parameters ADC and IAUGC60. The robust implementation of the K-means algorithm included in the software package CCHIPS [20] was incorporated into the code, as it provides reproducible results and is independent of starting conditions. The algorithm standardizes the data, accounting for the different scales of the parameters. Following cluster analysis, each voxel was assigned a colour, with different colours representing membership of a different cluster.

In order to obtain a generic classification of the data, applicable to the whole cohort, the clustering algorithm was applied to the cumulative distribution of voxels containing data from all the patients and including both baseline and post-IC data. The slice with the largest cross-section (referred to as “*central ROI*”) was selected for each lesion. The cumulative distribution containing pre and post-treatment *central ROIs* from both PTs and LNs was partitioned into  $k$  clusters ( $k = 2, 3$  or  $4$ ).

After classification, the number of voxels assigned to different clusters within each lesion pre and post-treatment was counted. For each cluster, paired groups of voxel counts were compared using the Wilcoxon signed-ranks test; a two-tailed  $p$ -value  $< 0.05$  was chosen to indicate a significant difference in the number of voxels between pre and post-treatment data.

In Appendix 1 pre and post-treatment voxels counts are reported in detail for the PT *central ROIs*, in the simplest case of  $k = 2$ , with the purpose of demonstrating how the proposed method analyses longitudinal data. Additionally, the PTs were re-analysed comparing the use of the *central ROIs* only versus the use of the entire PT volumes, in order to determine if the slice with largest tumour cross-section represents the tissue type content of the whole lesion.

**Principal component analysis.** In order to assess the relative contribution of each of the two parameters (IAUGC60 and ADC) to the partitioning, principal component analysis (PCA) [21] was performed on the cumulative distribution of voxels containing the *central ROIs* of both PTs and LNs.

**Cluster validation.** A fundamental problem in cluster analysis is the determination of the optimal number of clusters, i.e. the value of  $k$  which best describes the data [22]. The “jump method” [15] was employed to determine the value of  $k$  which maximizes efficiency while minimizing error by information theory standards. This method makes use of a measure of data dispersion within the cluster indicated as *transformed distortion* and defined as:

$$d_k = \frac{1}{p} \cdot \min_{c_1 \dots c_k} \text{MAHAL}(\mathbf{X}, \mathbf{c}_x)^{\frac{p}{2}} \quad (1)$$

Where  $k$  = number of clusters

$p$  = number of dimensions of the feature space

$\mathbf{X}$  =  $p$ -dimensional data

$c_1 \dots c_k$  = cluster centres, being  $c_x$  the one closest to  $\mathbf{X}$

MAHAL = Mahalanobis distance

This method was applied to the partitioned data and produced a *distortion* value for each  $k$ . The validation curve plots the difference between consecutive *distortion* values as a function of  $k$ , and was generated for the set of data containing the *central ROIs* of both PTs and LNs. The value of  $k$  associated with the largest “jump” in the validation curve represents the optimal number of clusters [15].

## Results

Fourteen separate lesions, 5 PTs and 9 LNs, were considered in the analysis of 9 patients. Although 25 lesions were depicted in this cohort, complete functional datasets were available only in 14 lesions, due to differences in superior/inferior coverage between the DCE and DWI sequences. A total of 2179 voxels from the *central ROIs* (1435 pre-treatment and 744 post-IC) were included in this analysis. After IMRT and at three months post-treatment the total number of voxels was 423 and 220, respectively. At three months post-treatment, a clinical assessment stratified patients into responders and non-responders. At six months post-treatment there was no residual disease in either group, as either the lesions completely responded to treatment (in 7 patients), or were resected with modified radical neck dissection (in 2 patients) [3].

The ROIs contouring the area of enhancement on DCE data were employed in the analysis. For large PTs (an example in Fig 1) this area is larger than the core of restricted diffusion contoured on DWI data (a total of 970 and 703 voxels for all *central ROIs* at baseline, for DCE and DWI respectively), while for the LNs, which are generally encapsulated structures, there is minimal difference in size between DCE and DWI ROIs (a total of 475 and 490 voxels for all *central ROIs* at baseline, for DCE and DWI respectively).

The mutual information algorithm detected an in-plane shift in the phase-encoding direction of the DWI EPI images, due to offsets of the resonance frequency set by the scanner [7]. This occurred in 54% of the cases and was corrected by applying the computed integer voxel shift to the DWI images. Distortion in the MR images for this imaging protocol was previously assessed [16] and was minimal, leading to a reliable pairing between DCE and DWI voxels.

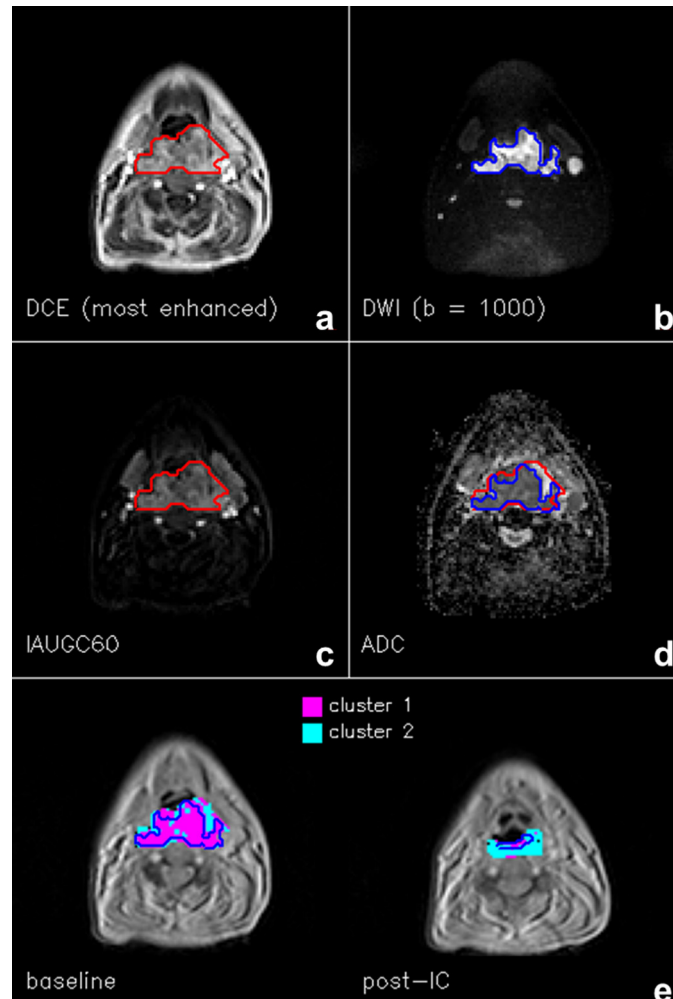
## Cluster Analysis

**Principal component analysis.** PCA analysis found that the two orthogonal components present in the data (Fig 2) are almost parallel to the original parameter axes (IAUGC60 and ADC), thus indicating that the two selected parameters are largely independent. The length of the two vectors in Fig 2 is proportional to the weight of each component. If it is assumed that the first component is almost coincident with ADC and the second with IAUGC60, it can be observed that their contribution is similar, with the first component accounting for 59.9% of the difference in the data.

**Cluster validation.** Validation curves for  $2 \leq k \leq 10$  are presented in Fig 3, showing *distortion* as a function of the number of clusters (grey line). A sharp rise in *distortion* at  $k = 4$  is associated with a sudden increase in performance of the clustering algorithm [15]. The validation curve (black line) offers an alternative visualization by plotting the difference between consecutive *distortion* values, and exhibits a “jump” at  $k = 4$ , which represents the optimal number of clusters.

**Clustering algorithm.** Considering the combination of PTs and LNs, Table 1 details the values of the cluster centres and the number of voxels in each cluster when  $k = 2, 3$  or  $4$ . The reported p-values indicate in which cluster the difference in the number of voxels pre and post-treatment reached significance; these clusters are characterized by high IAUGC60 values. Table 1 shows that with every choice of  $k$  it was possible to obtain these regions when partitioning the data, but with increasing  $k$  the algorithm was able to better separate clusters with a significant percentage reduction in voxel number following IC (83% and 81% with  $k = 4$ ,  $p < 0.02$ ) from clusters with a non-significant reduction (23% and 9% with  $k = 4$ ,  $p > 0.75$ ).

Referring to the data analysed with  $k = 4$ , Fig 4 provides an example of the obtained colour maps, contrasting two LNs found to have different clinical outcomes assessed during long-term follow-up. The LN in Fig 4b, containing predominantly voxels belonging to Clusters 1



**Fig 1. Slice with the largest tumour cross-section of a representative primary tumour. (a)** ROI (red) contoured on the most enhanced frame of the DCE series, **(b)** ROI (blue) contoured on  $b = 1000$  image of the DWI series, **(c)** IAUGC60 map, **(d)** ADC map, with the two ROIs overlapped, and **(e)** cluster analysis map pre and post-induction chemotherapy (IC), in the simplest case of  $k = 2$ . The red ROI also encompasses the region where cluster analysis was performed; the blue ROIs represent the area of restricted diffusion independently contoured on DWI data, almost coincident with Cluster 1, which is characterized by both low ADC and high IAUGC60 values.

doi:10.1371/journal.pone.0138545.g001

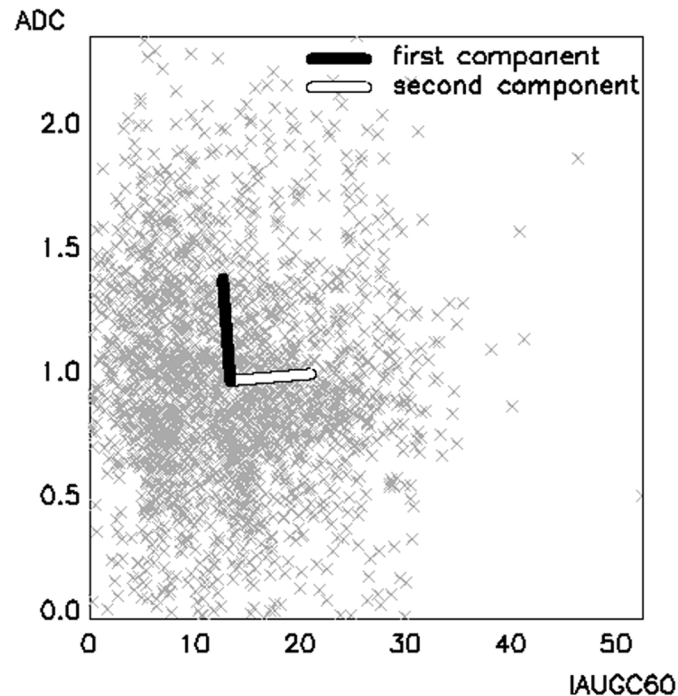
and 2 at baseline, responded to treatment, while the LN in Fig 4c, containing predominantly voxels belonging to Clusters 3 and 4, did not respond.

## Discussion

The proposed method robustly partitioned cumulative distributions of pre and post-treatment functional MRI data from SCCHN lesions into sub-sets of similar functional characteristics. In the cohort of patients analysed, this resulted in a description of functional heterogeneity.

The partitioning was based on two functional parameters derived from different MRI modalities: IAUGC60 and ADC. Wang et al considered two parameters derived from pharmacokinetic analysis of DCE data [9], but did not investigate a correlation between them. In our work PCA demonstrates that in this dataset IAUGC60 and ADC are independent parameters and have similar weight (Fig 2), thus providing distinct and complementary functional





**Fig 2. Principal component analysis.** Principal component analysis in the bi-dimensional space formed by the two parameters IAUGC60 and ADC. The two vectors represent the two orthogonal components. The length of the two vectors is proportional to the weight of each component. IAUGC60 is expressed in units of  $\text{mmol}\cdot\text{s}$ , ADC in units of  $\times 10^{-3} \text{mm}^2/\text{s}$ .

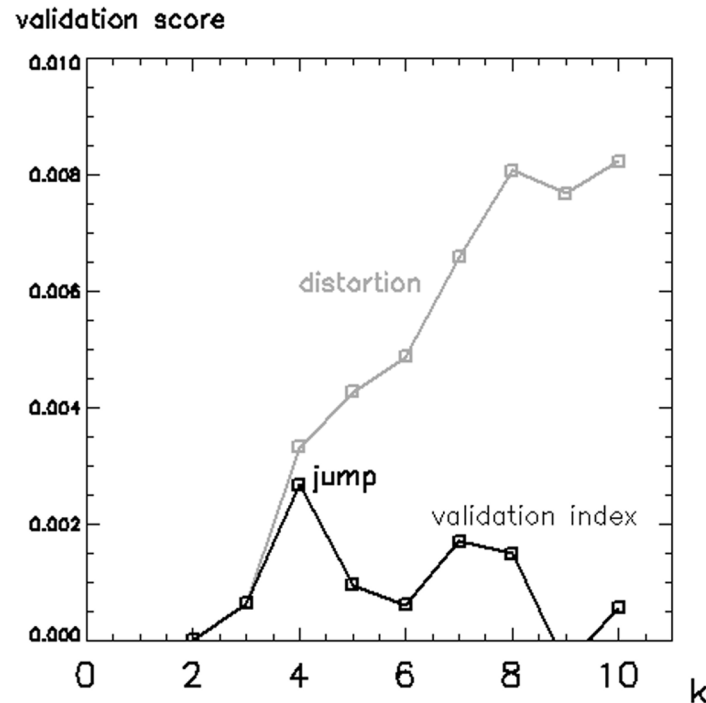
doi:10.1371/journal.pone.0138545.g002

information. While the choice of IAUGC60 limits the specificity of the physiological information, it was preferred to more specific kinetic parameters like  $K^{\text{trans}}$  and  $v_e$ , since it retains a relationship with these parameters but is unaffected by the accuracy of either the chosen contrast-agent uptake model or the arterial input function [4].

Our analysis was performed on data from both PTs and LNs. Limiting the analysis to the *central ROIs* reduces the unbalanced contribution of large lesions to the classification and minimizes inclusion of the voxels located at the lesion edges, where partial volume effects could be significant. However, one should verify that the slice with the largest tumour cross-section is representative of the tissue type content of the lesion (Appendix 1).

The evolution of the resulting sub-regions with IC treatment was assessed relative to the number of clusters ( $k = 2, 3$  or  $4$ ); for all  $k$ , a statistically significant reduction in the number of voxels was found in post-IC data for clusters with high IAUGC60. This suggests that an enhanced delivery of IC in highly perfused areas might be the main determinant of response. A higher  $k$  allowed for a more complete description of heterogeneity, with optimal separation between clusters where the reduction in the voxel count post-IC is minimal and those where it is maximal.

Unsupervised clustering may be employed when, as in this case, there is no prior knowledge of the data. Nevertheless, determining the optimal number of clusters is essential to achieve meaningful results: a lower value could fail to differentiate some of the main structures in the data, while a higher value would simply increase the complexity of the partitioning without necessarily providing additional information. The cluster validation method that we adopted is non-parametric, makes limited assumptions, and has been proven effective in a wide range of problems for identifying underlying structures [15]. Table 1 shows that the



**Fig 3. Cluster validation.** Cluster validation curves for the range  $2 \leq k \leq 10$ . The *distortion* curve (grey line) has a sharp rise in correspondence of the optimal number of clusters  $k = 4$ . The validation curve (black line) is produced by taking the difference between consecutive values of the *distortion* curve, and exhibits a “jump” in correspondence of  $k = 4$ .

doi:10.1371/journal.pone.0138545.g003

separation between reducing and non-reducing clusters improves as  $k$  rises towards the optimal number of clusters ( $k = 4$ ) mathematically derived with the jump method. This would appear to link mathematical computation with real functional differences, and informs the use of cluster validation to increase the reliability of unsupervised partitioning.

PCA and cluster validation can be employed as validation tools to investigate the composition of the data. In this case, both PCA (Fig 2) and the partitioning with the optimal number of clusters (Fig 4a) suggest that a binary classification of both parameters best separates the data, demonstrating consistency of results. Users should attempt to reproduce this consistency when applying this workflow to other datasets with different or higher number of clusters, higher data dimensions and possible internal correlations.

The analysed distribution of voxels included different tissue types before and after IC treatment. The adopted method performs an unsupervised partitioning of functional parameters in cumulative and longitudinal data, with the purpose of describing functional heterogeneity. Each pair of functional parameters describes tissue behaviour within a voxel, and as such, each voxel, rather than each lesion, contributes information to the global classification. However, the algorithm does not distinguish between pre and post-treatment data; it is therefore a remarkable result that this method independently separated voxels more likely to be affected by treatment.

This analysis was confined to baseline and post-IC data, because only small volumes of residual disease were present at subsequent time-points. Following IMRT treatment, response was dominated by volume reduction rather than by changes in functional parameters. In this situation, description of functional heterogeneity across a volume of few voxels is of limited value. Furthermore, in very small lesions both the weight and the reliability of the information



**Table 1. Cluster analysis as a function of the number of clusters.**

		Cluster 1		Cluster 2		Cluster 3		Cluster 4		Total	
<b>k = 2</b>	<b>Cluster centres</b>	IAUGC60 = 17.8, ADC = 0.73		IAUGC60 = 8.8, ADC = 1.21							
		<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>					<b>Baseline</b>	<b>Post-IC</b>
	<b>Voxel count</b>	828	285	607	459					1435	744
	<b>Reduction (%)</b>	66%		24%						48%	
	<b>p-value</b>	0.0074*		0.2005							
<b>k = 3</b>	<b>Cluster centres</b>	IAUGC60 = 22.5, ADC = 0.93		IAUGC60 = 9.6, ADC = 0.67		IAUGC60 = 9.1, ADC = 1.4					
		<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>			<b>Baseline</b>	<b>Post-IC</b>
	<b>Voxel count</b>	574	95	500	374	361	275			1435	744
	<b>Reduction (%)</b>	83%		25%		24%				48%	
	<b>p-value</b>	0.0045*		0.9045		0.8181					
<b>k = 4</b>	<b>Cluster centres</b>	IAUGC60 = 22.2, ADC = 0.82		IAUGC60 = 19.7, ADC = 1.55		IAUGC60 = 9.9, ADC = 0.63		IAUGC60 = 7.2, ADC = 1.27			
		<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>	<b>Baseline</b>	<b>Post-IC</b>
	<b>Voxel count</b>	486	84	199	37	422	324	328	299	1435	744
	<b>Reduction (%)</b>	83%		81%		23%		9%		48%	
	<b>p-value</b>	0.0024*		0.0198*		0.7642		0.8572			

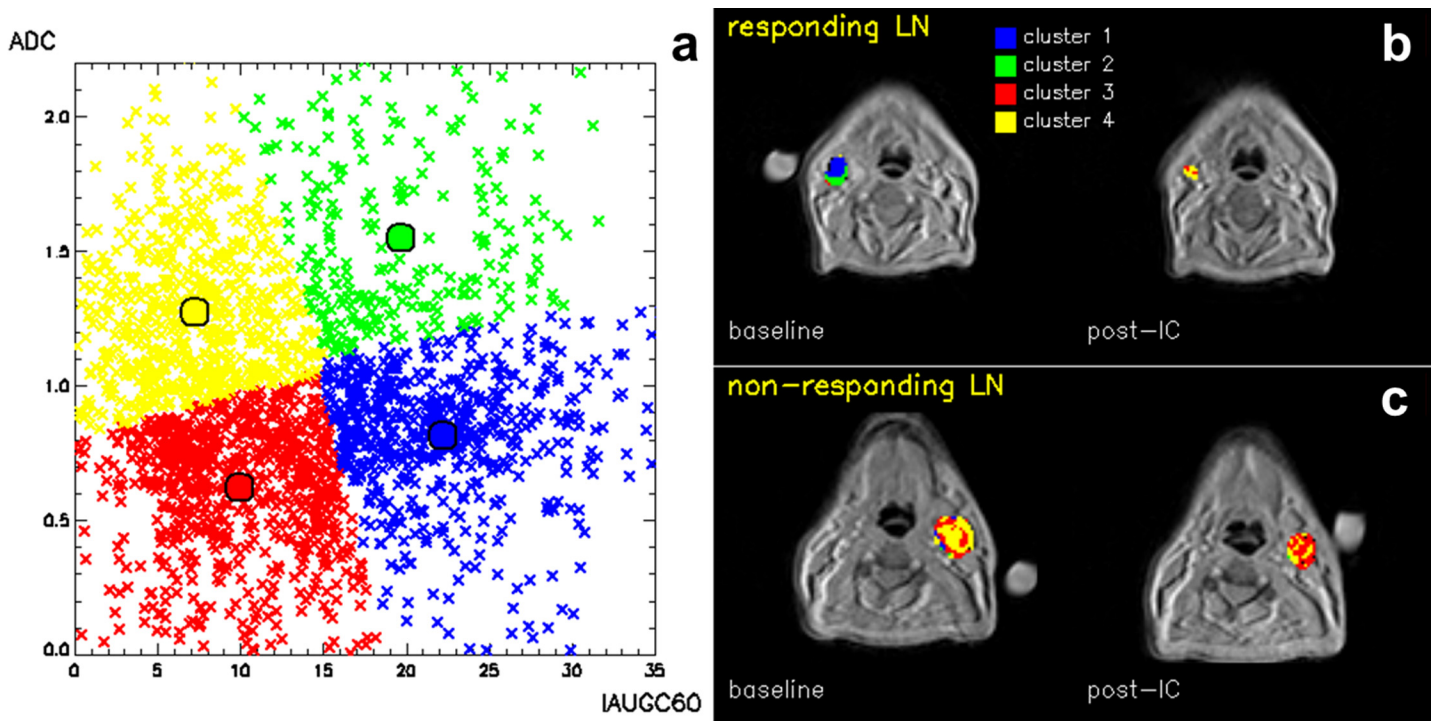
Results from cluster analysis (cluster centres and voxel counts) as a function of k (number of clusters), and percentage reduction after induction chemotherapy (IC) for the combination of primary tumours and lymph nodes. In total, 2179 voxels were considered in the analysis, from which 1435 were measured at baseline and 744 post-IC, with a global percentage reduction of 48%. Asterisks indicate significant two-tailed p-values ( $p < 0.05$ ) from the Wilcoxon signed-ranks test, comparing the number of voxels pre and post-treatment. IAUGC60 (initial area under the gadolinium curve integrated over 60 s) is expressed in units of mmol·s, ADC (apparent diffusion coefficient) in units of  $\times 10^{-3}$  mm<sup>2</sup>/s.

doi:10.1371/journal.pone.0138545.t001

extracted from the voxels are sensibly reduced. This type of analysis could have a potentially greater value at early stages of treatment response, where it could identify the sub-volumes whose functional parameter are more likely to be associated with treatment-resistant disease.

Each cluster represents a tissue class with a separate functional behaviour defined by the MRI parameters. Consideration of the cluster centre values, which summarize the functional parameters of each cluster, for the optimal  $k = 4$  (Table 1) suggests the presence of regions of high vascular activity and regions that are more likely to be necrotic. Cluster 1 has increased perfusion (higher IAUGC60) and increased cellularity (lower ADC) and may represent a phenotype with rapid growth [23]. Alternatively, as a low ADC and high  $K^{trans}$  [24] at baseline predict treatment response [25], this phenotype may select tumours which are likely to respond to treatment, as our analysis seems to suggest. Cluster 2 separates areas with high perfusion and high ADC, suggesting the presence of oedema. The low perfusion associated with Clusters 3 and 4 could indicate necrotic tissue, whilst those voxels exhibiting a higher ADC in Cluster 4 might indicate the presence of liquid necrosis. However, histological validation is required in order to assign biological meaning to the clusters.

Functional heterogeneity within each lesion can be described by the presence of different clusters in each ROI. As an example, Fig 4 displays the cluster composition at baseline in two representative LNs: the non-responding LN is predominantly described by clusters with low IAUGC60 (Cluster 1 and 2), unlike the responding LN. The presence at baseline of those clusters which demonstrate low vascularity (low IAUGC60) would appear to be linked to treatment resistance. This suggests that the proposed analysis, limited to the initial response to IC, could have the potential to predict favourable and unfavourable prognostic groups.



**Fig 4. Cluster analysis.** (a) Cumulative distribution of voxels (primary tumours + lymph nodes), partitioned with  $k = 4$ , in the bi-dimensional space formed by the two parameters IAUGC60 and ADC. IAUGC60 is expressed in units of  $\text{mmol}\cdot\text{s}$ , ADC in units of  $\times 10^{-3} \text{mm}^2/\text{s}$ . (b,c) Corresponding cluster analysis maps of a responding (b) and a non-responding (c) lymph node, at baseline and post-induction chemotherapy (IC). Different colours indicate a difference in functionality at baseline.

doi:10.1371/journal.pone.0138545.g004

The multi-parametric approach described in this paper differs from methodologies which consider single parameters separately and provide a description of their histogram distribution. In these cases, heterogeneity can be characterized by defining thresholds based on the analysis of the associated histogram. However, thresholds can not be determined independently and would necessitate evaluation of treatment response on a large number of cases. Furthermore, extension of threshold values to other datasets acquired with different MRI scanners and sequences is not straightforward, as thresholds rely on the numerical value of the parameter. Conversely, unsupervised cluster analysis, as presented here, produces sub-divisions as a function of distance in the multi-dimensional feature space without any prior knowledge. This type of analysis can therefore be transferred to other dataset, as it relies on the relationship between the parameters rather than their absolute numerical value.

The main aim of this paper is to present an approach to analyse longitudinal DCE and DWI data with the objective of characterizing tumour heterogeneity and response in SCCHN. The retrospective analysis performed on this pilot dataset is subject to a number of limitations. The low incidence of non-responding lesions precluded a statistically significant evaluation of the predictive value of our analysis. In addition, the absence of histological correlation limits the biological inferences which can be assigned to the clusters. This approach now warrants evaluation in a larger cohort of patients together with biological validation.

## Conclusion

We have presented a robust approach to partition cumulative distributions of functional parameters from SCCHN lesions, which employs unsupervised clustering, with no prior knowledge of

the data, to separate voxels with distinct functional properties. In a distribution of pre and post-treatment MRI functional parameters, belonging to both PTs and LNs, the proposed methodology was able to separate clusters where the reduction in voxel number post treatment is maximal from those where it is minimal. Prior assessment of data composition and mathematical optimization of the number of clusters increased the reliability of the partitioning. This analysis demonstrated that independent functional parameters from different MRI modalities can be considered jointly to provide a multi-parametric characterization of functional heterogeneity within tumour volumes. By identifying those functional parameters more likely to characterize treatment-resistant disease, this method could have the potential to predict treatment outcome, to identify a biological target volume for treatment escalation or to stratify patients for alternative treatment modalities.

### Appendix 1

This section presents in detail the longitudinal analysis of a subset of the data containing pre and post-treatment PT *central ROIs*, partitioned with the minimum number of clusters ( $k = 2$ ). [Table 2](#) reports the number of voxels in each sub-region for each of the 5 PTs and in total, separating pre and post-treatment data. A global reduction in the number of voxels of 52% was observed, indicating response to treatment. Cluster 1 ([Table 2](#), [Fig 1e](#)) is characterized by voxels with low ADC and high IAUGC60 values, associated with increased perfusion and cellularity. [Fig 1e](#) also demonstrates that this area is largely coincident with the area of restricted diffusion independently delineated on DWI data. The number of voxels assigned to this cluster reduced by 70% post-treatment. Voxels within this cluster were affected by treatment in greater proportion; these areas are both cellular and highly perfused, facilitating exposure of active tumour cells to induction chemotherapy. In contrast, Cluster 2 is characterized by low IAUGC60 and higher ADC values, and the voxel count reduced by only 7% post-treatment. While  $k = 2$  might appear sufficient for partitioning the dataset containing only the PTs, the inclusion of the LNs required  $k = 4$  to optimally characterize the higher data complexity introduced.

The 5 PT ROIs were also re-analysed comparing the use of the *central ROIs* only (1439 voxels) versus the use of the entire PT volumes (7067 voxels) for  $k = 2, 3$  and 4, allowing investigation of the approach of including only the largest cross-section into the cumulative

**Table 2. (Appendix 1). Cluster analysis of the primary tumours.**

Cluster centres	Cluster 1		Cluster 2		Total	
	IAUGC60 = 17.3, ADC = 0.75		IAUGC60 = 8.1, ADC = 1.30			
Number of voxels	Baseline	Post-IC	Baseline	Post-IC	Baseline	Post-IC
Patient 1	354	22	103	124	457	146
Patient 2	187	100	23	11	210	111
Patient 3	37	11	14	25	51	36
Patient 4	38	14	5	13	43	27
Patient 5	71	59	138	90	209	149
<b>Total</b>	<b>687</b>	<b>206</b>	<b>283</b>	<b>263</b>	<b>970</b>	<b>469</b>
<b>Reduction (%)</b>	70%		7%		52%	

Number of voxels, for each of the 5 primary tumours and in total, belonging to the two clusters determined by cluster analysis with  $k = 2$ . Pre and post-induction chemotherapy data are separated, and the total number of voxels in the ROIs is reported. IAUGC60 is expressed in units of  $\text{mmol}\cdot\text{s}$ , ADC in units of  $\times 10^{-3} \text{mm}^2/\text{s}$ .

doi:10.1371/journal.pone.0138545.t002

distribution. This is particularly relevant for PTs, which are often larger and asymmetrical structures. The two approaches provided similar sub-regions with any value of  $k$ , and cluster centre values no more than 9.1% apart (with an average difference of  $4.3 \pm 2.9\%$  in the cases examined). This indicates that the analysis can be limited to the *central ROI* as the slice with the largest cross-section represents the tissue type content of the whole lesion, independently of the value of  $k$  adopted.

## Acknowledgments

The authors would also like to thank Minty Ledger, Matt Orton, James d'Arcy, Jessica Boulton, Rafal Panek, and Liam Welsh for helpful discussions.

## Author Contributions

Conceived and designed the experiments: MB MAS CP DMK AMR MP SAB CMN KJH KLN MOL. Performed the experiments: MB MAS CP. Analyzed the data: MB MAS CP DMK AMR. Contributed reagents/materials/analysis tools: MB. Wrote the paper: MB MAS CP DMK AMR MP SAB CMN KJH KLN MOL. Designed the software used in analysis: MB. Designed the clinical study: SAB CMN KJH KLN. Designed the imaging protocol: MB MAS MP MOL. Reviewed the clinical images: CP DMK AMR.

## References

1. Quon H, Brizel DM. Predictive and prognostic role of functional imaging of head and neck squamous cell carcinomas. *Semin Radiat Oncol*. 2012 Jul; 22(3):220–32. doi: [10.1016/j.semradonc.2012.03.007](https://doi.org/10.1016/j.semradonc.2012.03.007) PMID: [22687947](https://pubmed.ncbi.nlm.nih.gov/22687947/)
2. Newbold K, Partridge M, Cook G, Sohaib SA, Charles-Edwards E, Rhys-Evans P, et al. Advanced imaging applied to radiotherapy planning in head and neck cancer: a clinical review. *Br J Radiol*. 2006 Jul 1; 79(943):554–61. PMID: [16823059](https://pubmed.ncbi.nlm.nih.gov/16823059/)
3. Powell C, Schmidt M, Borri M, Koh D-M, Partridge M, Riddell A, et al. Changes in functional imaging parameters following induction chemotherapy have important implications for individualised patient-based treatment regimens for advanced head and neck cancer. *Radiotherapy and Oncology*. 2013 Jan; 106(1):112–7. doi: [10.1016/j.radonc.2012.09.009](https://doi.org/10.1016/j.radonc.2012.09.009) PMID: [23089306](https://pubmed.ncbi.nlm.nih.gov/23089306/)
4. Walker-Samuel S, Leach MO, Collins DJ. Evaluation of response to treatment using DCE-MRI: the relationship between initial area under the gadolinium curve (IAUGC) and quantitative pharmacokinetic analysis. *Phys Med Biol*. 2006 Jul 21; 51(14):3593–602. PMID: [16825751](https://pubmed.ncbi.nlm.nih.gov/16825751/)
5. Padhani AR, Liu G, Mu-Koh D, Chenevert TL, Thoeny HC, Takahara T, et al. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. *Neoplasia*. 2009 Feb; 11(2):102–25. PMID: [19186405](https://pubmed.ncbi.nlm.nih.gov/19186405/)
6. McPhail LD, Robinson SP. Intrinsic susceptibility MR imaging of chemically induced rat mammary tumors: relationship to histologic assessment of hypoxia and fibrosis. *Radiology*. 2010 Jan; 254(1):110–8. doi: [10.1148/radiol.2541090395](https://doi.org/10.1148/radiol.2541090395) PMID: [20032145](https://pubmed.ncbi.nlm.nih.gov/20032145/)
7. Riches SF, Payne GS, Morgan VA, Sandhu S, Fisher C, Germuska M, et al. MRI in the detection of prostate cancer: combined apparent diffusion coefficient, metabolite ratio, and vascular parameters. *AJR Am J Roentgenol*. 2009 Dec; 193(6):1583–91. doi: [10.2214/AJR.09.2540](https://doi.org/10.2214/AJR.09.2540) PMID: [19933651](https://pubmed.ncbi.nlm.nih.gov/19933651/)
8. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. *International Journal of Cancer*. 2001; 94(2):153–6.
9. Wang P, Popovtzer A, Eisbruch A, Cao Y. An approach to identify, from DCE MRI, significant subvolumes of tumors related to outcomes in advanced head-and-neck cancer. *Medical Physics*. 2012; 39(8):5277–85. doi: [10.1118/1.4737022](https://doi.org/10.1118/1.4737022) PMID: [22894453](https://pubmed.ncbi.nlm.nih.gov/22894453/)
10. Overgaard J, Hansen HS, Overgaard M, Bastholt L, Berthelsen A, Specht L, et al. A randomized double-blind phase III study of nimorazole as a hypoxic radiosensitizer of primary radiotherapy in supraglottic larynx and pharynx carcinoma. Results of the Danish Head and Neck Cancer Study (DAHANCA) Protocol 5–85. *Radiother Oncol*. 1998 Feb; 46(2):135–46. PMID: [9510041](https://pubmed.ncbi.nlm.nih.gov/9510041/)
11. Overgaard J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck—a systematic review and meta-analysis. *Radiother Oncol*. 2011 Jul; 100(1):22–32. doi: [10.1016/j.radonc.2011.03.004](https://doi.org/10.1016/j.radonc.2011.03.004) PMID: [21511351](https://pubmed.ncbi.nlm.nih.gov/21511351/)

12. Ling CC, Humm J, Larson S, Amols H, Fuks Z, Leibel S, et al. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *International Journal of Radiation Oncology\*Biophysics*. 2000 Jun 1; 47(3):551–60.
13. Nuyts S. Defining the target for radiotherapy of head and neck cancer. *Cancer Imaging*. 2007; 7(Special Issue A):S50–S55.
14. Clarke LP, Velthuisen RP, Camacho MA, Heine JJ, Vaidyanathan M, Hall LO, et al. MRI segmentation: methods and applications. *Magn Reson Imaging*. 1995; 13(3):343–68. PMID: [7791545](#)
15. Sugar CA, James GM. Finding the Number of Clusters in a Dataset. *Journal of the American Statistical Association*. 2003; 98(463):750–63.
16. Ahmed M, Schmidt M, Sohaib A, Kong C, Burke K, Richardson C, et al. The value of magnetic resonance imaging in target volume delineation of base of tongue tumours—a study using flexible surface coils. *Radiother Oncol*. 2010 Feb; 94(2):161–7. doi: [10.1016/j.radonc.2009.12.021](#) PMID: [20096947](#)
17. D'Arcy JA, Collins DJ, Padhani AR, Walker-Samuel S, Suckling J, Leach MO. Informatics in Radiology (infoRAD): Magnetic Resonance Imaging Workbench: analysis and visualization of dynamic contrast-enhanced MR imaging data. *Radiographics*. 2006 Apr; 26(2):621–32. PMID: [16549620](#)
18. Craciunescu OI, Yoo DS, Cleland E, Muradyan N, Carroll MD, MacFall JR, et al. Dynamic contrast-enhanced MRI in head-and-neck cancer: the impact of region of interest selection on the intra- and interpatient variability of pharmacokinetic parameters. *Int J Radiat Oncol Biol Phys*. 2012 Mar 1; 82(3): e345–350. doi: [10.1016/j.ijrobp.2011.05.059](#) PMID: [21985945](#)
19. Roulston MS. Estimating the errors on measured entropy and mutual information. *Physica D: Nonlinear Phenomena*. 1999 Jan 15; 125(3–4):285–94.
20. Schmithorst VJ, Holland SK, Dardzinsk BJ. CCHIPS: Cincinnati Children's Hospital Image Processing Software. 2010. Available: <https://irc.cchmc.org/software/cchips.php>
21. Smith LI. A tutorial on principal components analysis. Cornell University, USA. 2002; 51:52.
22. Halkidi M, Batistakis Y, Vazirgiannis M. On Clustering Validation Techniques. *Journal of Intelligent Information Systems*. 2001 Dec 1; 17(2–3):107–45.
23. Leach MO, Morgan B, Tofts PS, Buckley DL, Huang W, Horsfield MA, et al. Imaging vascular function for early stage clinical trials using dynamic contrast-enhanced magnetic resonance imaging. *Eur Radiol*. 2012 Jul 1; 22(7):1451–64. doi: [10.1007/s00330-012-2446-x](#) PMID: [22562143](#)
24. Kim S, Loevner LA, Quon H, Kilger A, Sherman E, Weinstein G, et al. Prediction of response to chemoradiation therapy in squamous cell carcinomas of the head and neck using dynamic contrast-enhanced MR imaging. *AJNR Am J Neuroradiol*. 2010 Feb; 31(2):262–8. doi: [10.3174/ajnr.A1817](#) PMID: [19797785](#)
25. Vandecaveye V, De Keyzer F, Nuyts S, Deraedt K, Dirix P, Hamaekers P, et al. Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings. *Int J Radiat Oncol Biol Phys*. 2007 Mar 15; 67(4):960–71. PMID: [17141979](#)