Original Research Article

Development and optimisation of a preclinical cone beam computed tomography-based radiomics workflow for radiation oncology research

Kathryn H. Brown [a,*], Neree Payan [a], Sarah Osman [b], Mihaela Ghita [a], Gerard M. Walls [a,c],
Ileana Silvestre Patallo [d], Giuseppe Schettino [d], Kevin M. Prise [a], Conor K. McGarry [a,c],
Karl T. Butterworth [a]

[a] Patrick G. Johnston Centre for Cancer Research, Queen's University Belfast, Northern Ireland, UK
[b] University College London Hospitals NHS Foundation Trust Department of Radiotherapy, London, UK
[c] Cancer Centre, Belfast Health & Social Care Trust, Lisburn Road, Belfast BT9 7AB, Northern Ireland, UK
[d] National Physical Laboratory, London, UK

ARTICLE INFO

ABSTRACT

Background and purpose: Radiomics features derived from medical images have the potential to act as imaging biomarkers to improve diagnosis and predict treatment response in oncology. However, the complex relationships between radiomics features and the biological characteristics of tumours are yet to be fully determined. In this study, we developed a preclinical cone beam computed tomography (CBCT) radiomics workflow with the aim to use *in vivo* models to further develop radiomics signatures.

Materials and methods: CBCT scans of a mouse phantom were acquired using onboard imaging from a small animal radiotherapy research platform (SARRP, Xstrahl). The repeatability and reproducibility of radiomics outputs were compared across different imaging protocols, segmentation sizes, pre-processing parameters and materials. Robust features were identified and used to compare scans of two xenograft mouse tumour models (A549 and H460).

Results: Changes to the radiomics workflow significantly impact feature robustness. Preclinical CBCT radiomics analysis is feasible with 119 stable features identified from scans imaged at 60 kV, 25 bin width and 0.26 mm slice thickness. Large variation in segmentation volumes reduced the number of reliable radiomics features for analysis. Standardization in imaging and analysis parameters is essential in preclinical radiomics analysis to improve accuracy of outputs, leading to more consistent and reproducible findings.

Conclusions: We present the first optimised workflow for preclinical CBCT radiomics to identify imaging biomarkers. Preclinical radiomics has the potential to maximise the quantity of data captured in *in vivo* experiments and could provide key information supporting the wider application of radiomics.

1. Introduction

Medical imaging is central to clinical decision-making for the identification of tumours, delivery of treatment and follow-up assessments [1]. It is well established that these radiological images are data rich and can be used as imaging biomarkers [2]. With the commercialisation of parallel preclinical computed tomography (CT) and cone-beam CT (CBCT) imaging platforms onboard small animal irradiators [3]; imaging biomarkers can be determined from these preclinical scans [4–6].

Radiomics is a high-throughput form of image analysis to extract quantitative information from medical images which can be correlated to biological outcomes to improve diagnostic, prognostic and predictive accuracy [7–11]. Whilst radiomics has been termed a 'virtual biopsy' and associated with several clinical endpoints, the complex relationships between radiomics and clinical factors are still largely unknown [12]. Standardisation of image acquisition and analysis to identify and validate imaging biomarkers is a large focus within radiation oncology [13–15].

The prognostic potential of magnetic resonance (MR)- and CT-based clinical radiomics has already been well documented within the literature [1,16–18], with emerging evidence of feasibility using CBCT scans [19–22]. CBCT scans are acquired at multiple timepoints throughout

radiotherapy treatment and extraction of radiomics signatures from these could lead to surplus data in both clinical and preclinical settings [23].

Previously, Panth *et al* demonstrated that mouse models can be used to expand our knowledge of CT-based radiomics signatures [24]. Since then, preclinical radiomics analysis has evolved to include CT, MRI and PET imaging for the detection and prediction of tumour phenotypes, early metastases and treatment response [24–27]. However, preclinical radiomics lacks standardisation of methods and validation of results [28]. This is in addition to the lack of imaging standards and protocols which already exist within preclinical studies [29]. Repeatability and reproducibility analysis is therefore crucial to evaluate feature stability in a controlled scenario (test–retest) and the influence of different imaging acquisition or analysis parameters (scan-rescan) [30].

In this study, we assessed the repeatability and reproducibility of CBCT-based radiomics features toward standardising the first preclinical CBCT radiomics workflow. Different image acquisition protocols and feature extraction methods were trialled to identify a subset of features that are robust for analysis. These features were then applied to preclinical tumour models in a pilot feasibility analysis.

2. Materials and methods

2.1. Phantoms

Two phantoms were used in this study (Supplementary Fig. 1). Firstly, an anatomically correct, tissue-equivalent mouse phantom with densities and atomic composition for bone (1.39 g/cm^3), lung (0.68 g/cm^3) and soft tissue (1.01 g/cm^3) was used for workflow analysis [31,32]. Secondly, an in-house Perspex phantom (60x60x60 mm) with cylindrical inserts (20x60mm) for air, solid water (Bart's) (1.05 g/cm^3), PVC (1.47 g/cm^3) and acetal (1.52 g/cm^3) was used to compare how differences in material density effect texture features.

2.2. Imaging

CBCT imaging was performed using the Small Animal Radiation Research Platform (SARRP, Xstrahl Life Sciences, UK) (Supplementary Table 1). For the mouse phantom, scans were acquired twice at 40, 50 and 60 kV and 0.8 mA (0.5 mm Al filtration). For the texture phantom, scans were acquired twice at 60 kV. All energies had an imaging dose of 2.4 cGy.

2.3. Tumour models

CBCT scans from previous in vivo experiments were retrospectively analysed. Tumour xenograft studies were performed using the non-small cell lung cancer (NSCLC) cell lines, A549 and H460. Cells were cultured *in vitro* (Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal bovine serum and 1% penicillin/streptomycin) and prepared in phosphate-buffered saline (PBS) for subcutaneous injection into the flank of SCID mice. At 100 mm^3, tumours were imaged at 60 kV on the SARRP (n = 9 for each arm). All experimental procedures were carried out in accordance with the Home Office Guidance on the Operation of the Animals (Scientific Procedures Act 1986) (PPL2813).

2.4. Segmentation

Segmentations were created using ITK-SNAP software (version 3.8.0) [33]. Manual contours were created using the 3-D round brush in the abdominal region of the mouse phantom model (not including lung or bone). Standard spherical segmentations of 27.68, 34.38, 41.71, 92.24 and 237.5 mm^3 were used for scan-rescan analysis. Segmentation of tumours was completed using a standard spherical segmentation volume of 94.25 mm^3. This method was adopted to reduce the impact of inter-observer variabilities associated with manual contours [63].

2.5. Radiomics analysis

Radiomics analysis was performed using PyRadiomics (version 2.7.7, Harvard Medical School, Boston, MA, USA) [34], which is compliant with the Image Biomarker Standardisation Initiative (IBSI) [14]. 842 features were extracted including: shape (n = 14), first order statistics (n = 18), gray level cooccurrence matrix (GLCM) (n = 23), gray level run length matrix (GLRLM) (n = 16), gray level size zone matrix (GLSZM) (n = 16), gray level dependence matrix (GLDM) (n = 14) and neighbouring gray tone difference matrix (NGTDM) (n = 5). Wavelet filtering was also applied to these features. Shape features were only used for correlation analysis to segmentation volume.

To optimise our radiomics workflow, different pre-processing parameters were tested. The slice thickness of the CBCT scans were resampled to either 0.2, 0.26, 0.3, 0.5 or 1 mm by changing the "resampledPixelSpacing", without modifying the axial spacing. Image intensity discretization was performed to compare different fixed bin width values of 10, 25, 50 and 100 by altering the "binWidth".

2.6. Correlation to segmentation volume

Features highly correlated to volume changes was determined using correlation analysis (*cor* function within the *corrplot* library in RStudio software (version 4.1.2)). The Pearson correlation coefficient was calculated for each feature with respect to volume and a correlation coefficient > 0.8 applied.

2.7. Statistical analysis

The intraclass correlation coefficient (ICC) was used to determine the reliability and robustness of radiomics outputs through the production of a reliability index (Table 1). ICCs were calculated using the *irr* library from the *lpSolve* package in RStudio.

Reliability analysis was based on a single value with absolute-agreement and determined using 2-way mixed-effects models for the scan-rescan analysis of radiomics feature outputs across each variable [35]. Reproducibility analysis was based on an average of each scan and rescan (n = 6) with absolute-agreement and determined using 2-way mixed-effects models. Analysis was conducted between the tumour cohorts' through a 2-way mixed-effects ICC model. The Pearson correlation coefficient was also calculated for each feature (*cor* in RStudio) and a correlation coefficient > 0.8 was considered significant. Comparison of radiomics outputs for tumour models was performed using a paired *t*-test (two-tailed, $p < 0.05$) (n = 9). Analysis was performed using GraphPad Prism 7 (Version 7.0) with significance reported as p **** < 0.0001.

3. Results

3.1. Repeatability of preclinical radiomics features

Repeatability was assessed using scan-rescans of a mouse phantom acquired at different imaging energies or processed using different bin

Table 1
Classification of ICC results. Koo *et al* classifies ICC as poor (<0.5), moderate (0.5–0.7), good (0.7–0.9) and excellent (>0.9) [35,36]. A stricter ICC of > 0.8 was used to determine good/excellent reliability to better match with previous thresholds reported in test–retest analysis. The 95% confidence intervals (CIs) (>0.7) were used to remove errors and indicate robustness as recommended by Koo *et al*.

Intraclass correlation coefficient (ICC)	Reliability Index
0.8	Good reliability
>0.8	Excellent reliability
1	Perfect reliability
Classification of ICC in this study	
ICC > 0.8 & 95% confidence interval > 0.7	Highly robust

widths or slice thickness. This analysis aims to show how differences in the preclinical radiomics workflow may reduce the reliability of features. For imaging energies of 40, 50 and 60 kV there were 343, 420 and 388 reliable features respectively (ICC > 0.8) (Fig. 1 A). However, only 46, 53 and 57 features were robust (lower CI of the ICC > 0.7). Scans acquired at 40 kV had the greatest variability; potentially due to increased artefacts and noise in scans. Only 10 robust features (1%) overlapped across all 3 imaging energies; all of which were first order features (Fig. 1 A). CBCT scans acquired at different imaging energies can therefore limit the number of robust radiomics features for comparative analysis and the higher energies (60 kV) recommended for analysis.

Repeatability of features across different bin widths was compared at 60 kV (Fig. 1 B). Bin widths of 25 and 50 had the most robust features of 57 and 58 (7%) respectively, 43 of which were shared. Fig. 1 B includes a heatmap of the 31 (4%) robust and reliable features maintained across all bin widths.

Radiomics features were extracted with a resampled slice thickness of 0.2, 0.26, 0.3, 0.5 or 1 mm (Fig. 1 C). A slice thickness of 0.5 mm had the most robust features of 78 (9%). Only 12 (1%) overlapping features were identified across all slice thicknesses, all of which were first order. Additional analysis showed that increasing the slice thickness led to variability in shape and volume analysis ("original_shape_MeshVolume").

3.2. The volume effect

To determine if volume impacts feature reliability or if there is a minimum volume suitable for extracting reliable results, we compared radiomics outputs for a range of volumes in a mouse phantom model. As preclinical models are smaller than their clinical counterparts five relevant volumes for preclinical analysis were used (28, 34, 42, 92 and

238 mm^3) (Supplementary Fig. 2). The smallest volume, 28 mm^3, had the least repeatable features (101 features), in comparison, larger volumes of 92 and 238 mm^3 had 388 and 381 repeatable features respectively (Fig. 2 B). There was no overlap in robust features across the range of segmentation volumes evaluated. Supplementary Fig. 3 details overlapping features amongst similar volumes. These results suggest that volumes < 34 mm^3 may be too small to extract reliable data.

The number of robust features did not increase with increasing segmentation volume. Volumes of 42 and 92 mm^3 had the most robust and reliable features of 119 and 57 features respectively (Fig. 2 C). A volume range of 42–92 mm^3 may be suitable for preclinical radiomics analysis with 32 features maintained for both volumes (Fig. 2 D). These non-linear results may be influenced by the phantom model used in which we assume tissue regions are homogeneous. Our results show that first order and GLDM features have a higher reliability range when comparing different volumes. Whereas GLCM, GLSZM and NGTDM features are more sensitive to volume changes (Supplementary Fig. 4).

The correlation of segmentation volume to unfiltered radiomics features is shown in Fig. 2 E. Fifty-four features were highly correlated to an increase in segmentation volume (original_shape_MeshVolume). These included 9 shape, 12 first order, 13 GLCM, 6 GLRLM, 6 GLSZM, 6 GLDM and 2 NGTDM features (Supplementary Table 2). Of these, 7 have been determined as reliable features from scan-rescan analysis for volumes of 42 – 92 mm^3.

A workflow of scans imaged at 60 kV and features extracted at bin width of 25 and slice thickness maintained at 0.26 mm was determined. From repeatability analysis 119 (14%) robust features can be extracted at 42 mm^3 (Supplementary Table 3) and 57 (7%) robust features at 92 mm^3 (Supplementary Table 4) which are stable for preclinical analysis.
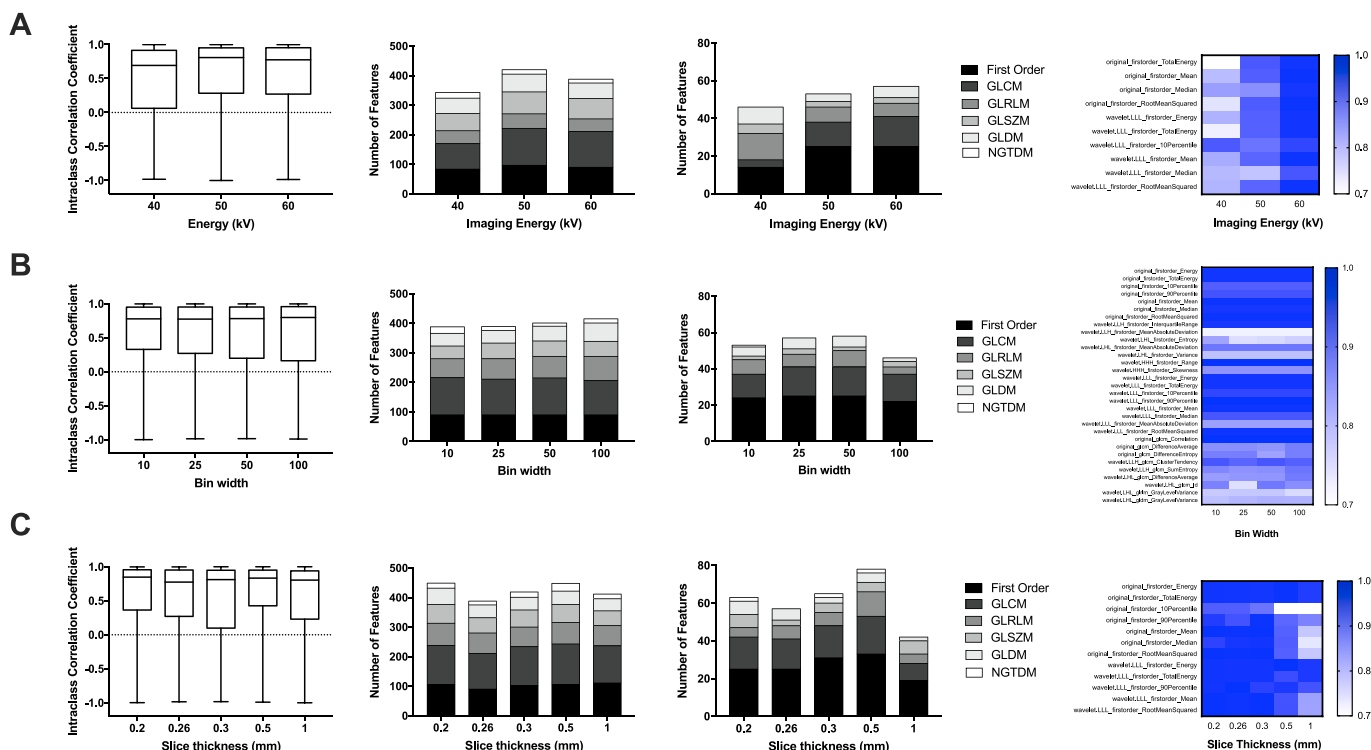

Fig. 1. Reliability and robustness of radiomics features with varying CBCT image acquisition and image discretisation methods. CBCT scans of a 3-D mouse phantom were acquired on the SARRP and analysed using PyRadiomics. Boxplots display ICC values of radiomics features (left). The number of reliable radiomics features (ICC > 0.8) (middle-left). The number of robust radiomics features (lower CI > 0.7) (middle-right). Heatmap of ICC values for overlapping robust features (right). Panel A: Reliability of radiomics features across imaging energies of 40, 50 and 60 kV. Panel B: Reliability of radiomics features after changing the intensity discretization via bin width to 10, 25, 50 or 100. Panel C: Reliability of radiomics features after changing the slice thickness during analysis (0.2, 0.26, 0.3, 0.5 & 1 mm).
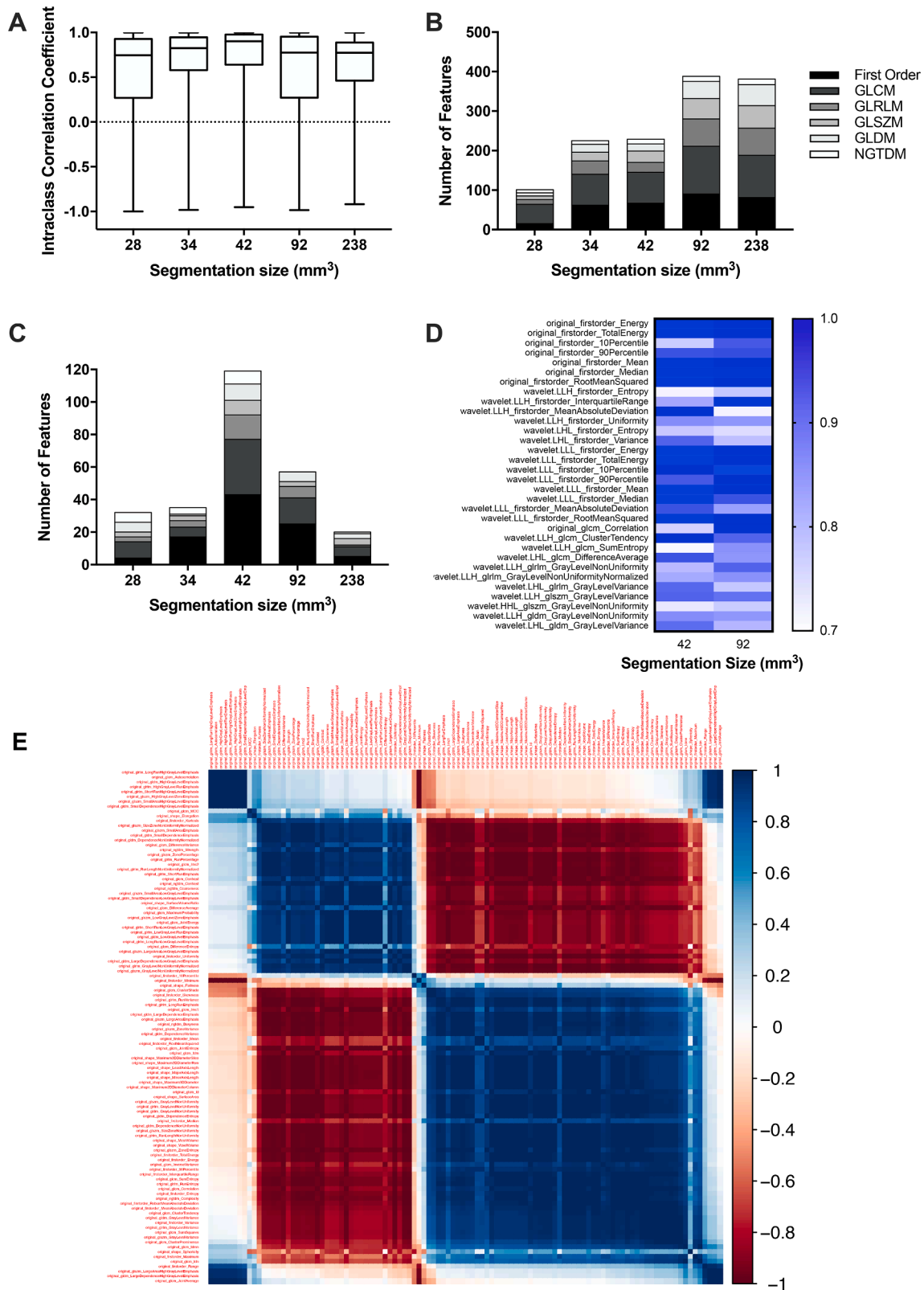
Fig. 2. Overview of radiomics outputs for a range of segmentation volumes. Panel A: Boxplots to display ICC values of radiomics features assessed across a range of segmentation volumes (28 – 238 mm³). Panel B: The number of reliable radiomics features by feature class for each segmentation volume. Panel C: The number of robust features for each segmentation volume Panel D: Heatmap of overlapping robust features for 42 and 92 mm³ segmentation volumes. Panel E: Hierarchical correlation matrix to identify unfiltered radiomics features that are highly correlated to an increase in segmentation volume. 54 unfiltered features were highly correlated to changes in the segmentation volume.

3.3. Reproducibility of preclinical radiomics features

To further optimise our results, we assessed the reproducibility of radiomics outputs. Changing the imaging energy had the biggest impact on the reproducibility of features with only 2 features identified. Altering the slice thickness resulted in 45 reproducible features. Variations in the bin width and segmentation sizes were least affected with 176 and 183 reproducible features respectively (Fig. 3 A). Overall, the most reproducible feature types were first order, GLCM and GLRLM.

No robust features overlapped from repeatability and reproducibility studies for varying imaging energies; however, there was an overlap of 45, 16 and 31 features for bin width, slice thickness and segmentation size respectively (Fig. 3 B). These features are therefore highly conserved for comparison of preclinical radiomics outputs when using different workflow parameters (Supplementary Table 5).

3.4. Texture analysis

A multi-density phantom was used to measure the variability of radiomics features to changes in texture. Bart's solid water (1.05 g/cm^3) and the mouse phantom (1.01 g/cm^3) have similar densities and visually look similar from CBCT scans yet the average gray level intensity (original_firstorder_Mean) values differ from 2,940 to 16,844 (Fig. 4 A). Scan-rescan analysis was conducted with ICC outputs for wavelet features shown in Fig. 4 B. GLSZM features had the lowest median ICC for all textures apart from acetal (Fig. 4 B). NGTDM features were further analysed and shown to be influenced by changes in density (Fig. 4 C). This confirms that preclinical radiomics analysis can be used to differentiate materials with differing density through textural radiomics analysis.

3.5. Differentiation of tumour models using radiomics features: Pilot analysis

Pre-treatment CBCT scans from two NSCLC tumour models were retrospectively analysed (Fig. 5 A). There were 773 and 776 highly correlated features for A549 and H460 tumours respectively with 731 shared (Fig. 5B). Test-retest analysis identified 26 and 89 reliable features for the A549 and H460 cohorts respectively (Fig. 5 C/D). After comparison with robust features (Supplementary Table 5), 4 features can be used to differentiate A549 and H460 tumours on preclinical CBCT scans (Fig. 5 E).

4. Discussion

Since the first application of radiomics analysis for phenotype prediction, it has led to the discovery of imaging biomarkers and evolved to include multiple imaging modalities [1,14,23,36,37]. Radiomics analysis also has major clinical and economic benefits for the replacement of invasive and expensive procedures to determine tumour heterogeneity, such as biopsies [38]. Yet, real-world application of radiomics in oncology is limited by the lack of "big" and standardised clinical data due to different imaging protocols, variability in patient history and restrictions by law and ethics [39].

Mouse models are hugely beneficial in radiation oncology for the understanding of cancer progression and treatment development [40]. In addition, preclinical radiomics analysis has been successful using preclinical CT and MR scans [24,25,41]. Despite evidence that mouse models can expand our knowledge in radiomics signatures, there are currently no established guidelines to ensure consistency in preclinical analysis [28]. We aimed to optimise and standardise the first preclinical CBCT-radiomics workflow to improve the accuracy and reproducibility of outputs.

A typical radiomics workflow includes 4 main steps: image acquisition, tissue delineation, feature extraction and analysis. Clinical studies have shown that changes to these can reduce the number of robust features to 6 – 43% [14,42]. Some steps depend on expertise (tissue delineation) or research question (analysis), but others can be standardised (image acquisition and feature extraction) [43,44]. We have shown preclinical analysis to be more sensitive to these changes with 0.2–22% robust features identified.

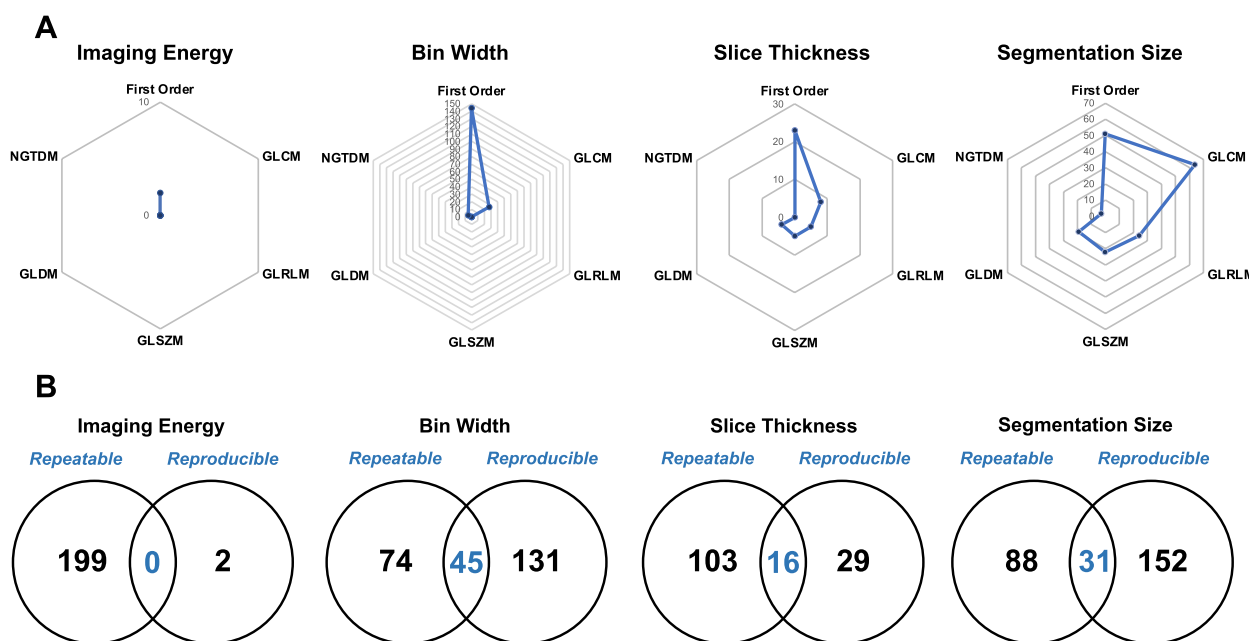Preclinical CBCT scans are acquired at lower energies than used

Fig. 3. Results of the reproducibility analysis for preclinical radiomics. Panel A: Reproducibility of radiomics outputs was compared within each variable and the number of features with a good ICC (>0.8) was plotted for imaging energy, bin width, slice thickness and segmentation size. Panel B: Venn diagrams to show the overlap of repeatable and reproducible features extracted from a mouse phantom using varying preclinical radiomics analysis methods. Repeatable features include the 119 robust features detailed in Table 2.
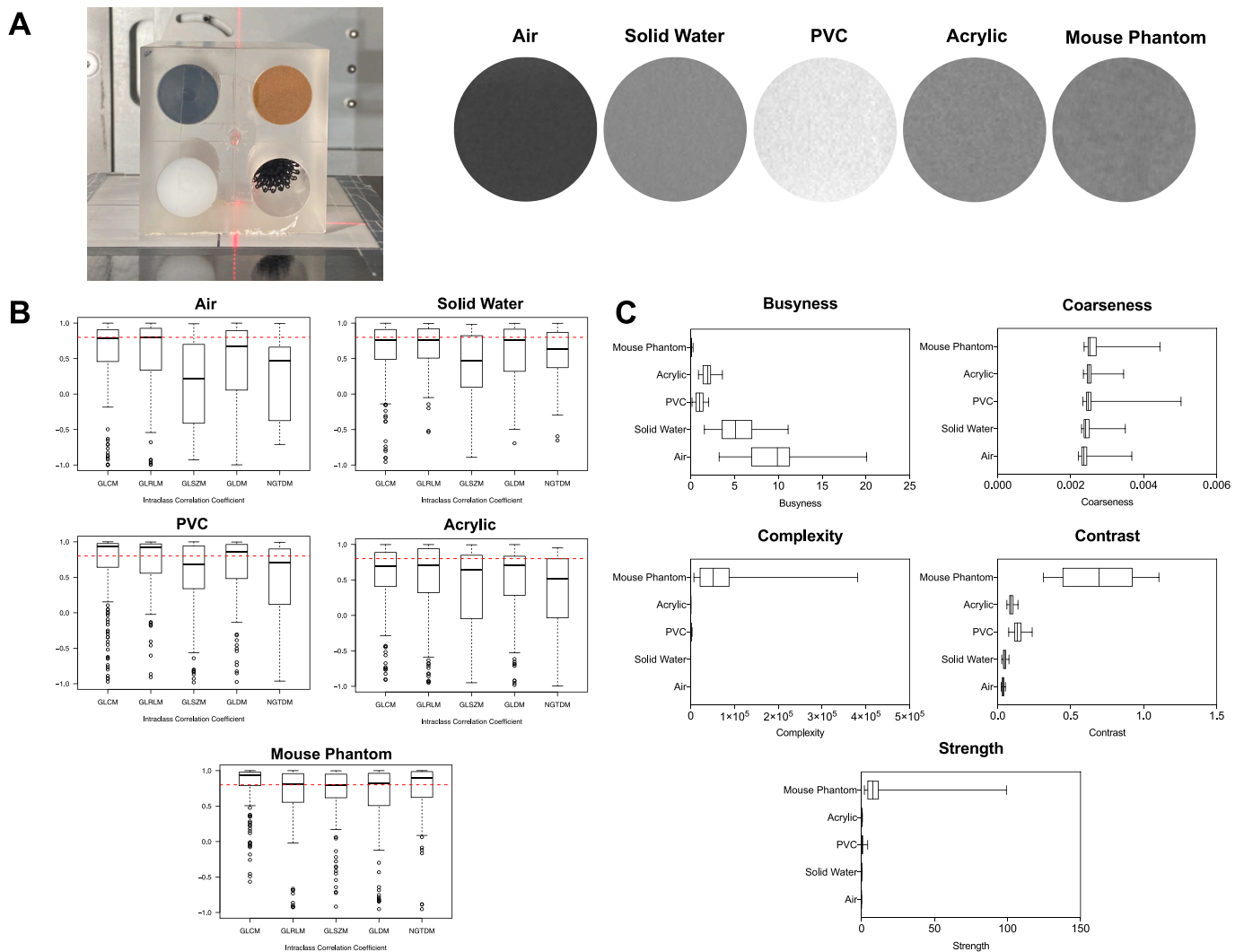
Fig. 4. Radiomics features are affected by changes in texture. Panel A: Textural phantom on imaging bed with labels for each textural insert (left). CBCT cross section of the 4 different cylindrical inserts and the mouse phantom (right). The average gray level intensity (original_firstorder_Mean) for each material was 1861 for air, 2940 for solid water, 4138 for PVC, 2917 for acetal and 16,844 for the mouse phantom. Panel B: Boxplots of ICC outputs for wavelet radiomics features across textures at a segmentation volume of 42 mm^3. Panel C: NGTDM feature values for air, solid water, PVC, acetal and the mouse phantom.

clinically [15,45], CBCT scan quality is known to have scattering and beam hardening artefacts in comparison to CT scans causing additional variabilities between scans [20]. Reduction of variabilities during image acquisition was achieved through use of a single, high imaging energy (60 kV). Advanced imaging methods such as dual-energy CT (DECT) improve image quality and could potentially reduce variabilities in radiomics analysis. However, imaging doses associated with preclinical DECT (60 cGy) are higher than single energy exposures (2.4 cGy) and repeated longitudinal imaging may have increased biological implications [46,47].

Studies also recommend standardising image intensity discretisation through bin widths as a normalisation step for comparative analysis [48–50]. A fixed bin width was used for intensity discretisation for filtered features [51]. Our analysis identified bin widths of 25 or 50 to have the most robust features for analysis. Changing the slice thickness or pixel size can also reduce the impact of noise within the scans for the extraction of more reproducible and robust features [48]. First order, GLCM and GLRLM feature classes were the most robust to changes in slice thickness in agreement with other studies [48,52]. However, altering the slice thickness during analysis caused changes to shape features which could significantly impact analysis. Further normalisation methods may be of interest for future preclinical radiomics

studies [53,54].

Studies have shown different segmentation volumes have a more significant effect on CT-derived features than MR- features [55]. [56–58] Roy *et al* showed that volume size had the largest influence on GLSZM features followed by GLCM, GLRLM and NGTDM features [28]. Some clinical analysis excludes tumours if they have a volume under a defined limit [56–58]. Segmentation volumes are typically smaller in preclinical models making them more challenging to delineate and contain fewer voxels or quantitative information for analysis. Our study is the first to evaluate the volume effect on preclinical radiomics outputs. Similar to clinical results, GLCM, GLSZM and NGTDM features were affected the most by changes in volume. As some features classes are more heavily influenced or dependent on volume to maximise reliability, first order and GLDM features should be used for analysis, or similar segmentation volumes should be compared [28,56].

In clinical analysis, tumour volume has been shown to complement texture analysis of intra-tumoral heterogeneity [57]. Our results have determined 54 features highly correlated to changes in volume (Supplementary Table 2). Removing features dependent on volume changes should therefore be excluded from studies assessing tumour heterogeneity.

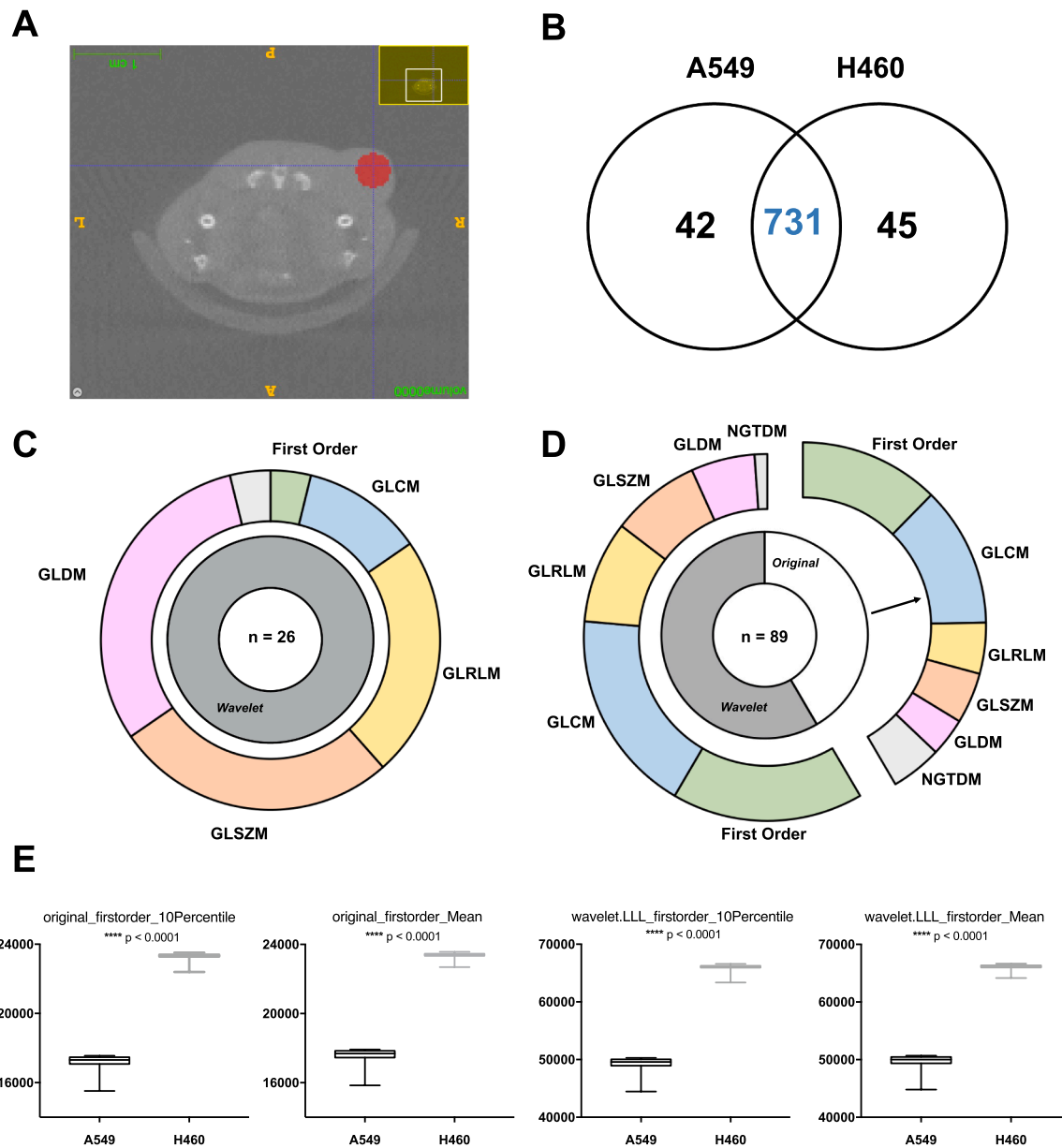Phantoms are invaluable to radiation research to mimic tissue

Fig. 5. Application of radiomics analysis to preclinical CBCT scans of lung tumour models (A549 and H460). Panel A: Example of pre-treatment CBCT scan acquired at 60 kV used for analysis. An example of the spherical segmentation can be visualised in red. Panel B: Venn diagram of highly correlated features overlapping between tumour cohorts. Panel C: Schematic to represent the 26 reliable radiomics features for A549 tumours (ICC > 0.8) subdivided by feature type and class. Panel D: The 89 reliable radiomics features for H460 tumours (ICC > 0.8) broken down by feature type and class. Panel E: Example of 4 repeatable and reproducible radiomics features which can be used to differentiate the two tumour cohorts. Significance reported as p **** < 0.0001. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

texture and density without repeated imaging dose to human or animal subjects [59]. Through the inclusion of a density phantom, similar to that of soft tissue (solid water) and bone (PVC), we demonstrated preclinical radiomics can differentiate between density changes. NGTDM features were further analysed as understandable texture properties [60,61]. The creation of a dedicated preclinical radiomics phantom with differing densities and textural components may be more applicable for comparison of texture outputs with tissue equivalents.

Whilst our study provides a thorough analysis of robust and reliable features for preclinical radiomics, it has several limitations. Shape features was excluded from the repeatability and reproducibility analysis to remove user bias from manual contouring methods. Results from tumour models only provide proof of principle in extracting useful information from preclinical scans with additional analysis required to correlate

features to biological parameters. This study is the first effort to optimise and standardise preclinical CBCT-radiomics analysis with further scope to compare radiomics outputs between research centres and across imaging modalities [62].

We present the first preclinical CBCT-radiomics workflow comparing changes to the repeatability and reproducibility of features across image acquisition, pre-processing parameters and segmentation sizes. Our results recommend that preclinical CBCT scans should be acquired at higher imaging energy (60 kV) and features extracted using a set bin width (25) and slice thickness (0.26 mm). Feasibility of extracting meaningful data was validated in a multi-texture phantom and preclinical models of NSCLC. Our data demonstrates that preclinical radiomics analysis is a novel tool that has the potential to develop imaging biomarkers to support the wider application of radiomics.

Funding

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2023.100446.

References

[1] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014:5. https://doi.org/10.1038/ncomms5006.

[2] O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14:169–86. https://doi.org/10.1038/nrclinonc.2016.162.

[3] Verhaegen F, Butterworth KT, Chalmers AJ, Coppes RP, de Ruysscher D, Dobiasch S, et al. Roadmap for Precision preclinical x-ray radiation studies. Phys Med Biol 2022. https://doi.org/10.1088/1361-6560/acaf45.

[4] Tillner F, Thute P, Butof R, Krause M, Enghardt W. Pre-clinical research in small animals using radiotherapy technology–a bidirectional translational approach. Z Med Phys 2014;24:335–51. https://doi.org/10.1016/j.zemedi.2014.07.004.

[5] Verhaegen F, Granton P, Tryggestad E. Small animal radiotherapy research platforms. Phys Med Biol 2011;56:R55–83. https://doi.org/10.1088/0031-9155/56/12/R01.

[6] Brown KH, Ghita M, Dubois LJ, de Ruysscher D, Prise KM, Verhaegen F, et al. A scoping review of small animal image-guided radiotherapy research: Advances, impact and future opportunities in translational radiobiology. Clin Transl Radiat Oncol 2022;34:112–9. https://doi.org/10.1016/j.ctro.2022.04.004.

[7] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62. https://doi.org/10.1038/nrclinonc.2017.141.

[8] Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol 2016;120:258–66. https://doi.org/10.1016/j.radonc.2016.05.024.

[9] Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. Radiology 2021;298:505–16. https://doi.org/10.1148/radiol.2021202553.

[10] Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: A potential marker of survival. Eur Radiol 2012;22:796–802. https://doi.org/10.1007/s00330-011-2319-8.

[11] Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, et al. Defining the biological basis of radiomic phenotypes in lung cancer. Elife 2017:6. https://doi.org/10.7554/eLife.23421.

[12] Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: The need for safeguards. Radiother Oncol 2019;130:2–9. https://doi.org/10.1016/j.radonc.2018.10.027.

[13] Leijenaar RTH, Nalbantov G, Carvalho S, van Elmpt WJC, Troost EGC, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: The need for standardized methodology in tumor texture analysis. Sci Rep 2015:5. https://doi.org/10.1038/srep11075.

[14] Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:328–38. https://doi.org/10.1148/radiol.2020191145.

[15] Berenguer R, del Rosario P-J, Canales-Vázquez J, Castro-García M, Villas MV, Legorburo FM, et al. Radiomics of CT features may be nonreproducible and redundant: Influence of CT acquisition parameters. Radiology 2018;288:407–15. https://doi.org/10.1148/radiol.2018172361.

[16] Walls GM, Osman SOS, Brown KH, Butterworth KT, Hanna GG, Hounsell AR, et al. Radiomics for predicting lung cancer outcomes following radiotherapy: a systematic review. Clin Oncol 2021. https://doi.org/10.1016/j.clon.2021.10.006.

[17] Fried Dv, Tucker SL, Zhou S, Liao Z, Mawlawi O, Ibbott G, et al. Prognostic value and reproducibility of pretreatment ct texture features in stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2014;90:834–42. https://doi.org/10.1016/j.ijrobp.2014.07.020.

[18] Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol 2015;114:345–50. https://doi.org/10.1016/j.radonc.2015.02.015.

[19] Spuhler KD, Teruel JR, Galavis PE. Assessing the reproducibility of CBCT-derived radiomics features using a novel three-dimensional printed phantom. Med Phys 2021;48:4326–33. https://doi.org/10.1002/mp.15043.

[20] Delgadillo R, Spieler BO, Ford JC, Kwon D, Yang F, Studenski M, et al. Repeatability of CBCT radiomic features and their correlation with CT radiomic features for prostate cancer. Med Phys 2021;48:2386–99. https://doi.org/10.1002/mp.14787.

[21] van Timmeren JE, Leijenaar RTH, van Elmpt W, Reymen B, Lambin P. Feature selection methodology for longitudinal cone-beam CT radiomics. Acta Oncol 2017;56:1537–43. https://doi.org/10.1080/0284186X.2017.1350285.

[22] Shi L, Rong Y, Daly M, Dyer B, Benedict S, Qiu J, et al. Cone-beam computed tomography-based delta-radiomics for early response assessment in radiotherapy for locally advanced lung cancer. Phys Med Biol 2020:65. https://doi.org/10.1088/1361-6560/ab3247.

[23] van Timmeren JE, van Elmpt W, Leijenaar RTH, Reymen B, Monshouwer R, Bussink J, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: evaluation of the added prognostic value for overall survival and locoregional recurrence. Radiother Oncol 2019;136:78–85. https://doi.org/10.1016/j.radonc.2019.03.032.

[24] Panth KM, Leijenaar RTH, Carvalho S, Lieuwes NG, Yaromina A, Dubois L, et al. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. Radiother Oncol 2015;116:462–6. https://doi.org/10.1016/j.radonc.2015.06.013.

[25] Eresen A, Yang J, Shangguan J, Li Y, Hu S, Sun C, et al. MRI radiomics for early prediction of response to vaccine therapy in a transgenic mouse model of pancreatic ductal adenocarcinoma. J Transl Med 2020;18:1–9. https://doi.org/10.1186/s12967-020-02246-7.

[26] O'Farrell AC, Jarzabek MA, Lindner AU, Carberry S, Conroy E, Miller IS, et al. Implementing systems modelling and molecular imaging to predict the efficacy of BCL-2 inhibition in colorectal cancer patient-derived xenograft models. Cancers (Basel) 2020;12:1–19. https://doi.org/10.3390/cancers12102978.

[27] Becker AS, Schneider M, Wurnig MC, Wagner M, Clavien PA, Boss A. Radiomics of liver MRI predict metastases in mice. Eur Radiol Exp 2018:2. https://doi.org/10.1186/s41747-018-0044-7.

[28] Roy S, Whitehead TD, Quirk JD, Salter A, Ademuyiwa FO, Li S, et al. Optimal co-clinical radiomics: Sensitivity of radiomic features to tumour volume, image noise and resolution in co-clinical T1-weighted and T2-weighted magnetic resonance imaging. EBioMedicine 2020:59. https://doi.org/10.1016/j.ebiom.2020.102963.

[29] Ghita M, Brown KH, Kelada OJ, Graves EE, Butterworth KT. Integrating small animal irradiators with functional imaging for advanced preclinical radiotherapy research. Cancers (Basel) 2019:11. https://doi.org/10.3390/cancers11020170.

[30] Jha AK, Mithun S, Jaiswar V, Sherkhane UB, Purandare NC, Prabhash K, et al. Repeatability and reproducibility study of radiomic features on a phantom and human cohort. Sci Rep 2021:11. https://doi.org/10.1038/s41598-021-81526-8.

[31] Soultanidis G, Subiel A, Renard I, Reinhart AM, Green VL, Oelfke U, et al. Development of an anatomically correct mouse phantom for dosimetry measurement in small animal radiotherapy research. Phys Med Biol 2019:64. https://doi.org/10.1088/1361-6560/ab215b.

[32] Silvestre Patallo I, Subiel A, Westhorpe A, Gouldstone C, Tulk A, Sharma RA, et al. Development and implementation of an end-to-end test for absolute dose verification of small animal preclinical irradiation research platforms. Int J Radiat Oncol Biol Phys 2020;107:587–96. https://doi.org/10.1016/j.ijrobp.2020.03.001.

[33] Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31:1116–28. https://doi.org/10.1016/j.neuroimage.2006.01.015.

[34] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339.

[35] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.

[36] Vallières M, Freeman CR, Skamene SR, el Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 2015;60:5471–96. https://doi.org/10.1088/0031-9155/60/14/5471.

[37] Defeudis A, de Mattia C, Rizzetto F, Calderoni F, Mazzetti S, Torresin A, et al. Standardization of CT radiomics features for multi-center analysis: impact of software settings and parameters. Phys Med Biol 2020;65:195012. https://doi.org/10.1088/1361-6560/ab9f61.

[38] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6. https://doi.org/10.1016/j.ejca.2011.11.036.

[39] Rogers W, Seetha T, Refaee S, Lieverse TIY, Granzier RWY, Ibrahim R. Radiomics: from qualitative to quantitative imaging. vol. 93. 2020.

[40] Butterworth KT. Evolution of the supermodel: progress in modelling radiotherapy response in mice. Clin Oncol (R Coll Radiol) 2019;31:272–82. https://doi.org/10.1016/j.clon.2019.02.008.

[41] Holbrook MD, Blocker SJ, Mowery YM, Badea A, Qi Y, Xu ES, et al. Mri-based deep learning segmentation and radiomics of sarcoma in mice. Tomography 2020;6: 23–33. https://doi.org/10.18383/j.tom.2019.00021.

[42] Meyer M, Ronald J, Vernuccio F, Nelson RC, Ramirez-Giraldo JC, Solomon J, et al. Reproducibility of CT radiomic features within the same patient: Influence of radiation dose and CT reconstruction settings. Radiology 2019;293:583–91. https://doi.org/10.1148/radiol.2019190928.

[43] Fave X, MacKin D, Yang J, Zhang J, Fried D, Balter P, et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? Med Phys 2015;42:6784–97. https://doi.org/10.1118/1.4934826.

[44] Wang H, Zhou Y, Wang X, Zhang Y, Ma C, Liu B, et al. Reproducibility and repeatability of CBCT-derived radiomics features. Front. Oncol. 2021:11. https://doi.org/10.3389/fonc.2021.773512.

[45] Ligero M, Jordi-Ollero O, Bernatowicz K, Garcia-Ruiz A, Delgado-Muñoz E, Leiva D, et al. Minimizing acquisition-related radiomics variability by image resampling and batch effect correction to allow for large-scale data analysis. Eur. Radiol. 2021;31:1460–70. https://doi.org/10.1007/s00330-020-07174-0.

[46] Vaniqui A, Schyns LEJR, Almeida IP, van der Heyden B, van Hoof SJ, Verhaegen F. The impact of dual energy CT imaging on dose calculations for pre-clinical studies. Radiat Oncol 2017:12. https://doi.org/10.1186/s13014-017-0922-9.

[47] Schyns LEJR, Almeida IP, Van Hoof SJ, Descamps B, Vanhove C, Landry G, et al. Optimizing dual energy cone beam CT protocols for preclinical imaging and radiation research. Br J Radiol 2017:90. https://doi.org/10.1259/bjr.20160480.

[48] Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Med Phys 2017;44:1050–62. https://doi.org/10.1002/mp.12123.

[49] Palani D, Shanmugam S, Govindaraj K. Analyzing the possibility of utilizing CBCT radiomics as an independent modality: a phantom study. Asian Pac J Cancer Prev 2021;22:1383–91. https://doi.org/10.31557/APJCP.2021.22.5.1383.

[50] Bagher-Ebadian H, Siddiqui F, Liu C, Movsas B, Chetty IJ. On the impact of smoothing and noise on robustness of CT and CBCT radiomics features for patients with head and neck cancers. Med Phys 2017;44:1755–70. https://doi.org/10.1002/mp.12188.

[51] Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative 2016. doi: 10.1148/radiol.2020191145.

[52] Larue RTHM, van Timmeren JE, de Jong EEC, Feliciani G, Leijenaar RTH, Schreurs WMJ, et al. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. Acta Oncol (Madr) 2017;56:1544–53. https://doi.org/10.1080/0284186X.2017.1351624.

[53] Park D, Oh D, Lee MH, Lee SY, Shin KM, Jun JS, et al. Importance of CT image normalization in radiomics analysis: prediction of 3-year recurrence-free survival in non-small cell lung cancer. Eur Radiol 2022;32:8716–25. https://doi.org/10.1007/s00330-022-08869-2.

[54] Shafiq-Ul-Hassan M, Latifi K, Zhang G, Ullah G, Gillies R, Moros E. Voxel size and gray level normalization of CT radiomic features in lung cancer. Sci Rep 2018:8. https://doi.org/10.1038/s41598-018-28895-9.

[55] Jensen LJ, Kim D, Elgeti T, Steffen IG, Hamm B, Nagel SN. Stability of radiomic features across different region of interest sizes-A CT and MR phantom study. Tomography 2021;7:238–52. https://doi.org/10.3390/tomography7020022.

[56] Zhang X, Zhong L, Zhang B, Zhang L, Du H, Lu L, et al. The effects of volume of interest delineation on MRI-based radiomics analysis: Evaluation with two disease groups. Cancer Imaging 2019:19. https://doi.org/10.1186/s40644-019-0276-7.

[57] Hatt M, Majdoub M, Vallières M, Tixier F, le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: Investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. J Nucl Med 2015;56:38–44. https://doi.org/10.2967/jnumed.114.144055.

[58] Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol 2016;61:R150–66. https://doi.org/10.1088/0031-9155/61/13/R150.

[59] McGarry CK, Grattan LJ, Ivory AM, Leek F, Liney GP, Liu Y, et al. Tissue mimicking materials for imaging and therapy phantoms: a review. Phys Med Biol 2020:65. https://doi.org/10.1088/1361-6560/abbd17.

[60] Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern 1989;19:1264–74. https://doi.org/10.1109/21.44046.

[61] Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. Invest Radiol 2015;50: 757–65. https://doi.org/10.1097/RLI.0000000000000180.

[62] van Timmeren JE, Leijenaar RTH, van Elmpt W, Wang J, Zhang Z, Dekker A, et al. Test-retest data for radiomics feature stability analysis: generalizable or study-specific? Tomography 2016;2:361–5. https://doi.org/10.18383/j.tom.2016.00208.

[63] Brown KH, Illyuk J, Ghita M, Walls GM, McGarry CK, Butterworth KT. Assessment of Variabilities in Lung-Contouring Methods on CBCT Preclinical Radiomics Outputs. Cancers (Basel) 2023;15:2677. https://doi.org/10.3390/cancers15102677.