scientific reports

Check for updates

OPEN

Reliability of radiomic analysis on multiparametric MRI for patients affected by autosomal dominant polycystic kidney disease

Francesca Lussana[1], Ettore Lanzarone[1], Giulia Villa[2], Alfonso Mastropietro[3], Anna Caroli[2] & Elisa Scalco[4✉]

Autosomal dominant polycystic kidney disease (ADPKD) is a prevalent hereditary disorder characterized by the development and growth of fluid-filled cysts, resulting in a decline in kidney function. Beyond total kidney and cyst volume quantification, non-cystic tissue characterization by multi-parametric MRI (mp-MRI) and radiomics holds promise. We conducted a radiomic analysis based on reproducible and informative features extracted from non-cystic tissue on mp-MRI in ADPKD patients. T2-weighted (T2-w), T1-weighted MRI (T1-w), and IntraVoxel Incoherent Motion (IVIM) maps from Diffusion Weighted Imaging (DWI) were considered. The reliability of radiomic features was evaluated using five different segmentation methods. The impact of segmentation variability on radiomic reproducibility was assessed through Intraclass Correlation Coefficients (ICC), and a preliminary correlation analysis with relevant clinical parameters, such as age and eGFR, was also performed. The results from 14 patients indicate that radiomic features derived from IVIM maps exhibit greater reliability compared to features from T1-w and T2-w for characterizing non-cystic tissue in ADPKD patients, also showing a moderate correlation with age and eGFR. Additionally, lower-order features, including those computed from histograms and co-occurrence matrices, demonstrate higher reproducibility than other texture features.

Keywords ADPKD, mp-MRI, Radiomics

Autosomal dominant polycystic kidney disease (ADPKD) stands as a prevalent hereditary kidney disorder, often leading to end-stage renal disease in adults. ADPKD is denoted by the formation and progressive enlargement of fluid-filled cysts within the kidneys, and often the liver[1]. Disease monitoring and prognosis assessment usually rely on metrics such as total kidney and cyst volume. Total kidney volume (TKV), often corrected for the patient's height (the so-called height-adjusted TKV[1]) is crucial to assess the stage of ADPKD because it reflects the decompensation of renal parenchyma[2]. The strong link between height-adjusted TKV and disease progression is specified by the Mayo Imaging Classification, which classifies patients according to the risk of progression. While TKV is recognized as an indicator of disease progression, focusing specifically on cyst burden may offer a more precise understanding of the disease's condition[3]. Beyond TKV and cyst volume, non-cystic tissue appears to play a key role in ADPKD pathology; in fact, non-cystic volume has been shown to be characterized by peritubular interstitial fibrosis, microcysts, and dilated tubules that stand as important factors in the course of the disease[1].

Imaging techniques currently applied to diagnose and monitor ADPKD include ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI). Although all are excellent at visualizing the cystic and non-cystic components in patients, MRI is the most advantageous because it allows tissue segmentation and quantitative analysis of the severity of the pathology[4]. Multiparametric non-contrast-enhanced MRI (mp-MRI), which combines anatomical with microstructural and functional sequences, provides a non-invasive means to discern and characterize the non-cystic component, offering information on morphological and microstructural properties[3]. In particular, Diffusion-Weighted Imaging (DWI) allows for the assessment

[1]Department of Management, Information and Production Engineering, University of Bergamo, 24044 Dalmine, BG, Italy. [2]Bioengineering Department, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, 24020 Ranica, BG, Italy. [3]Institute of Intelligent Industrial Technologies and Systems, Italian National Research Council (STIIMA-CNR), 20133 Milan, Italy. [4]Institute of Biomedical Technologies, Italian National Research Council (ITB-CNR), 20054 Segrate, MI, Italy. ✉email: elisa.scalco@cnr.it

of ADPKD non-cystic tissue diffusion and perfusion using the IntraVoxel Incoherent Motion (IVIM) model[5]. IVIM is a DWI technique proposed by Le Bihan et al.[6], based on the acquisition of multiple b-values, which evaluates diffusion and pseudo-diffusion tissue characteristics by computing three parametric maps: D for the true diffusion, D^* for the pseudo-diffusion, and f for the perfusion fraction. On the other hand, anatomical MRI, including T2-weighted (T2-w) and T1-weighted (T1-w) sequences, enables the visualization and quantification of renal cysts[5].

From a clinical standpoint, TKV and cyst volume segmentation on radiological scans is crucial to guide therapy. Manual segmentation is currently the gold standard technique, but, despite its accuracy, it is extremely time-consuming. For this reason, several studies have addressed the automatic segmentation of kidneys and cysts on radiological scans in ADPKD patients[7–9]. Similarly to the total kidney, in recent years several automatic methods have been proposed to segment cysts, either on CT[10] or MR images[11,12]. Segmentation of the newly interesting non-cystic component is either done manually[2] or automatically on MRI images[5].

Radiomics, an evolving research domain, integrates high-throughput extraction of different features from medical images with machine learning techniques to construct classification and prediction models[13]. Radiomic analysis has previously been explored in the context of ADPKD. Li et al.[14] performed a radiomic analysis of T1-w and T2-w images to assess kidney tissue composition, including cysts, and kidney function decline. Calvaruso et al.[15] implemented a radiomic model on CT scans of patients with ADPKD to assess radiomic characteristics that can predict the decline in renal function over time. Additionally, Xie et al.[2] found that radiomics is crucial for identifying the decrease in renal parenchyma volume, closely associated with a decline in renal function (impairment rate) and eGFR. However, a primary challenge in radiomics lies in the sensitivity of features to various sources of variability within the radiomic workflow, covering image acquisition, processing, feature computation, and Region of Interest (ROI) delineation[16]. Specifically, variability in segmentation may significantly impact radiomic reliability, especially in delineating challenging regions such as non-cystic tissue, as evidenced in different works in the literature[16,17]. To guarantee a robust assessment of non-cystic tissue and thus be able to detect subtle variations in non-cystic microstructure across patients, it is essential to base radiomic analysis on stable and informative features, ensuring that observed differences reflect biological variations rather than inconsistencies in the feature extraction pipeline.

This study aimed to identify a set of radiomic features that remain reproducible across different reliable segmentation methods for non-cystic tissue while capturing meaningful microstructural differences between ADPKD patients using mp-MRI, encompassing both anatomical MRI and IVIM parametric maps. Specifically, different segmentation methods for non-cystic regions, including manual and automated approaches, were examined as a key source of variability in the radiomic pipeline. Building upon deep learning (DL)-based methods initially designed for cyst identification on T2-w images, and subsequently refined, our investigation proposes a novel framework to evaluate radiomics' reliability in terms of its capability to detect reproducible and informative features.

Methods
Dataset
The dataset used included 14 ADPKD patients from a prospective longitudinal multi-centric clinical study (CYSTic I—An international consortium to build a longitudinal observational cohort of patients with Autosomal Dominant PolyCYSTic Kidney Disease (ADPKD)). Details about demographic and clinical information are reported in Table 1. All subjects provided written informed consent upon enrollment in the clinical study. For this research, it was necessary to share the data from the Bergamo participating center in accordance with the Cystic Consortium's data sharing guidelines. The present study was approved by the local Ethics Committee "Comitato Etico Territoriale Lombardia 6" (Reg. N. 2024-3.11/486), and it was performed in accordance with the Declaration of Helsinki.

Image acquisition
Multi-parametric MRI coronal images were acquired at the Unit of Diagnostic Radiology ASST Papa Giovanni XXIII, Bergamo, Italy using a 1.5 T clinical scanner (Optima MR450w GEM, GE Healthcare, Chicago, IL, US) without injection of contrast media, following the protocol of the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP)[18]. Anatomical images were obtained using T2-w single-shot fast spin-echo

Variable	Value
Gender	Female: 9 (64%)
	Male: 5 (36%)
Age [years]	45 [39–52]
eGFR [mL/min/1.73 m^2]	87 [69–100]
Mayo classification	1B: 3 (21.4%)
	1C: 4 (28.6%)
	1D: 4 (28.6%)
	1E: 3 (21.4%)

Table 1. Demographic and clinical information of the ADPKD population considered in this study. Values are expressed as median [IQR] or number (%).

with fat saturation (matrix = 256 × 256, field of view = 380 × 380, thickness/gap = 4/0 mm, TE/TR = 190/1400 ms), 3D spoiled gradient interpolated T1-weighted (T1-w) without fat saturation, and Fast Imaging Employing Steady-State Acquisition (FIESTA). Diffusion-Weighted Imaging (DWI) was acquired with a single-shot spin-echo sequence (matrix = 256 × 256, thickness/gap = 5/1 mm, b-values: 0, 15, 50, 100, 200, 350, 500, 700, and 1000 s/mm^2, three averaged diffusion encoded directions, NEX = 1, ASSET = 2, fat suppression). DWI images were then processed to obtain the IntraVoxel Incoherent Motion (IVIM) model. A researcher with expertise in renal imaging followed all acquisitions, to guarantee high image quality and to ensure consistency of the acquisition parameters with the standard protocol. Moreover, all acquired images underwent visual inspection by renal imaging experts. In cases where relevant artifacts were detected, the acquisition was either repeated when possible or discarded to maintain data integrity. An example of the different MRI acquisitions is reported in Fig. 1a.

Preprocessing

Bias correction and intensity normalization

Artifacts from magnetic field inhomogeneities, called bias fields, require removal since they cause blurs that may distort the information contained in the images. Bias fields were addressed using the N4ITK method across the T1-w, T2-w, and DWI datasets[19] using Slicer 3D 4.10.0. Accordingly, the bias field estimated from $b0$ was removed from all b-values.

To ensure consistent and comparable data, intensity normalization was carried out on the T1-w and T2-w images by aligning intensity distributions through the Nyul histogram-matching approach[20], which has been demonstrated to provide the highest radiomic reproducibility compared to other normalization techniques[16]. Image histograms were thus aligned with a reference distribution by focusing on specific points. For each modality, a reference histogram was computed from the population, with intensity values linearly mapped from the histogram's reference points to the corresponding points in the reference histogram. In this study, the configuration involved 256 histogram bins, with reference points set at the 9 deciles as well as the 99th and 1st percentiles. Along with normalization, a resampling process was also involved, aimed at both matching the sizes of images and removing artifacts by applying a Gaussian filter.

Image registration

Affine registration was conducted between anatomical and DWI images to align all images with the T2-w space, employing Elastix[21]. Similarity measures related to mutual information were sought to enable optimal alignment. To minimize interpolation errors that may affect radiomic computation, the estimated spatial transformation was applied directly to the ROI masks, rather than to the original images. Additionally, an affine registration was applied to all b-values, relative to the previously registered value of b, to address distortions and movements.

IVIM maps generation

IVIM parametric maps were generated from the bi-exponential signal decay over the acquired b-values, to estimate the model's coefficients D, f, and D^*. For this purpose, a Deep Neural Network (DNN) approach
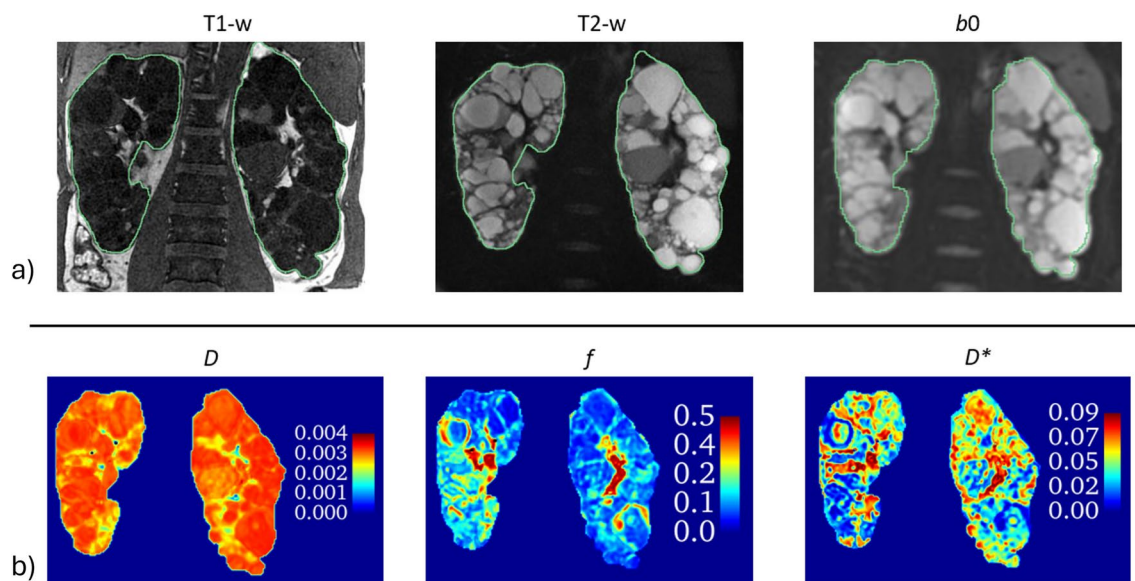

Figure 1. (**a**) Illustrative example of the multi-parametric MRI acquisitions considered in this study, for a patient affected by ADPKD. From left to right: T1-w MRI, T2-w MRI and b0 from DW-MRI. kidney contours are highlighted in green. (**b**) Parametric maps computed within kidneys from multi-b values DW-MRI acquisition following the IntraVoxel Incoherent Motion (IVIM) model. From left to right: true diffusion (D), flowing fraction (f) and pseudo-diffusion (D^*) maps. Values for D and D^* are expressed in mm^2/s.

was chosen, as it has been recently highlighted that this kind of technique is able to increase accuracy and computation time compared to traditional methods, such as Non-Linear Least Squares and Bayesian fitting[22]. In this study, IVIM parametric maps were generated from DWI images using a supervised fully connected DNN approach developed by Mastropietro et al.[22], trained on simulated signals generated using the specific b-values considered in this study. The network was applied to preprocessed b-values sequences, taking as input the nine b-values along with the log(b); the obtained parametric maps were then masked on b0 kidney segmentation. The resulting parametric maps are depicted in Figure 1b. Image preprocessing and registration were performed using Python libraries, while MATLAB scripts (MathWorks, Natick, MA, USA, R2023a) were utilized for IVIM quantification.

Non-cystic tissue identification

Non-cystic tissue was automatically identified using two different approaches:

- *Extended non-cystic tissue*. The extended non-cystic tissue segmentations were derived by subtracting the binary masks of the cysts, obtained by a DNN approach, from the binary mask of the entire kidney. For the identification of the cysts, two freely available segmentation tools based on DNN were considered:

  - The online tool Traceorg (https://www.traceorg.com/), developed by Moghadam et al.[8], leverages a self-configured nnU-Net architecture utilizing both 2D and 3D models for organ segmentation on abdominal T2-w MRI scans, specifically targeting liver and renal cysts in patients with Autosomal Dominant Polycystic Kidney Disease. The tool outputs a mask with multiple labels, each corresponding to the segmentation of cysts within specific organs.
  - The Convolutional Neural Network (CNN) developed by Gregory et al.[12] was employed to automatically segment cysts on the T2-w image and the corresponding reference kidney mask. The network was developed and trained by the authors to recognize cysts. Input images were processed through multiple connected layers of convolutional filters, which learn to recognize patterns associated with cysts. The output is a probability map indicating the likelihood that each voxel belongs to a cyst.

 The two different non-cystic segmentations obtained from the subtraction of cysts from the whole renal masks were named as NCext_1 and NCext_2, respectively. As this process did not involve any further modifications, the segmented region included both the tissue of interest and other tissues, such as vessels, fat, and hemorrhagic cysts.

- *Refined non-cystic tissue*. mp-MRI data were employed to refine the non-cystic tissue segmentations obtained from NCext_1 and NCext_2, as illustrated in Figure 2. This process led to the creation of Non-Cystic Refined segmentation 1 (NCref_1), which represents the refined non-cystic tissue obtained by subtracting the cyst segmentations generated by Traceorg[8] from the TKV. Non-Cystic Refined segmentation 2 (NCref_2) was derived by refining the subtraction of the total kidney mask and the cyst segmentations described by Gregory et al.[12]. Similar to the study conducted by Caroli et al.[5], vessels were excluded using the perfusion fraction (f) map with a threshold of 0.4 (Fig. 2b), and fat was removed by excluding all b0 values below 400 (Fig. 2c). Values were chosen experimentally, and small variations in these thresholds result in negligible modifications in the final mask. Finally, hemorrhagic cysts that were not detected by the original cyst segmentations were identified through the combination of anatomical MRI. In fact, they exhibit distinct intensity patterns on T1-w and T2-w images, showing hyperintense values on the first and hypointense values on the latter. They were thus identified based on pixel values exceeding the 98th percentile of the T1-w histogram and falling below the 2nd percentile of the T2-w histogram, (Fig. 2d), where these values were chosen based on empirical evaluations. Morphological operations were applied to enhance contour precision. The resulting non-cystic ROI was thus obtained by removing cysts, vessels, fat, and hemorrhagic cysts from the manual kidney segmentation.

Additionally, the manual segmentation of the non-cystic tissue was considered as Ground Truth (GT). Specifically, two renal MRI experts in consensus segmented the cysts on T2-w scans using a previously published semi-automatic method based on the ITK software library (Insight Segmentation and Registration Toolkit, http://www.itk.org/)[23]. Non-cystic tissue was finally obtained by subtracting cyst segmentation from the TKV and refining the delineation by excluding vessels and fat.

 The agreement between GT-derived ROIs and those obtained through automatic approaches (NCext_1, NCext_2, NCref_1, and NCref_2), as well as the accuracy in identifying renal cysts, was quantified using the Dice Similarity Coefficient (DSC) index.

Radiomic feature extraction

Feature extraction was conducted utilizing the Pyradiomics open-source software[24] (version 3.0.1), incorporating IVIM parametric maps (D, f, and D^*) as well as T1-w and T2-w images. Given the high anisotropic image resolution, features were computed slice-by-slice to account for the substantial disparity between in-plane pixel size and slice thickness. Texture matrices were generated through image discretization, employing fixed bin widths of 0.00005, 0.003, and 0.005 for D, D^*, and f maps, respectively, to accommodate their quantitative nature. For T1-w and T2-w images, a fixed bin number of 64 was chosen in accordance with the Image Biomarker Standardization Initiative (IBSI) guidelines[25]. Radiomic features encompassing 18 First Order, 22 Gray Level Co-Occurrence Matrix (GLCM), 16 Gray Level Run Length Matrix (GLRM), 16 Gray Level Size Zone Matrix (GLSZM), 14 Gray Level Dependence Matrix (GLDM), and 5 Neighboring Gray Tone Difference
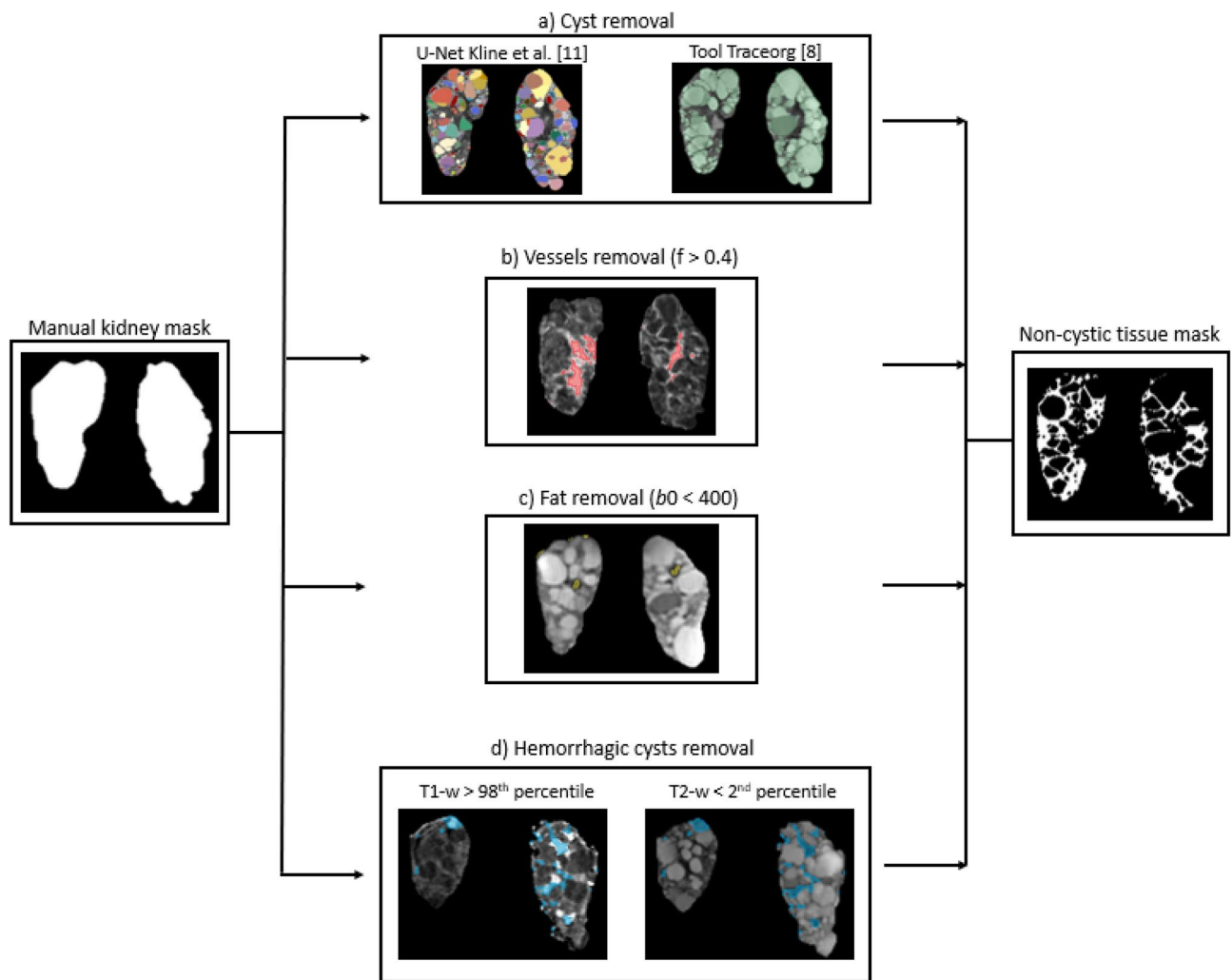
Figure 2. Schematic diagram showing the steps employed for the NCref_1 and NCref_2 segmentations of non-cystic tissue (non-cystic binary mask shown on the right) starting from kidney manual segmentation on T2-weighted MRI (binary mask on the left). The procedure includes (**a**) segmentation of cysts using the CNN developed by Kline et al.[11] and Traceorg[8] and removal of them, (**b**) identification of vessels on the f parametric map (f greater than 0.4) and removal of them, (**c**) identification of fat on the $b0$ image ($b0$ lower than 400) and removal of it, and (**d**) identification of hemorrhagic cysts on T1-w and T2-w images (pixels that simultaneously have values greater than the T1-w 98th percentile and less than the T2-w 2nd percentile).

Matrix (NGTDM) were extracted, resulting in a total of 91 features for each image, considering the two kidneys separately, and for each considered segmentation method.

Reliability assessment

The reliability of radiomic features was assessed in relation to the different non-cystic segmentation approaches using Python libraries. For the evaluation, the Intraclass Correlation Coefficient (ICC) was utilized, employing the two-way mixed effect model with a single rater type formulation[26]. The ICC was computed between features extracted from non-cystic ROIs in each kidney separately for all images, considering the pairs GT and NCext_1, GT and NCref_1, GT and NCext_2, and GT and NCref_2. Thus, a dataset comprising 28 data points was available for each comparison.

Specifically, a radiomic feature was deemed reproducible if it remained stable (ICC > 0.8) when extracted from ROIs with similar information content, such as GT and NCref_1, and GT and NCref_2. Conversely, a radiomic feature was considered informative for characterizing non-cystic tissue if it yielded different values (ICC < 0.5) when extracted from ROIs with distinct information content, such as GT and NCext_1, and GT and NCext_2, as they differ across themselves for the presence of vessels, hemorrhagic cysts, and fat. A similar methodology for identifying reproducible and informative radiomic features was proposed by Bologna et al.[27]. Subsequently, the percentage of reliable features for each image type was determined, considering non-cystic tissue identified from cysts segmented using the tool provided by Gregory et al.[12] (segmentation_2), and using the Traceorg segmentation tool[8] (segmentation_1), separately.
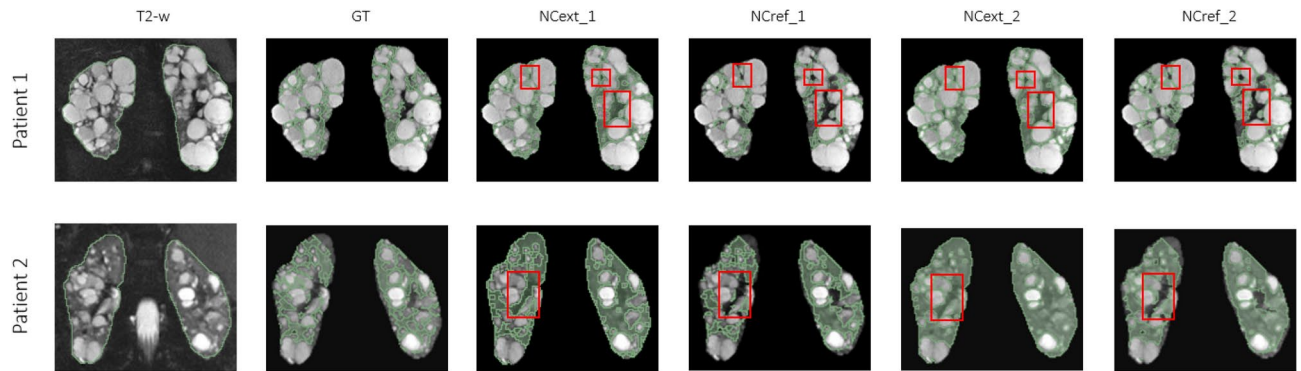
Figure 3. Results of non-cystic tissue segmentation obtained on T2-w scans from two representative ADPKD patients, using the different segmentation methods. Hemorrhagic cysts and vessels are highlighted in the red boxes.
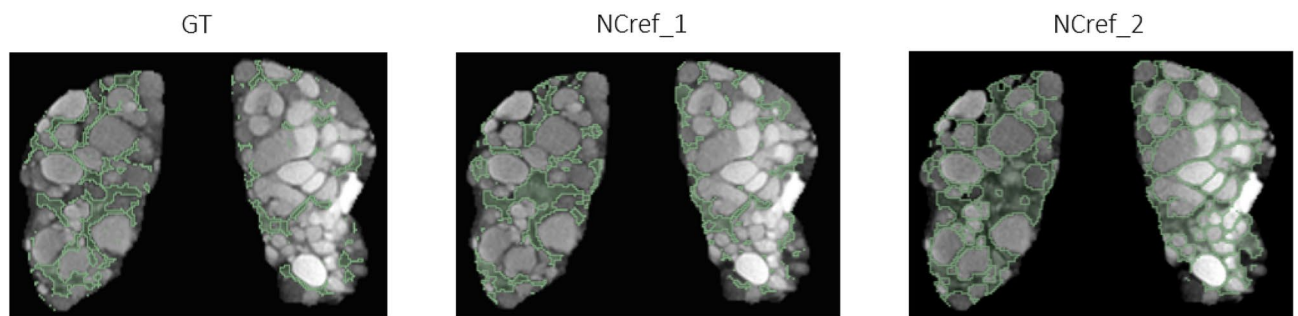

Figure 4. Comparison of non-cystic tissue segmentation obtained by manual Ground Truth (GT), NCref_1, and NCref_2 methods.

Correlation analysis

A preliminary investigation of the association between reliable radiomic features and clinical information was provided, to outline a potential clinical application of the proposed framework. For this analysis, only reproducible and informative radiomic features from segmentation_1 and segmentation_2 were considered, and the correlation with age and eGFR was explored. Additionally, features with null variance or high correlation (>0.8) were removed, retaining only one representative feature per cluster. Finally, Spearman correlation was used to assess associations between selected radiomic features and age/eGFR. Only statistically significant correlations ($p < 0.05$) were reported.

Results

Non-cystic tissue identification

Representative segmentation results for T2-w images are depicted in Figure 3, showing that the NCref_1 and NCref_2 segmentation exhibits greater similarity to the GT compared to NCext_1 and NCext_2. As can be seen from the image, the NCext_1 and NCext_2 segmentations include vessels, fat, and hemorrhagic cysts that are, in contrast, removed in the NCref_1 and NCref_2 segmentations. NCref_1 and NCref_2 accurately segment non-cystic tissue, with the primary difference being that NCref_1 excludes certain portions of tissue between cysts, unlike NCref_2, as illustrated in Fig. 4. Additionally, during the refinement process, some sections along the kidney border are lost.

Quantitatively, renal cysts were accurately segmented by both DL-based approaches, obtaining a Dice score of 0.85 ± 0.05 using Traceorg tool and a Dice score of 0.82 ± 0.09 using the model provided by Gregory et al.[12]. Regarding non-cystic ROI segmentation, Dice scores confirmed a higher concordance between NCref_1 and GT (Dice = 0.68 ± 0.12), and NCref_2 and GT (Dice = 0.68 ± 0.11) than between NCext_1 and GT (Dice = 0.49 ± 0.16), and NCext_2 and GT (Dice = 0.50 ± 0.16). In addition, the two refined segmentations NCref_1 and NCref_2 appear to be aligned with each other (Dice = 0.70 ± 0.20) as well as with GT.

Radiomics reliability

The results of radiomic reliability are presented in Table 2, showing median ICC values across all the radiomic features in each class. The reported metrics indicate the ICC between NCext_1, NCref_1, NCext_2, and NCref_2 segmentations relative to the GT. The refined segmentations demonstrated higher reproducibility compared to the extended segmentations when considering IVIM parametric maps, with an average ICC consistently

Feature	Segmentation	D	f	D^*	T1	T2
Average firstorder	NCref_1	**0.80** [0.59–0.84]	**0.91** [0.75–0.97]	**0.95** [0.94–0.97]	0.69 [0.53–0.80]	0.63 [0.41–0.88]
	NCext_1	0.08 [−0.05–0.47]	0.17 [−0.01–0.28]	0.47 [0.27–0.60]	**0.85** [0.54–0.90]	0.58 [0.45–0.82]
	NCref_2	**0.92** [0.89–0.94]	**0.96** [0.93–0.97]	**0.98** [0.98–0.99]	0.66 [0.52–0.79]	0.66 [0.45–0.84]
	NCext_2	0.18 [0.03–0.47]	0.26 [0.06–0.6]	0.56 [0.25–0.62]	**0.83** [0.54–0.89]	0.69 [0.52–0.77]
Average glcm	NCref_1	0.71 [0.62–0.81]	**0.88** [0.81–0.92]	**0.94** [0.92–0.95]	0.44 [0.30–0.52]	0.53 [0.34–0.70]
	NCext_1	0.23 [−0.05–0.45]	0.09 [−0.02–0.28]	0.48 [0.19–0.65]	0.79 [0.75–0.84]	0.54 [0.45–0.69]
	NCref_2	**0.91** [0.85–0.92]	**0.91** [0.91–0.94]	**0.98** [0.97–0.98]	0.51 [0.35–0.59]	0.60 [0.44–0.68]
	NCref_2	0.28 [0.04–0.54]	0.10 [0–0.57]	0.55 [0.16–0.69]	0.77 [0.73–0.83]	0.72 [0.50–0.84]
Average glrlm	NCref_1	0.78 [0.7–0.85]	**0.89** [0.71–0.97]	**0.90** [0.79–0.96]	0.43 [0.34–0.58]	0.37 [0.14–0.61]
	NCext_1	0.55 [0–0.74]	0.66 [0.09–0.84]	0.74 [0.03–0.91]	0.76 [0.71–0.84]	0.53 [0.3–0.72]
	NCref_2	**0.80** [0.73–0.92]	**0.88** [0.68–0.97]	**0.92** [0.77–0.99]	0.50 [0.40–0.59]	0.52 [0.09–0.61]
	NCext_2	0.62 [0.12–0.73]	0.70 [0.13–0.88]	0.72 [0.24–0.95]	0.70 [0.46–0.84]	0.52 [0.37–0.78]
Average glszm	NCref_1	0.67 [0.58–0.86]	**0.86** [0.82–0.93]	**0.91** [0.85–0.96]	0.54 [0.38–0.58]	0.61 [0.15–0.62]
	NCext_1	0.03 [0–0.57]	0 [−0.01–0.5]	0.20 [0.01–0.79]	0.79 [0.73–0.83]	0.67 [0.33–0.78]
	NCref_2	0.74 [0.66–0.85]	**0.86** [0.79–0.93]	**0.90** [0.84–0.94]	0.57 [0.42–0.65]	0.47 [0.17–0.58]
	NCext_2	0.10 [−0.01–0.40]	0.04 [0–0.18]	0.24 [0.01–0.57]	0.75 [0.71–0.80]	0.69 [0.39–0.84]
Average gldm	NCref_1	0.66 [0.52–0.82]	**0.89** [0.84–0.95]	**0.91** [0.85–0.96]	0.54 [0.38–0.57]	0.61 [0.44–0.6]
	NCext_1	0.23 [0.06–0.38]	0 [−0.02–0.59]	0.12 [0–0.71]	0.79 [0.78–0.84]	0.50 [0.35–0.81]
	NCref_2	0.74 [0.68–0.88]	**0.86** [0.69–0.96]	**0.94** [0.76–0.97]	0.46 [0.37–0.54]	0.44 [0.28–0.56]
	NCext_2	0.23 [0.08–0.42]	0 [0–0.51]	0.13 [0.01–0.69]	0.77 [0.72–0.78]	0.69 [0.38–0.88]
Average ngtdm	NCref_1	**0.82** [0.81–0.87]	0.70 [0.7–0.88]	**0.81** [0.67–0.91]	0.44 [0.36–0.45]	0.50 [0.41–0.5]
	NCext_1	0 [−0.03–0.14]	0.28 [0–0.38]	0.33 [−0.01–0.49]	0.66 [0.5–0.76]	0.46 [0.1–0.56]
	NCref_2	**0.89** [0.81–0.89]	**0.85** [0.78–0.86]	**0.84** [0.81–0.94]	0.56 [0.52–0.56]	0.63 [0.61–0.69]
	NCext_2	0 [−0.03–0.28]	0.38 [0–0.43]	0.36 [−0.01–0.62]	0.70 [0.53–0.72]	0.70 [0.38–0.73]

Table 2. Intraclass correlation coefficient (ICC, median[interquartile range]) of segmentation methods as compared to the ground truth segmentation, across all the radiomic features for D, f, D^*, T1-w, and T2-w images. Median ICC values greater than 0.8 are shown in bold.

exceeding 0.8. In particular, the f and D^* maps showed the highest reproducibility, above 0.85. Within the same segmentations, higher ICC values were observed in NCref_2, while the lowest ICC values occurred in NCext_1. Looking at the specific radiomic classes, FO and GLCM features revealed the highest ICC values for all three IVIM maps, often achieving an ICC above 0.9. As expected, the NCext_1 and NCext_2 segmentation showed lower reproducibility in the IVIM maps, not exceeding 0.28 for the D and f maps and 0.56 for the D^* map, except for the GLRLM feature where the ICC values were higher.

Conversely, features extracted from refined segmentations were not reproducible when derived from anatomical T1-w and T2-w images, yielding an average ICC of approximately 0.5, while extended segmentations exhibited average ICC values of 0.76 and 0.61 for T1-w and T2-w images, respectively. In particular, T1-w turns out to be slightly more reproducible than T2-w, with higher ICC values for both extended and refined segmentations.

The complete results for each feature, including ICC values, 95% Confidence Intervals and p-values, can be found in Table S1.

Figure 5 depicts the percentages of reliable features for each image, categorized by radiomic feature classes, for segmentation_1 (A) and segmentation_2 (B). In both barplots, it can be observed that the percentage of reliable features is higher for the IVIM maps, particularly for the f map in both grouped segmentations and for the D map in segmentation_2. Notably, for segmentation_1, the most reliable features were FO, GLCM, and GLSZM for the f parametric map, while NGTDM was the most reliable for the D parametric map. For segmentation_2, the most reliable features included FO and GLCM for both the D and f parametric maps, along with NGTDM for the D map and GLSZM for the f map. Importantly, the reliability percentage of these features consistently exceeded 60%, reaching as high as 80% for the NGTDM feature of the D map in both segmentations, and 87% for the GLSZM features of the f map for segmentation_1. Interestingly, none of the radiomic features computed from anatomical T1-w and T2-w images exhibited good reliability, attributed to the good reproducibility (ICC > 0.5) between GT and the refined segmentations, which makes these imaging acquisitions less informative for non-cystic tissue characterization.

The list of reliable and independent features, found for both segmentation_1 and segmentation_2 approaches, is separately reported in Table 3 for each image type. Notably, no features were available from T1-w and T2-w images.

Correlation analysis

The reliable and independent features features used to assess Spearman correlation with age and eGRF are listed in Table S1. The results showed a moderate correlation between some first-order features from the D map and
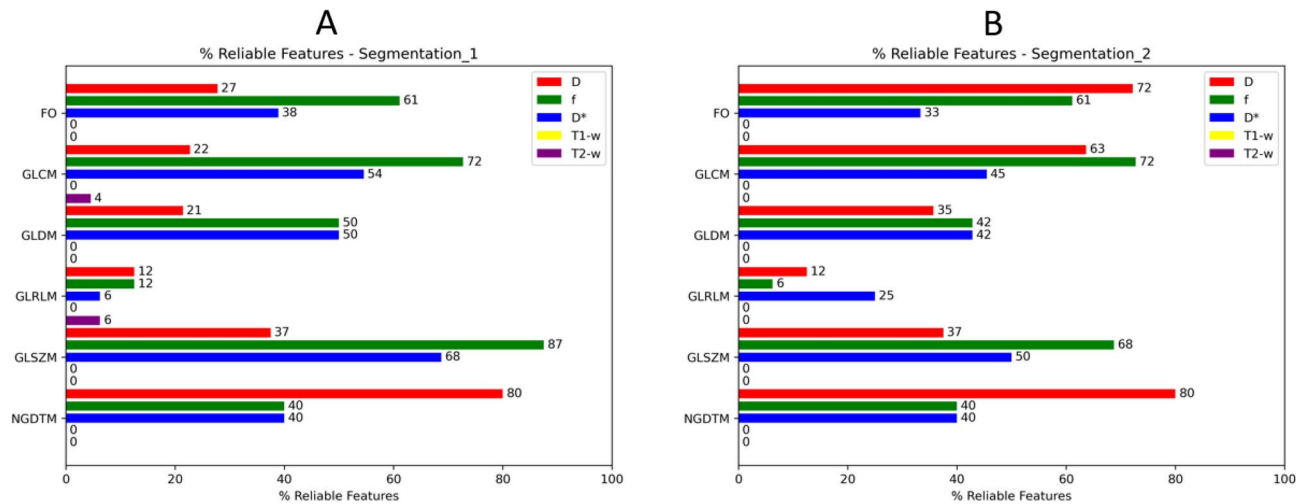
Figure 5. Percentage of reliable features in each radiomic group for D, f, D^*, T1-w, and T2-w images, divided by radiomic classes for segmentation_1 (**A**) and segmentation_2 (**B**).

D map	f map	D^* map
firstorder_90Percentile	glcm_MaximumProbability	firstorder_Entropy
firstorder_MeanAbsoluteDeviation	glcm_ClusterShade	firstorder_Skewness
ngtdm_Complexity	glcm_ClusterProminence	glcm_DifferenceEntropy
	glrlm_RunLengthNonUniformity Normalized	glcm_Imc2
	ngtdm_Complexity	glcm_MaximumProbability
		gldm_DependenceEntropy
		glrlm_LowGrayLevel RunEmphasis

Table 3. List of reliable and independent features that are in common between segmentation_1 and segmentation_2.

Image	Segmentation method	Radiomic feature	Target variable	Spearman correlation	p-value
D map	NCref_1	firstorder_90Percentile	Age	0.657	0.011
		firstorder_Mean	eGFR	− 0.647	0.012
		firstorder_90Percentile	eGFR	− 0.590	0.026
	NCref_2	firstorder_90Percentile	Age	0.596	0.025
		firstorder_Entropy	eGFR	0.561	0.037
D^* map	NCref_2	glszm_ZoneEntropy	Age	− 0.587	0.027

Table 4. Results of the Spearman correlation analysis between radiomic features and age/eGFR. Only significant correlations ($p<0.05$) are reported.

age and eGFR (see Table 4), whilst no correlations were found for the f map and only one (glszm_ZoneEntropy vs. age) for the D^* map. The scatterplots of the significant correlations, along with the estimated linear regression lines and 95% confidence intervals, are reported in Figure S1.

Discussion

In this study, we conducted the first assessment of radiomic reliability of mp-MRI scans in patients with ADPKD, focusing on the characterization of non-cystic tissue.

Segmenting and characterizing non-cystic tissue has proven to be fundamental for the study of ADPKD pathology[28–30]. Accurately delineating each cyst and tissue outside the renal parenchyma presents a formidable challenge for both manual and automatic methods, often necessitating multi-parametric information to identify various structures of interest, including fluid-filled and hemorrhagic cysts, vessels, and fat. To the best of our knowledge, this approach has been used only recently and has been minimally explored apart from Lai et al.[29]. Unlike renal cysts, the non-cystic region lacks a standardized definition and is inherently more difficult to segment due to several factors. First, it is a highly heterogeneous tissue that includes peritubular

interstitial volume, fibrosis, microcysts, dilated tubules, blood vessels, and fat, leading to complex and irregular shapes. Additionally, microcysts that are not visible on MRI may be included in the non-cystic region, further complicating segmentation. To refine non-cystic segmentation, our approach identifies hemorrhagic cysts, which are often misclassified in automated segmentations, and removes fat and vascular structures to improve specificity. Although this process improves segmentation accuracy, some non-cystic tissue may still be inadvertently excluded, impacting Dice score results. Nonetheless, the refined segmentation achieved a Dice score of 0.68, which we consider to be a reasonable agreement given the complexity of the non-cystic region.

Radiomics derived from IVIM parametric maps presents a compelling avenue for analysis, owing to its enhanced capacity to evaluate micro-structural tissue organization and composition in comparison to conventional T1-w and T2-w acquisitions, which turn out, across various studies, to have a small amount of reliable features[31,32]. Specifically, parameters such as D and f can elucidate diffusion and flowing-related properties critical for assessing the pathological condition of renal parenchymal tissue in ADPKD patients[5].

The extraction of radiomic features from IVIM maps has been recognized as being prone to errors and variability[33], owing to the diverse quantification methods available. Also, the application of DL techniques in computing IVIM maps enables more precise evaluation of radiomic parameters, thereby mitigating quantification errors and enhancing reproducibility and accuracy[33]. In our study, we relied on a cutting-edge quantification method based on DL to ensure the accurate estimation of IVIM maps[22]. In their original work, the authors provided errors estimation for the proposed supervised DL model, with a Median Absolute Error lower than 0.05 for D, lower than 0.15 for f and D^*, and a Median Bias centered around 0 for all three coefficients, confirming that the IVIM parametric maps estimated using this method are at the state-of-the-art level in terms of accuracy. Therefore, the two mentioned studies[22,33] support our decision to adopt this method for IVIM quantification, and confirm that radiomics can be reliably performed on the estimated parametric maps.

Recent investigations have assessed radiomic reproducibility in ADPKD patients concerning image pre-processing procedures, such as intensity normalization and image discretization, particularly focusing on T2-w and T1-w acquisitions[34,35]. In this study, following the recommendations from a previous review[16], several preprocessing steps were implemented to enhance the reproducibility of radiomic features in anatomical images. These steps included a standardized image acquisition protocol, followed by expert visual inspection, image registration, intensity normalization and discretization, to guarantee high image quality and preserve textural information. In contrast, the evaluation of radiomic reliability from IVIM maps in ADPKD is still an open research field.

Our findings indicated that radiomic features demonstrate reliability for D and f maps when non-cystic tissue is identified using segmentation_2, and for the f map using segmentation_1. Additionally, they exhibit an intermediate behavior for D^* maps and are entirely unreliable for T1-w and T2-w images in all segmentations. IVIM maps, in fact, turn out to have the highest ICC values and a very good percentage of reliable features for segmentation_1, and segementation_2, unlike anatomical images, which, in contrast, have no reliable features.

The limited efficacy of anatomical images in providing informative features can be attributed to the lower contrast among different tissues, which diminishes the sensitivity of radiomic feature computation to large segmentation errors and their relevance for characterizing non-cystic tissue specifically. This is consistent with the findings of Kremer *et al.*, who evaluated radiomic model performance for ADPKD genotype classification using T2-w images and observed higher area under the curve (AUC) values for models constructed on the entire kidney parenchyma compared to non-cystic tissue[36].

In contrast, texture information derived from diffusion and flowing fraction maps appears capable of characterizing the non-cystic parenchymal region effectively, being both informative and resilient to minor segmentation discrepancies, particularly in features from FO and GLCM classes of D and f maps for segmentation_2 and FO, GLCM, and GLSZM classes of f map for segmentation_1. The differences between reliability of features extracted from D maps using segmentation_1 and segmentation_2 approaches may be explained by the ability to identify small portions of non-cystic tissue at the borders of the cysts by NCref_2, whereas in NCref_1 packed cysts are grouped together. Being a high-contrast region in D maps, it has a high impact on feature reproducibility.

The reproducibility results suggest that there is high homogeneity and robustness for parameters describing pixel intensity distribution and relationships between neighboring pixels within the considered non-cystic segmentation[37]. Previous reports have suggested that IVIM parameters have the potential to detect fluid-filled microcysts and peritubular interstitial fibrosis in this region[5]. It is worth noting that caution should be exercised in interpreting information from D^* maps, as these maps are susceptible to larger quantification errors and noise, as highlighted in previous research[33].

We also provided a preliminary analysis of potential correlations between reliable radiomic features from IVIM maps and some clinical parameters, such as age and eGFR. Despite the small sample size here available, some noteworthy trends emerge. All significant correlations but one (D^* map, glszm_ZoneEntropy) were derived from the D map, emphasizing its importance in capturing microstructural properties. The most relevant features belonged to the first-order category, probably because higher-order texture features require larger datasets to achieve statistical significance. These results are in line with biological expectations. In fact, higher diffusion intensity values (firstorder_90Percentile in D map) were positively correlated with age, suggesting a higher presence of microcysts in elderly patients. Similarly, higher diffusion intensity values (firstorder_Mean and firstorder_90Percentile in D map) correlated with lower eGFR, reflecting more advanced ADPKD progression. Higher image entropy (firstorder_Entropy in D map) was associated with higher eGFR, indicating that greater microstructural complexity may be linked to better kidney function. While remaining an exploratory analysis due to the small sample size, it provides useful preliminary insights and highlights the potential for further investigation in larger cohorts.

For a better interpretation of the radiomic analysis and to increase the explainability of the resulting reliable features, potential future studies would be directed towards integration with histological samples. Validating radiomic analyses through histology could theoretically be pursued via pathomics, which applies texture-based feature extraction to pathological images. However, direct comparison between radiomic features derived from macroscopic in-vivo imaging and subcellular-scale pathomic data presents significant challenges and may not provide direct insights into the biological basis of radiomics[38].

A major limitation of our study lies in the relatively small sample size available for the radiomic analysis. In fact, the multi-parametric MRI protocol adopted in this study, which includes IVIM acquisitions, is not yet part of the standard imaging workflow for ADPKD patients. While IVIM has significant potential for non-invasive assessment of cystic and non-cystic microstructure, it is not routinely acquired in large ADPKD cohorts, making it difficult to assemble larger datasets. Future investigations are needed to validate our findings in a larger cohort, preferably from multiple centers and different patient populations, also including different clinical outcomes, such as Mayo classification, genetic information, ADPKD stage and follow-up. Additionally, international validation cohorts would provide stronger evidence for wider applicability of our findings.

Another limitation concerns segmentation accuracy, which needs improvements. In fact, examined segmentations manifest certain irregularities which can distort the final analysis, and the inclusion of additional pre-processing techniques, as well as dedicated DL-based models, would lead to greater precision and more accurate results.

In conclusion, in this study we assessed the reliability of a radiomic analysis aimed at characterizing non-cystic tissue in patients affected by ADPKD on mp-MRI, focusing on reproducibility across different segmentation methods. Our study highlights the robustness of radiomic features derived from D and f IVIM maps to segmentation variability, and their reliability for characterizing non-cystic kidney tissue in patients with ADPKD, especially for FO, GLCM, and GLSZM texture information. Conversely, features derived from conventional anatomical acquisitions appeared to lack the necessary informativeness for this type of analysis.

Data availability

The clinical data used in this study cannot be publicly shared due to restrictions. Researchers interested in accessing similar data may contact Dr. Anna Caroli directly for further information and collaboration.

Code availability

The Matlab code developed to obtain the refined non-cystic ROI for NCref segmentations, and the python codes used to assess radiomic reliability, as well as tables containing all the extracted radiomic features, are publicly available at https://zenodo.org/records/15063396.

References

1. Caroli, A. et al. Intermediate volume on computed tomography imaging defines a fibrotic compartment that predicts glomerular filtration rate decline in autosomal dominant polycystic kidney disease patients. *Am. J. Pathol.* **179**, 619–627. https://doi.org/10.1016/j.ajpath.2011.04.036 (2011).
2. Xie, Y. et al. The predictive value of renal parenchymal information for renal function impairment in patients with ADPKD: A multicenter prospective study. *Abdom. Radiol.* **47**, 2845–2857. https://doi.org/10.1007/s00261-022-03554-w (2022).
3. Caroli, A. & Kline, T. L. Abdominal imaging in ADPKD: Beyond total kidney volume. *J. Clin. Med.* **12**, 5133. https://doi.org/10.3390/jcm12155133 (2023).
4. Magistroni, R., Corsi, C., Marti, T. & Torra, R. A review of the imaging techniques for measuring kidney and cyst volume in establishing autosomal dominant polycystic kidney disease progression. *Am. J. Nephrol.* **48**, 67–78. https://doi.org/10.1159/000491022 (2018).
5. Caroli, A. et al. Diffusion magnetic resonance imaging for kidney cyst volume quantification and non-cystic tissue characterisation in ADPKD. *Eur. Radiol.* https://doi.org/10.1007/s00330-023-09601-4 (2023).
6. Bihan, D. et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505. https://doi.org/10.1148/radiology.168.2.3393671 (1988).
7. Kline, T. L. et al. Image texture features predict renal function decline in patients with autosomal dominant polycystic kidney disease. *Kidney Int.* **92**, 1206–1216. https://doi.org/10.1016/j.kint.2017.03.026 (2017).
8. Moghadam, M. C. et al. Deep learning-based liver cyst segmentation in MRI for autosomal dominant polycystic kidney disease. *Radiol. Adv.* **1**, umae014. https://doi.org/10.1093/radadv/umae014 (2024).
9. Zhu, C. et al. A primer for utilizing deep learning and abdominal MRI imaging features to monitor autosomal dominant polycystic kidney disease progression. *Biomedicines* **12**, 1133 (2024).
10. Rombolotti, M., Sangalli, F., Cerullo, D., Remuzzi, A. & Lanzarone, E. Automatic cyst and kidney segmentation in autosomal dominant polycystic kidney disease: Comparison of u-net based methods. *Comput. Biol. Med.* **146**, 105431. https://doi.org/10.1016/j.compbiomed.2022.105431 (2022).
11. Kline, T. L. et al. Automatic semantic segmentation of kidney cysts in MR images of patients affected by autosomal-dominant polycystic kidney disease. *Abdom. Radiol.* **46**, 1053–1061. https://doi.org/10.1007/s00261-020-02748-4 (2021).
12. Gregory, A. et al. Semantic instance segmentation of kidney cysts in MR images: A fully automated 3D approach developed through active learning. *J. Digit. Imaging* **34**, 1–15. https://doi.org/10.1007/s10278-021-00452-3 (2021).
13. Lambin, P. et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**, 749–762. https://doi.org/10.1038/nrclinonc.2017.141 (2017).
14. Li, X. et al. A MRI-based radiomics nomogram for evaluation of renal function in ADPKD. *Abdom. Radiol.* **47**, 1385–1395. https://doi.org/10.1007/s00261-022-03433-4 (2022).
15. Calvaruso, L. et al. #6060 A novel CT-based radiomics approach for kidney function evaluation in ADPKD. *Nephrol. Dial. Transplant.* https://doi.org/10.1093/ndt/gfad063c_6060 (2023).
16. Scalco, E., Rizzo, G. & Mastropietro, A. The stability of oncologic MRI radiomic features and the potential role of deep learning: A review. *Phys. Med. Biol.* **67**, 09TR03. https://doi.org/10.1088/1361-6560/ac60b9 (2022).
17. Pfaehler, E. et al. A systematic review and quality of reporting checklist for repeatability and reproducibility of radiomic features. *Phys. Imaging Radiat. Oncol.* **20**, 69–75. https://doi.org/10.1016/j.phro.2021.10.007 (2021).

18. Chapman, A. B. et al. Renal structure in early autosomal-dominant polycystic kidney disease (ADPKD): The consortium for radiologic imaging studies of polycystic kidney disease (CRISP) cohort. *Kidney Int.* **64**, 1035–1045. https://doi.org/10.1046/j.1523-1755.2003.00185.x (2003).
19. Tustison, N. J. et al. N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* **29**, 1310–1320. https://doi.org/10.1109/TMI.2010.2046908 (2010).
20. Nyúl, L. G., Udupa, J. K. & Zhang, X. New variants of a method of MRI scale standardization. *IEEE Trans. Med. Imaging* **19**, 143–150. https://doi.org/10.1109/42.836373 (2000).
21. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. Elastix: A toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* **29**, 196–205. https://doi.org/10.1109/TMI.2009.2035616 (2009).
22. Mastropietro, A., Procissi, D., Scalco, E., Rizzo, G. & Bertolino, N. A supervised deep neural network approach with standardized targets for enhanced accuracy of IVIM parameter estimation from multi-SNR images. *NMR Biomed.* **35**, e4774. https://doi.org/10.1002/nbm.4774 (2022).
23. Caroli, A. et al. Effect of longacting somatostatin analogue on kidney and cyst growth in autosomal dominant polycystic kidney disease (ALADIN): A randomised, placebo-controlled, multicentre trial. *The Lancet* **382**, 1485–1495 (2013).
24. Van Griethuysen, J. J. et al. Computational radiomics system to decode the radiographic phenotype. *Can. Res.* **77**, e104–e107. https://doi.org/10.1158/0008-5472.CAN-17-0339 (2017).
25. Zwanenburg, A. et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **295**, 328–338. https://doi.org/10.1148/radiol.2020191145 (2020).
26. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012 (2016).
27. Bologna, M. et al. Assessment of stability and discrimination capacity of radiomic features on apparent diffusion coefficient images. *J. Digit. Imaging* **31**, 879–894. https://doi.org/10.1007/s10278-018-0092-9 (2018).
28. Antiga, L. et al. Computed tomography evaluation of autosomal dominant polycystic kidney disease progression: A progress report. *Clin. J. Am. Soc. Nephrol. CJASN* **1**, 754–60. https://doi.org/10.2215/CJN.02251205 (2006).
29. Lai, S. et al. Magnetic resonance imaging 3T and total fibrotic volume in autosomal dominant polycystic kidney disease: Markers of progression in ADPKD. *Intern. Med. J.* **48**, 1505–1513. https://doi.org/10.1111/imj.14039 (2018).
30. Li, L.-P. et al. Radiomics-based image phenotyping of kidney apparent diffusion coefficient maps: Preliminary feasibility & efficacy. *J. Clin. Med.* **11**, 1972. https://doi.org/10.3390/jcm11071972 (2022).
31. Saha, A., Yu, X., Sahoo, D. & Mazurowski, M. Effect of MRI scanner parameters on breast cancer radiomics. *Expert Syst. Appl.* **87**, 384–391. https://doi.org/10.1016/j.eswa.2017.06.029 (2017).
32. Mayerhoefer, M., Szomolanyi, P., Jirak, D., Materka, A. & Trattnig, S. Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: An application-oriented study. *Med. Phys.* **36**, 1236–43. https://doi.org/10.1118/1.3081408 (2009).
33. Scalco, E., Rizzo, G. & Mastropietro, A. The quantification of intravoxel incoherent motion-MRI maps cannot preserve texture information: An evaluation based on simulated and in-vivo images. *Comput. Biol. Med.* **154**, 106495. https://doi.org/10.1016/j.compbiomed.2022.106495 (2023).
34. Kremer, L. E., Fosso, B., Groothuis, L., Chapman, A. & Armato III, S. G. Radiomics-based classification of autosomal dominant polycystic kidney disease (ADPKD) Mayo imaging classification (MIC) and the effect of gray-level discretization. In *Medical Imaging 2023: Computer-Aided Diagnosis*, vol. 12465, 695–703 (SPIE, 2023). https://doi.org/10.1117/12.2654476.
35. Kremer, L. E., Perri, N., Sorber, E., Chapman, A. & Armato III, S. G. Normalization of MRI signal intensity in polycystic kidney disease and the effect on radiomic features. In *Medical Imaging 2022: Computer-Aided Diagnosis*, vol. 12033, 774–781 (SPIE, 2022). https://doi.org/10.1117/12.2613164.
36. Kremer, L. E., Chapman, A. B. & Armato, S. G. III. Magnetic resonance imaging preprocessing and radiomic features for classification of autosomal dominant polycystic kidney disease genotype. *J. Med. Imaging* **10**, 064503–064503. https://doi.org/10.1117/1.JMI.10.6.064503 (2023).
37. Wu, Y. et al. Repeatability, reproducibility and sources of variability in the assessment of backscatter coefficient and texture parameters from high-frequency ultrasound acquisitions in human median nerve. *Ultrasound Med. Biol.* **49**, 122–135. https://doi.org/10.1016/j.ultrasmedbio.2022.08.007 (2023).
38. Tomaszewski, M. R. & Gillies, R. J. The biological meaning of radiomic features. *Radiology* **298**, 505–516 (2021).

Acknowledgements

Author contributions

E.L., A.C., and E.S. designed the study; G.V. and A.C. provided the dataset; F.L., G.V., A.M., A.C., and E.S. supplied codes; F.L., E.L., and E.S. performed the analyses and examined the results. F.L., E.S., and E.L. wrote the first draft of the manuscript. All authors discussed and reviewed the manuscript.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-99982-x.

Correspondence and requests for materials should be addressed to E.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.