


Dynamic parametric MRI and deep learning: Unveiling renal pathophysiology through accurate kidney size quantification

Tobias Klein^{1,2}  | Thomas Gladytz¹ | Jason M. Millward¹ | Kathleen Cantow³ | Luis Hummel³ | Erdmann Seeliger³ | Sonia Waiczies¹ | Christoph Lippert^{2,4} | Thoralf Niendorf^{1,5}

¹Berlin Ultrahigh Field Facility (B.U.F.F.), Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

²Digital Health - Machine Learning Research Group, Hasso Plattner Institute for Digital Engineering, University of Potsdam, Potsdam, Germany

³Institute of Translational Physiology, Charité – Universitätsmedizin, Berlin, Germany

⁴Hasso Plattner Institute for Digital Health, Icahn School of Medicine at Mount Sinai, New York City, New York, USA

⁵Experimental and Clinical Research Center, a joint cooperation between the Charité Medical Faculty and the Max Delbrück Center for Molecular Medicine, Berlin, Germany

Correspondence

Thoralf Niendorf, Berlin Ultrahigh Field Facility (B.U.F.F.), Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Robert Rössle Strasse 10, 13125 Berlin, Germany.
Email: thoralf.niendorf@mdc-berlin.de

Funding information

This study was funded in part (T.K., T.N.) by the iNAMES - MDC-Weizmann Helmholtz International Research School (HIRS) for Imaging and Data Science from the Nano to the Meso.

Abstract

Renal pathologies often manifest as alterations in kidney size, providing a valuable avenue for employing dynamic parametric MRI as a means to derive kidney size measurements for the diagnosis, treatment, and monitoring of renal disease. Furthermore, this approach holds significant potential in supporting MRI data-driven preclinical investigations into the intricate mechanisms underlying renal pathophysiology. The integration of deep learning algorithms is crucial in achieving rapid and precise segmentation of the kidney from temporally resolved parametric MRI, facilitating the use of kidney size as a meaningful (pre)clinical biomarker for renal disease. To explore this potential, we employed dynamic parametric T₂ mapping of the kidney in rats in conjunction with a custom-tailored deep dilated U-Net (DDU-Net) architecture. The architecture was trained, validated, and tested on manually segmented ground truth kidney data, with benchmarking against an analytical segmentation model and a self-configuring no new U-Net. Subsequently, we applied our approach to in vivo longitudinal MRI data, incorporating interventions that emulate clinically relevant scenarios in rats. Our approach achieved high performance metrics, including a Dice coefficient of 0.98, coefficient of determination of 0.92, and a mean absolute percentage error of 1.1% compared with ground truth. The DDU-Net enabled automated and accurate quantification of acute changes in kidney size, such as aortic occlusion (−8% ± 1%), venous occlusion (5% ± 1%), furosemide administration (2% ± 1%), hypoxemia (−2% ± 1%), and contrast agent-induced acute kidney injury (11% ± 1%). This approach can potentially be instrumental for the development of dynamic parametric MRI-based tools for kidney disorders, offering unparalleled insights into renal pathophysiology.

KEYWORDS

deep learning, kidney, kidney size, MRI, parametric mapping, segmentation

Abbreviations used: ABSM, automated bean-shape model; ADPKD, autosomal-dominant polycystic kidney disease; AKI, acute kidney injury; CKD, chronic kidney disease; CM, contrast medium; CO, cortex; CPU, central processing unit; CRF, contrast reduction factor; DDU-Net, deep dilated U-Net; deoxyHb, deoxygenated hemoglobin; DL, deep learning; EMA, European Medicines Agency; FDA, US Food and Drug Administration; GFR, glomerular filtration rate; GPU, graphics processing unit; i.v., intravenous; IM, inner medulla; KS, kidney size; LAGeSo, Landesamt für Gesundheit und Soziales; MAPE, mean absolute percentage error; ML, machine learning; nnU-Net, no new U-Net; OM, outer medulla; PKD, polycystic kidney disease; pO₂, partial oxygen pressure; R, Pearson correlation coefficient; R², coefficient of determination; ReLU, rectified linear unit; USPIO, ultrasmall superparamagnetic iron oxide; ΔKS, change in kidney size.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *NMR in Biomedicine* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

With ever-increasing incidence, kidney diseases affect approximately 10% of the worldwide population, constituting a major public health burden.^{1–6} Despite the clear need, current options for the prevention and therapy of kidney disorders remain inadequate. A major reason for this is that the clinically available diagnostic and prognostic markers—such as serum creatinine-based metrics—only reveal kidney abnormalities at a late stage, and fail to detect early events in the pathological process. Several alternative blood- or urine-based diagnostic markers reflecting renal injury have been proposed.^{3,7–13} The modest performance of these markers has not yet delivered on their promise, and none has advanced to the point of use in clinical practice.^{8,11,12,14,15}

This major health issue requires noninvasive imaging to complement conventional approaches for the assessment of kidney diseases. Meeting this urgent need calls for synergistic approaches that include MRI.^{3,8,11,16–21} MRI allows for noninvasive imaging of kidney structure with whole organ coverage, and assessment of functional kidney features, with high spatial resolution.^{18,20–29} Many functional MRI methods have emerged that are sensitive to pathophysiological changes associated with renal hemodynamics, oxygenation, fibrosis, inflammation, and microstructure. MRI has the potential to provide quantitative imaging markers to detect various stages of pathophysiology, improve prediction and interception of disease progression, and evaluate treatment of renal disease.^{3,8,17,18,21,30–34}

Several renal pathologies are associated with changes in kidney size (KS), thus offering an opportunity for MRI-derived KS to be used as a parameter for the assessment of renal disease.^{14,35} In polycystic kidney disease (PKD), KS correlates with disease progression and the decline in glomerular filtration rate (GFR).³⁶ KS has been approved as a prognostic marker for use in clinical trials of new therapies for autosomal-dominant PKD.^{37,38} Detecting KS reduction because of parenchymal atrophy, sclerosis, and fibrosis has been recognized as a marker to identify chronic kidney disease (CKD) and to determine its severity.³⁹ Longitudinal MR-based KS monitoring is proposed as a key measure for several kidney disorders, including hyperfiltration in early diabetic nephropathy, renal transplants, renal artery stenosis, and vesicoureteral reflux.¹⁹

Serial MRI in experimental models has demonstrated that KS changes can indicate pathophysiological developments. These include models of clinical conditions such as acute ureteral obstructions because of urolithiasis or during upper urinary tract endourologic procedures, administration of X-ray contrast media (CM) for cardiac procedures, obstructions of the renal vein during partial nephrectomy or because of renal cell carcinoma-derived thrombus formation, and clamping of the suprarenal aorta or renal artery during surgery, or the low arterial target pressure during cardiopulmonary bypass.⁴⁰ MR-based KS assessment has also been used in experimental diabetes, mutant models mimicking PKD, and renal allografts.^{41–43} A recent preclinical study demonstrated that monitoring KS allows for physiological interpretation of MRI-based oxygenation changes in acute pathophysiologically relevant scenarios.⁴⁴

In vivo KS assessment requires a robust and fast segmentation of the kidney from MR images. Manual segmentation is time-consuming and prone to observer bias. This poses a major impediment for dynamic and longitudinal studies, and severely limits the potential for KS assessment in translational research. These constraints can be offset by analytical models using pre-established geometric shapes like ellipsoids or hybrid level-set methods.^{45–50} These approaches are semiautomated and provide up to 70-fold improvement in segmentation speed compared with manual segmentation.⁴⁰ Nevertheless, analytical model-based renal segmentation requires an expert observer to manually preselect measurements and/or landmarks to initialize the models.

Machine learning (ML)-based renal segmentation has gained momentum as a tool for decoding the links between KS as a mesoscopic marker, and data from histopathological, physiological, and functional measurements, with the goal of deciphering determinants and associations of renal disease.^{48,51,52} ML approaches provide a viable solution for deforming the kidney shape using a constrained statistical model-based algorithm trained upon a dataset.⁵³ This approach requires minimal user interaction. Fully automated supervised and unsupervised ML algorithms have been explored for renal segmentation from MRI using convolutional neuronal network models.^{53–58} MRI studies using neural networks for renal segmentation reported processing times as good as 1–10 s per subject.^{58–60} Most of these deep learning (DL) approaches take advantage of U-Net variants.^{61–65} Recent application of DL-based kidney segmentation has focused on automation of renal cyst and kidney volume measurements in healthy subjects and patients with autosomal-dominant PKD and CKD.^{54,56,60,66–70} The feasibility and reliability of dynamic or longitudinal MRI-based KS monitoring using DL in acute pathophysiological scenarios, where changes may be more subtle than in autosomal-dominant PKD or CKD, has not yet been investigated.

Recognizing this opportunity, the current study examines the feasibility and reliability of dynamic parametric MRI-based automated KS assessment in rats using DL. Addressing this challenge, a custom-tailored deep dilated U-Net (DDU-Net) was developed and validated to facilitate quantification of acute changes in KS (Δ KS) in pathophysiologically relevant experimental setups mimicking realistic clinical scenarios. We demonstrate that our approach has the potential for establishing MRI-based diagnostic tools for various kidney disorders, and for gaining new insights into the mechanisms of renal pathophysiology.

2 | EXPERIMENTAL

2.1 | Animal preparation

Investigations were approved by the Landesamt für Gesundheit und Soziales (LAGeSo) of Berlin in accordance with German Animal Protection Law and EU Directive 2010/63/EU. Male Wistar rats ($n = 52$, aged 12–13 weeks, 270–300 g, Harlan-Winkelmann, Borchon, Germany) were anesthetized with urethane (1.2 g/kg; Sigma-Aldrich, Steinheim, Germany) during MRI examinations. Fourteen animals were surgically implanted with MR-safe remote-controlled inflatable occluders, applied around the suprarenal aorta and the left renal vein, as previously described.^{71–73} During MR scanning, respiration and body temperature were monitored continuously.

2.2 | MRI experiments

Experiments were performed on a 9.4-T animal MR system (Bruker Biospec 94/20; Bruker Biospin, Ettlingen, Germany). A linear birdcage radio-frequency (RF) volume resonator (inner diameter = 72 mm; Bruker Biospin) was employed for transmission, and a four-channel surface RF coil array customized for rats was used for signal reception.

Following T_2 -weighted pilot scans, selective shimming of magnetic field homogeneity on a volume of interest accommodating the left kidney was conducted. T_2 mapping was performed with a respiratory-gated (Model 1025; SA Instruments, New York, NY, USA) multispin-echo technique: TR = 500 ms, number of echoes = 13, first TE = 6.4 ms, interecho time $\Delta TE = 6.4$ ms, number of averages = 1, and $t_{\text{acquisition}} = 58$ s. A midcoronal oblique slice was acquired: in-plane spatial resolution = $226 \times 445 \mu\text{m}^2$, field of view = $38.2 \times 50.3 \text{ mm}^2$, matrix size = 169×113 (zero-filled to 169×215), and slice thickness = 1.4 mm. Parametric maps of absolute T_2 were calculated by pixel-wise monoexponential fitting to the signal intensities of the T_2 -weighted images acquired at different echo times (in-house developed program; MATLAB, R2010a, MathWorks, Natick, MA, USA). T_2 maps were clipped at $T_2 = 0$ and 200 ms and rescaled to [0, 1].

The MR relaxation parameter T_2 was monitored to examine renal blood oxygenation as a surrogate of renal tissue oxygenation. These data were also used for dynamic assessment of the renal size. Dynamic T_2 mapping of the central coronal slice of the left kidney was performed in 43 rats under baseline conditions (without any intervention, at three subsequent time points), and in 30 of the 43 rats during or following the respective interventions and recovery periods, where applicable, according to the specific subgroup.

In addition, dynamic T_2 mapping was performed in another nine rats before and after bolus injection of the X-ray CM iodixanol into the thoracic aorta.⁴⁴ T_2 maps were obtained from an analogous MSME protocol (TR = 550 ms, number of echoes = 7, first TE = 10 ms, and interecho time $\Delta TE = 10$ ms).

2.3 | Longitudinal quantification of changes in KS upon pathophysiological interventions

Six pathophysiologically relevant interventions that alter renal tissue oxygenation were performed.^{71,72}

Rats with vascular occluders ($n = 14$) underwent serial T_2 mapping prior to, during, and following occlusion of the aorta ($n = 12$ successful). The occlusion was applied for 3.8 ± 0.3 min, then deflated, and the rats were allowed to recover for 7 min to ensure restoration of preocclusion hemodynamics and oxygenation. Time-of-flight MR angiography was performed immediately after inflation/deflation of the occluder to confirm occlusion/reperfusion of the vessels.^{71–73} After recovery, the same procedure was applied for renal venous occlusion ($n = 12$ successful) and for combined aortic-venous occlusion ($n = 8$ successful). Occlusion of the aorta results in decreased renal blood volume, occlusion of the renal vein induces an increase in renal blood volume, and simultaneous occlusion of both vessels does not affect renal blood volume.

A separate group of five rats, received an intravenous (i.v.) bolus of 5 mg/kg furosemide (Ratiopharm, Ulm, Germany) followed by a 0.2-mL saline chaser.⁴⁴

In a separate group, 11 rats underwent T_2 mapping during a control period of normoxia (inspiratory oxygen fraction [FiO_2] of 21%) during 3.8 ± 0.1 min of hypoxemia ($\text{FiO}_2 = 10\%$), followed by 10 min of recovery at normoxia.

A final group of nine rats received an i.v. bolus of 1.5 mL of the X-ray CM iodixanol (320 mg/mL iodine, Visipaque; GE Healthcare Buchler, Braunschweig, Germany), followed by a 0.2-mL saline chaser. Administration of CM and furosemide are expected to have longer term effects, increasing the renal tubular volume fraction.

For the occlusion of the suprarenal aorta, occlusion of the renal vein, and the bolus administration of furosemide, interventions were repeated following i.v. administration of ultrasmall superparamagnetic iron oxide (USPIO), 2 mg Fe/kg, ferumoxylol (Feraheme; AMAG Pharmaceuticals, Lexington, USA).⁷³ Ferumoxylol reduces the signal intensity and image contrast due to T_2 shortening,⁷³ but does not affect renal physiology at this dose.⁷⁴

2.4 | Manual segmentation of the kidney by independent observers

Accuracy and precision of DDU-Net was evaluated versus the ground truth KS under baseline physiological conditions, based on manual segmentation of T_2 maps by a reader with 1 year of experience in renal MRI analyses in rats. Following the manual segmentation procedure, a consensus segmentation involving the five expert observers was conducted to determine the ground truth. The manual segmentation prepared by the first observer was set as a starting point for the consensus segmentation. Subsequently, each of the 129 T_2 maps was presented to all observers in ITK-SNAP together with a transparent overlay of the area determined as belonging to the kidney. Training was performed using two T_2 maps as examples: one T_2 map with very good image quality (classified as acceptable), and one T_2 map with lower image quality (classified as marginal). From these training data, a consensus was reached regarding the inclusion/exclusion of voxels with apparent partial volume effects. For the remaining consensus reading, renal segmentation T_2 maps were used. During the consensus reading, the observers remained blinded to the individual T_2 map number, the animal identification, the numerical KS derived from the independent renal segmentation, and the number of voxels of the overlay. Upon presentation of each T_2 map along with the corresponding transparent overlay, all observers agreed in real time if it was necessary to add or delete (a) certain voxel(s) to/from the overlay, to improve the accuracy of the renal boundaries, whereupon a consensus on the total area assigned as renal tissue was reached. Changes were adopted in real time using ITK-SNAP. A time limit of 90 s was set for the consensus renal segmentation of each T_2 map. The ground truth comprised consensus readings of manual segmentations from a total of 129 T_2 maps ($n = 43$ rats, with three baseline scans per rat).

2.5 | Data handling and data simulation

The 52 rats used in the study were split into five subgroups (Figure 1). In the first subgroup ($n = 13$), the training set, three baseline T_2 maps were obtained per rat, and used for training and validation of DDU-Net. Another cohort ($n = 30$), the test set, was divided into three subgroups according to the interventions applied, as described above. For these animals, three baseline T_2 maps were acquired prior to the intervention, followed by serial T_2 maps obtained during the intervention and recovery period; the baseline T_2 maps were used for training, validation, and testing of the neural network against the ground truth manual segmentation. The serial T_2 maps regarding the intervention were used for the application of the network. For the remaining cohort ($n = 9$), the longitudinal data following administration of the X-ray CM were included.^{71,72,75}

For the definition of the ground truth, all three baseline T_2 maps of the $13 + 30 = 43$ animals were anonymized. For these 43 rats $\times 3 = 129$ T_2 maps, manual segmentation was performed as described in on 2.4. Figure 2 shows the result of the manual segmentation of an exemplary renal T_2 map, compared with the segmentations predicted by the proposed DDU-Net, the self-configuring no new U-Net (nnU-Net), and an automated bean-shape model (ABSM) used in a prior study.⁴⁰

Based on the ground truth, 12 rats of the test set with baseline T_2 maps showing the highest signal-to-noise ratio (SNR) were selected to simulate T_2 maps emulating the interventions. The SNR was defined as

$$\text{SNR} = \frac{0.9535}{\text{median}(\text{abs}(\text{diff}(S_{\text{ROI}})))}, \quad (1)$$

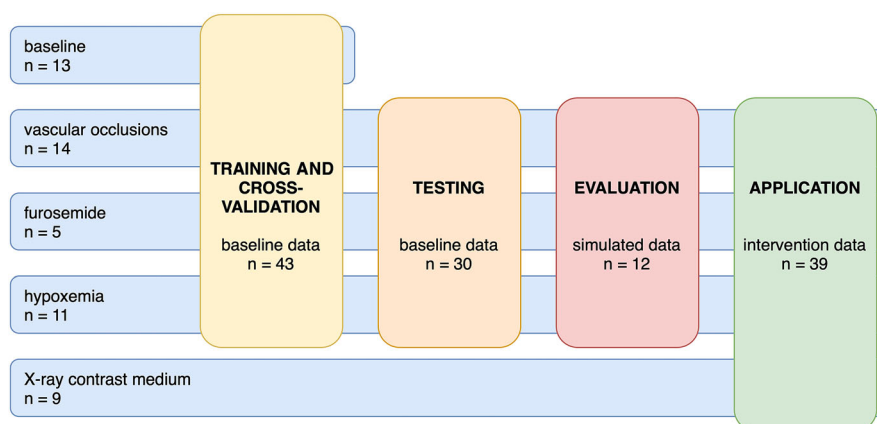


FIGURE 1 Overview of the data handling: the total number of 52 rats were split into five distinct subgroups (blue) as shown in the figure. While the baseline subgroup only was used during the training and the cross-validation (yellow) and the subgroup of the administration of the X-ray contrast media (CM) only during the application on the in vivo longitudinal MRI data (green), the remaining subgroups were used during all steps of the study, including the testing (orange) and the evaluation (red).

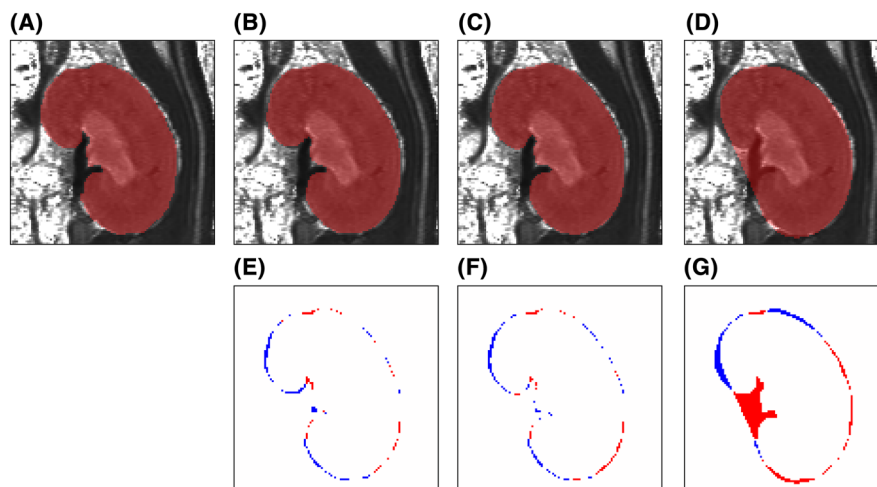


FIGURE 2 Exemplary overview of renal segmentation: (A) Manual segmentation obtained from the consensus reading and segmentations predicted by (B) Deep dilated U-Net (DDU-Net), (C) No new U-Net (nnU-Net), and (D) The automated bean-shape model (ABSM). Difference maps of the predicted segmentation versus the ground truth obtained from the consensus reading are shown for (E) DDU-Net, (F) nnU-Net, and (G) The ABSM. Although the manual and the deep learning (DL)-based segmentations accurately include the pelvis and exclude the extrarenal vessels, the analytical model (D) draws a straight border from both poles of the kidney.

with abs referring to the absolute value function, $diff$ calculating the differences between adjacent signal intensities, and S_{ROI} the signal intensities inside the region of interest (ROI).

A series of five copies randomly chosen from the three baseline T_2 maps for each subject was created for all combinations of 10 different changes in KS ranging from -5% to 5% , 10 SNR levels of the T_2 maps ranging from 2.5 to 50, and four tissue contrast reduction factors (CRFs) ranging from 1 to 0.25.

To generate different SNR levels, a nonlocal means filter was applied to ensure that the intended SNR was higher than the original SNR. For lower SNR, the image was manipulated as follows:

$$S_{\text{simulated}} = \exp \left(\log(S_{\text{original}}) + X_{\text{normal}} \cdot \sqrt{\frac{0.9535}{\text{SNR}_{\text{simulated}}} - \frac{0.9535}{\text{SNR}_{\text{original}}}} \right). \quad (2)$$

Here, $S_{\text{simulated}}$ and S_{original} refer to the simulated and original signal intensity, and X is a random variable with the standard normal distribution. The contrast reduction was performed by applying the following formula:

$$S_{\text{simulated}} = \text{CRF} \cdot (S_{\text{original}} - \mu_S) + \mu_S. \quad (3)$$

Here, μ_S denotes the mean of the original signal intensities.

These simulations resulted in 400 series of five simulated images for each of the five randomly chosen baseline maps, for each of the 12 rats: one baseline map, three maps right after an intervention exponentially recovering back to the baseline, and a final baseline map. This set of $5 \times 12 \times 400 = 24,000$ synthetically generated T_2 maps was used for benchmarking DDU-Net against the ABSM and nnU-Net.

For automatic DDU-Net based quantification of KS changes in in vivo longitudinal MRI data, the serial T_2 maps of the rats that underwent the interventions were used ($n = 30 + 9$).^{71,72,75}

2.6 | Architecture of DDU-Net

Our DDU-Net incorporates the general structure of the original U-Net and consists of an encoder and a decoder part with four resolution levels each.⁶¹ While the decoder part remains the same, the two original convolutional layers in each level of the encoder are replaced by a sequence of

four convolutional layers with growing dilation and residual connections. This results in an increased receptive field and an asymmetric network architecture with a larger encoding part.^{62,65}

A schematic representation of the network architecture is shown in Figure 3A. The numbers of filters are 16, 32, 64, and 128, and 256 in the bottleneck. Each level of the encoding part consists of two 3×3 convolutional layers without dilations, two 3×3 convolutional layers with dilations of 2 and 4, respectively, and a final downsampling 2×2 convolutional layer with stride of 2. All convolutional layers are followed by a leaky rectified linear unit (Leaky ReLU) with a slope of 0.1 and a batch normalization. The outputs of the first four convolutional layers are concatenated and function as the input of the final downsampling layer. The composition of an encoding block is illustrated in Figure 3B. The bottleneck of the network consists of two 3×3 convolutional layers (followed by Leaky ReLU and batch normalization) without dilations. For the decoding part, the original U-Net structure is maintained: each level starts with an upsampling 2×2 transposed convolutional layer with stride of 2, followed by two 3×3 convolutional layers without dilations. As before, all convolutional layers succeed a Leaky ReLU and a batch normalization. In each level, the concatenated layer outputs of the corresponding encoding stage are concatenated to the output of the transposed convolutional layer,

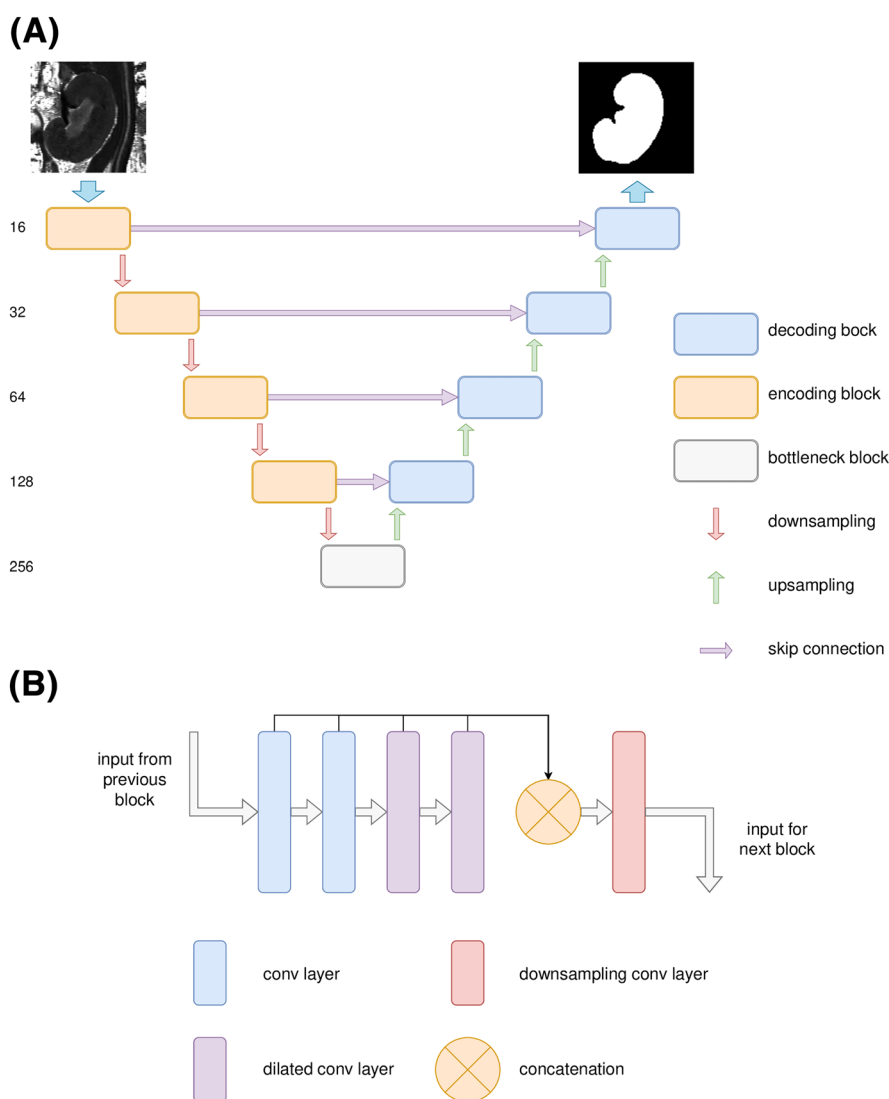


FIGURE 3 Schematic representation of the deep dilated U-Net (DDU-Net) and its encoding blocks: (A) The general structure of the DDU-Net is illustrated with its encoding blocks (orange), decoding blocks (blue), and the bottleneck (gray). The red arrows represent the downsampling, the green arrows the upsampling. Violet arrows show the skip connection between encoding and decoding blocks. Numbers on the left represent the numbers of channels of each level of encoding/decoding. (B) A schematic illustration of the encoding blocks is shown with two regular 3×3 convolutional (conv) layers without dilation (blue), followed by two dilated conv layers (violet) with dilations of 2 and 4. The concatenation of the output of those four layers is shown in orange. The 2×2 downsampling conv layer is represented in red. Because of the extra depth and dilations, the receptive field size created by one encoding block (including downsampling) is 16 compared with 6 by the original encoding block of the U-Net. This corresponds to an expansion of more than 2.5 times.

and serve as the input of the first convolutional layer. The final layer of the decoder is a 1×1 transposed convolutional layer with a sigmoid activation function.

Because the network design presumes an input shape of 192×224 because of architectural reasons, the input was rescaled to 192×224 and the output rescaled back to 169×215 .

2.7 | Training, data augmentation, and postprocessing

DDU-Net was pretrained using the kidney labels and MRI data from the CHAOS dataset.⁷⁶ For the actual training, the subjects of the test set ($n = 30$) were randomly split into three distinct test subsets of size $n_{\text{test}} = 10$. For each subset, a 5-fold cross-validation of DDU-Net was performed based on the two remaining subsets ($n = 20$) and the subgroup that was only used for training and validation ($n = 13$). This resulted in training sets of sizes $n_{\text{train}} = 26$ and $n_{\text{train}} = 27$ and validation sets of sizes $n_{\text{validation}} = 7$ and $n_{\text{validation}} = 6$. To achieve a more robust prediction and a lower generalization error, the five models of each cross-validation were combined into a 5-fold DDU-Net ensemble: the output logits before the final sigmoid activation were averaged, passed through the sigmoid activation function and rounded to zero and one, resulting in binary segmentation masks. For subjects not included in the training or validation (rats administered X-ray CM), networks of all cross-validations were combined into a 15-fold ensemble. Each single network was trained with a stochastic gradient descent optimizer with an initial learning rate of $1e^{-2}$ and a momentum of 0.8, and a loss function combining the binary focal loss and the Dice loss. The learning rate was reduced by a factor of 0.5 to a minimal learning rate of $1e^{-4}$ in case the validation loss did not improve for 50 epochs. Training was performed for a maximum number of 700 epochs, switching to stochastic weight averaging after 500 epochs to improve the generalization of the model.⁷⁷ The schedule used for stochastic weight averaging was the cyclic cosine annealing from the initial learning rate to $1e^{-8}$ with a duration of 25 epochs.

Each epoch, every training map underwent a data-augmentation procedure including random affine transformations (x- and y-zoom, x- and y-shear, rotation, shift), elastic transformation, randomized manipulations of the signal intensities (brightness, contrast, and gamma adjustment), and a random Gaussian noise.

After prediction by the network, the output masks were cleaned automatically and only the largest segmentation region was kept. Smaller distinct regions, often representing parts of the other kidney, were removed from the mask.

Preprocessing, network design, training, postprocessing, and evaluation were programmed in Python (PyTorch 1.13).

2.8 | Validation of DDU-Net

For validation, DDU-Net was benchmarked against nnU-Net and the ABSM. To benchmark DDU-Net against a state-of-the-art DL approach, we trained nnU-Net based on the exact same distinct testing subsets. During the training, nnU-Net performs a 5-fold cross-validation by default for several network configurations and chooses the best model automatically. For the cross-validation, the exact same splits were used as for DDU-Net. The resulting trained nnU-Net was applied on the baseline test set, the simulated dataset, and the intervention data. As for DDU-Net, all predicted masks were cleaned as described above. For the baseline maps of the test sets, the Dice scores between the manual segmentations and the predicted masks by DDU-Net and nnU-Net were calculated. Furthermore, the R^2 of the renal area obtained from manual segmentations and DDU-Net/nnU-Net was determined and linear regression was performed for all three baseline time points. To meet the assumption of normality, two rats with unusually small KS of less than 180 mm^2 were identified as outliers and excluded from this analysis. Relative residuals were plotted and the mean absolute percentage errors (MAPEs) were determined.

To further benchmark DDU-Net, we used an established analytical ABSM. Although the manual segmentation and both DL-based predictions include the papilla while excluding the incoming and outgoing vessels, the ABSM is a geometrical model of a skewed ellipse with a straight line (border) drawn that connects the outermost points of the upper and lower medial curvatures of the kidney (Figure 2D).⁴⁰ Therefore, the three methods can only be compared by evaluating their prediction of relative changes in KS.

For the validation of DDU-Net, synthetic data were generated as described in section 2.5. KS was determined by all three methods and the relative size of the peak of the intervention was compared with the calculated baseline. By considering only the relative KS changes upon an intervention, all three methods could be compared, regardless of the differing segmentation policies of the DL approaches and the ABSM. Based on this relative KS, DDU-Net, nnU-Net, and ABSM were compared by calculating the respective MAPEs against the ground truth change.

2.9 | Statistical analysis

For the in vivo study of the intervention-induced changes in ΔKS , KS assessed in the T_2 mapping before the respective intervention was used as the control KS and ΔKS was calculated for the subsequent T_2 maps. T_2 values for the renal cortex (CO), outer medulla (OM), and inner medulla

(IM) were determined using baseline T_2 maps before the respective interventions and the T_2 map during or right after the respective intervention. The placement of ROIs was conducted using a standardized semiautomatic approach, following a previously described method.⁷⁸ This method involved positioning the ROIs (five for the CO and OM, and three for the IM) in a way that excludes the transitional regions between renal layers, ensuring that partial volume effects are avoided. Because of the T_2 shortage induced by ferumoxytol, T_2 values were not analyzed after the administration of USPIO.

Relative changes in KS over time and changes in T_2 were analyzed using the nonparametric repeated-measures Friedman test, followed by Dunn's post-hoc test with the Benjamini–Hochberg correction for multiple comparisons. p values less than 0.05 were considered to be statistically significant. All statistical analysis was performed in R v. 3.6.3 using the libraries “rstatix” and “dunn.test”.

3 | RESULTS

3.1 | DDU-Net evaluation on the ground truth obtained from manual segmentation of the kidney

For the baseline T_2 maps of the test set, the mean KS predicted by DDU-Net was similar to the mean KS derived from the ground truth manual segmentation: $211 \pm 1 \text{ mm}^2$ (mean \pm standard error of the mean [SEM]). Figure 4A–C depict the linear regression plots for the KS predicted by DDU-Net versus the ground truth for each of the three baseline T_2 maps. A mean coefficient of determination (R^2) = 0.92, a mean Pearson correlation coefficient (r) = 0.96, and an average MAPE = 1.1% were determined over the three baseline T_2 maps. Figure 4D shows the distribution of relative residuals between both methods. DDU-Net yielded a median intrasubject variability of 0.3% compared with 0.8% for the ground truth and 0.6% for the previously published ABSM.⁴⁰ For the segmentations, DDU-Net reached an accuracy of more than 0.99, a recall of 0.98, a

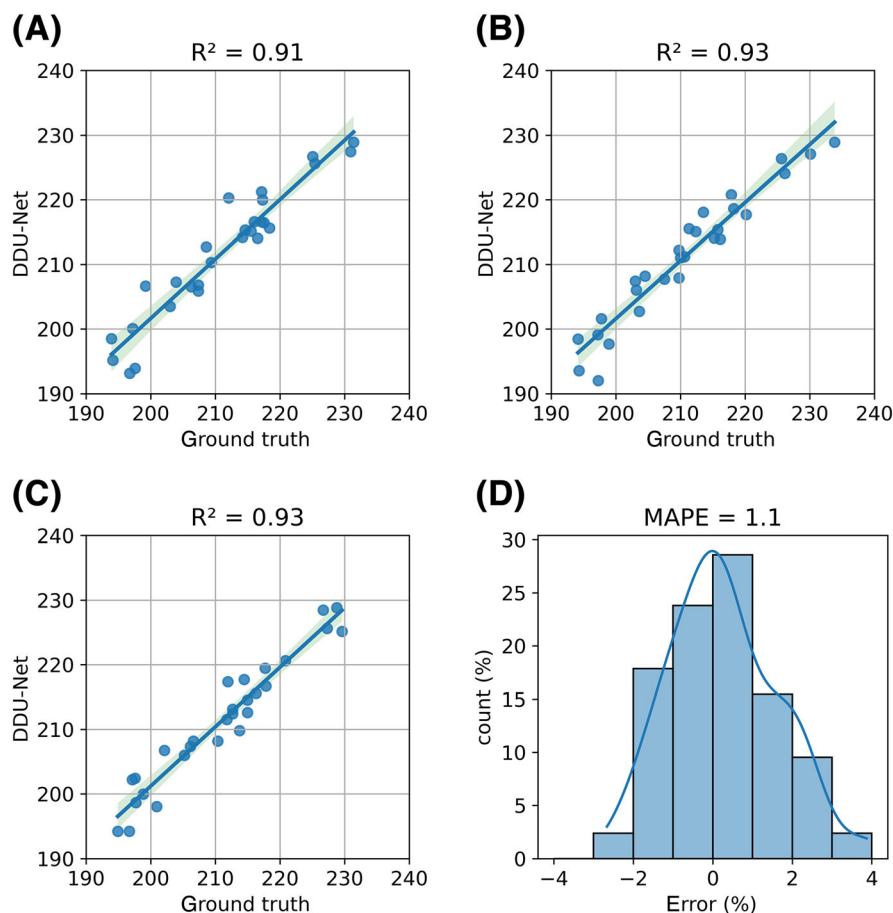


FIGURE 4 Performance of deep dilated U-Net (DDU-Net) on the baseline test set. (A–C) For the ground truth kidney size (KS) and the KS prediction by DDU-Net, linear regression plots (blue) including the 95% confidence interval (green) are shown for the three baseline timepoints. To meet the assumption of normality, two outlier rats with a KS of less than 180 mm^2 were excluded from this analysis. (D) The distribution of the percentile error between the ground truth and the predicted KS for all three baseline maps yielding an averaged mean absolute percentage error (MAPE) of 1%.

precision of 0.98, a mean intersection-over-union of 0.96, and a Dice score of 0.98. Renal T_2 mapping under baseline conditions revealed a median T_2 of 47 ± 1 ms in the CO, 46 ± 2 ms in the OM, and 66 ± 5 ms in the IM.

On the ground truth baseline data, we further benchmarked DDU-Net against the self-configuring state-of-the-art nnU-Net. nnU-Net-predicted area achieved a mean $R^2 = 0.91$, a mean $r = 0.95$, an averaged MAPE = 1.2%, and a median intrasubject variability of 0.3%. For the predicted segmentations, nnU-Net reached an accuracy of more than 0.99, a recall of 0.98, a precision of 0.98, a mean intersection-over-union of 0.96, and a Dice score of 0.98.

3.2 | Validation of DDU-Net against the ABSM and nnU-Net using simulated time series data with changes in renal size

Fitting the 24,000 simulated T_2 maps took less than 4 min on a graphics processing unit (GPU) and less than 30 min on a central processing unit (CPU) for DDU-Net and nnU-Net, compared with approximately 100 h for the ABSM. This translates to a speed gain of ~ 1500 for the DL-based approaches versus the geometric model. Figure 5 depicts the MAPEs of the relative KS for all combinations of SNRs, CRFs, and changes in relative KS for all three methods. In 292 of the 400 scenarios (73%), DDU-Net outperformed the ABSM in terms of MAPE (one-sided binomial $p < 10^{-20}$). While failing for the lowest (nonrealistic) SNR and lowest contrast compared with the ABSM (Figure 5), for SNRs of more than 2.5 and CRFs of more than 0.25, DDU-Net was superior in 240 of the 270 scenarios (89%) ($p < 10^{-41}$) and achieved a median MAPE of 1.0% compared with 1.5% for the ABSM. DDU-Net was superior to nnU-Net in 205 of 400 scenarios (51%) ($p > 0.05$). For realistic SNRs of more than 2.5 and CRFs of more than 0.25, DDU-Net outperformed nnU-Net in 155 of the 270 scenarios (57%) ($p < 10^{-2}$). The median MAPE achieved by nnU-Net was 1.1%, slightly higher than DDU-Net.

3.3 | In vivo quantification of changes in renal size upon pathophysiologically relevant interventions

The various interventions the animals underwent altered the signal intensity and contrast of the respective T_2 maps. KS quantified by the DDU-Net and by the nnU-Net showed similar longitudinal changes in renal size for the vascular occlusions, for the furosemide application and for

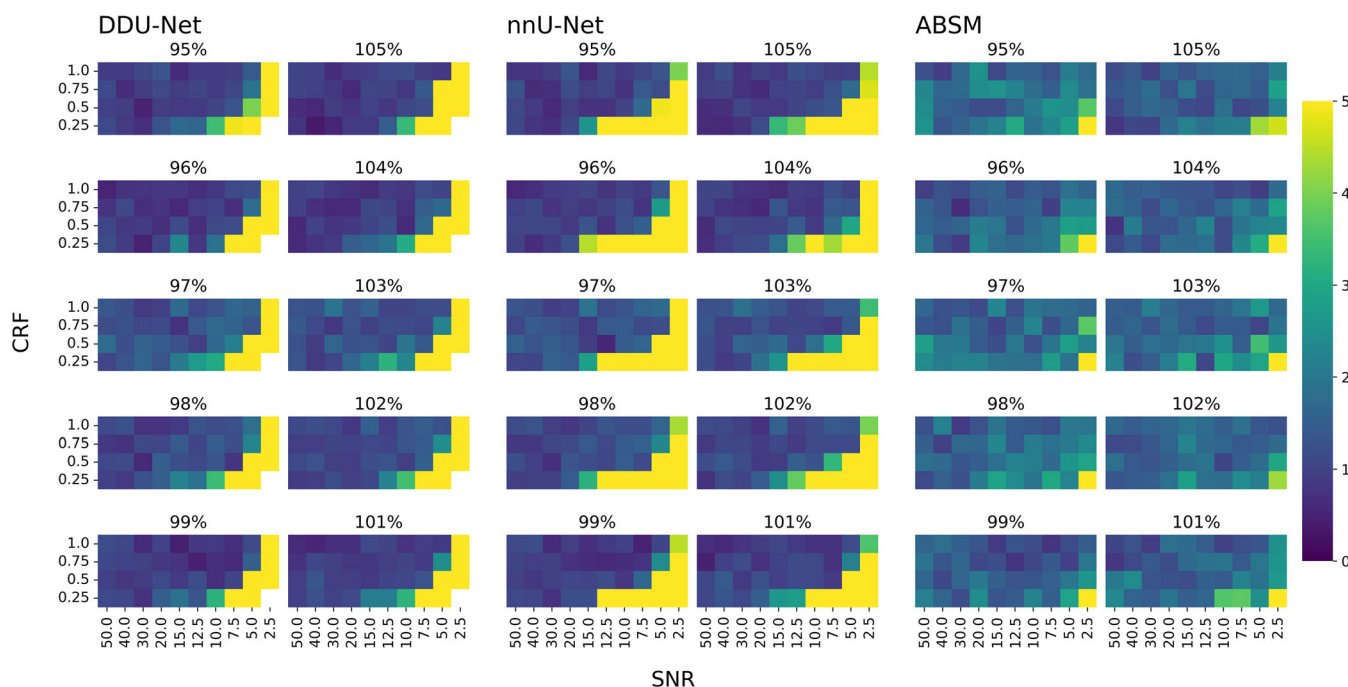


FIGURE 5 Assessment of deep dilated U-Net (DDU-Net), no new U-Net (nnU-Net), and the automated bean-shape model (ABSM) approach using simulated data and the metric mean absolute percentage errors (MAPEs). The figure summarizes MAPEs achieved by DDU-Net, nnU-Net, and the ABSM for every combination of signal-to-noise ratio (SNR) (from 2.5 to 50), contrast reduction factor (CRF) (from 0.25 to 1.0), and relative kidney size (KS) (from 95% to 105%). White spots indicate a failure to segment the kidney for simulated data using very low SNRs, which are not realistic for real-world MRI examinations.

hypoxemia. However, nnU-net KS assessment demonstrated higher SEM. For the low SNR data obtained upon administration of the X-ray CM, nnU-Net could not confirm the trend shown by DDU-Net. A closer examination of the nnU-Net segmentations revealed mis-segmented areas. Examples of suboptimal performance or failure of nnU-Net segmentation on the in vivo MRI data are shown in Figure 6. Examples of mis-segmentation include shifts of the segmented regions, extension of the boundary into adjacent tissue and organs, and inclusion of both kidneys or segmentation of the contralateral kidney by DDU-Net. Because of this limitation, the delta KS graphs obtained from nnU-Net are erroneous.

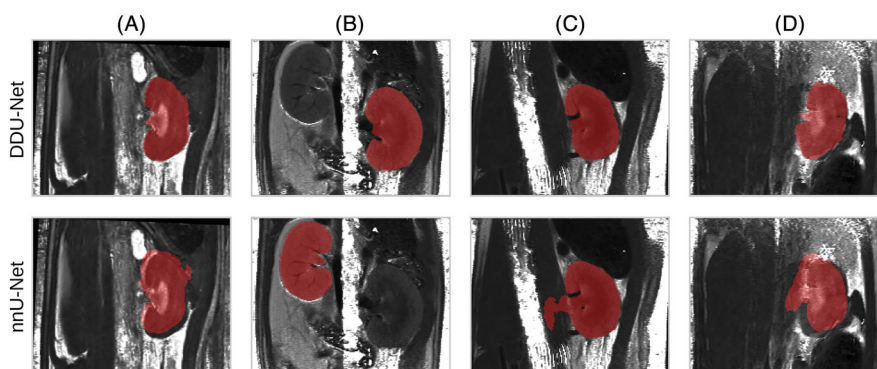


FIGURE 6 Exemplary T_2 maps and the corresponding prediction by deep dilated U-Net (DDU-Net) and no new U-Net (nnU-Net). The T_2 maps illustrate cases in which nnU-Net failed to correctly segment the kidney, while DDU-Net performed effectively. (A) The whole mask predicted by nnU-Net was shifted and expanded. (B) An exemplary T_2 map including both kidneys. While both networks were trained on the central slice of the left kidney, nnU-Net incorrectly segmented the right kidney even although the texture differs from the central slices. (C) The segmentation of nnU-Net extended irregularly far beyond the pelvic region, and in (D) The prediction was shifted and expanded in the pelvic region. These kinds of failure were not observed for DDU-Net.

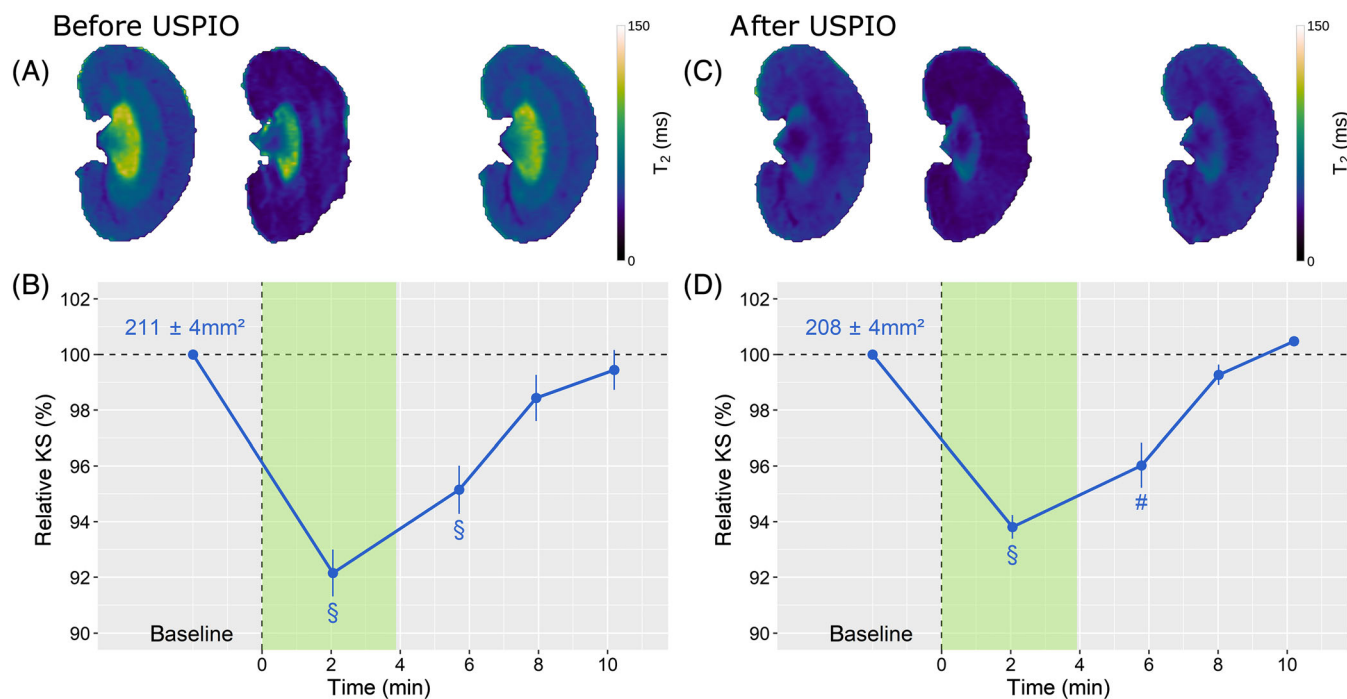


FIGURE 7 Relative kidney size (KS) over time for the occlusion of the suprarenal aorta. (A) Exemplary T_2 maps are shown before (left) and during (center) the intervention, as well as after recovery (right). (B) Development of relative KS changes upon the occlusion of the suprarenal aorta before the administration of ultrasmall superparamagnetic iron oxide (USPIO). The blue line depicts the mean of each time point and the corresponding interval of the standard error of the mean (SEM) ($n = 12$ rats). The mean KS at baseline was $211 \pm 4 \text{ mm}^2$. The duration of the occlusion is highlighted in green. (C) Analogous exemplary T_2 maps of the same rat. (D) Time course of relative KS changes after administration of USPIO with a mean KS at baseline of $208 \pm 4 \text{ mm}^2$. §: $p < 0.001$, #: $p < 0.01$.

Following aortic occlusion, DDU-Net–based segmentation revealed a significant reduction in KS of $-8\% \pm 1\%$ before (Figure 7A,B) and $-6\% \pm 0\%$ after USPIO (Figure 7C,D), accompanied by T_2 reductions of $22\% \pm 2\%$ in the CO, $27\% \pm 2\%$ in the OM, and $13\% \pm 3\%$ in the IM, highlighting the impact of altered blood flow on renal physiology.

Similarly, during occlusion of the left renal vein, DDU-Net segmentation detected a significant increase in KS of $5\% \pm 1\%$ before (Figure 8A,B) and $6\% \pm 1\%$ after USPIO (Figure 8C,D), along with pronounced T_2 reductions of $33\% \pm 2\%$ in the CO, $39\% \pm 2\%$ in the OM, and $32\% \pm 6\%$ in the IM, reflecting the interplay between venous outflow and KS dynamics, as well as tissue oxygenation.

Additionally, simultaneous aortic-venous occlusion showed no significant change in KS, but immediate declamping revealed a significant decrease in KS of $-3\% \pm 0\%$ (Figure 9A,B), corresponding to T_2 reductions of $23\% \pm 1\%$ in the CO, $28\% \pm 1\%$ in the OM, and $20\% \pm 3\%$ in the IM, further emphasizing the acute response of the kidneys to vascular adjustments.

Furthermore, furosemide administration induced a significant KS increase of $2\% \pm 1\%$ both before (Figure 10A,B) and following USPIO (Figure 10C,D), accompanied by T_2 increases of $11\% \pm 2\%$ in the CO and $19\% \pm 2\%$ in the OM, indicating the functional impact of diuretics on renal fluid dynamics. The KS remained significantly elevated throughout the observation period without USPIO (Figure 10A,B), while even with reduced contrast, a significant furosemide-induced KS increase was detected at one timepoint, providing insights into the persistence of its effects (Figure 10C,D).

In the context of hypoxemia, DDU-Net detected a significant decrease in KS of $-2\% \pm 1\%$ (Figure 11A,B), with corresponding T_2 reductions of $18\% \pm 3\%$ in the CO, $15\% \pm 3\%$ in the OM, and $8\% \pm 3\%$ in the IM, revealing the response of the kidneys to altered oxygenation levels. Notably, following the administration of the X-ray contrast agent, DDU-Net detected a significant increase in KS of up to $11\% \pm 1\%$ (Figure 11C,D), while T_2 increased by $14\% \pm 3\%$ in the CO and by $17\% \pm 5\%$ in the OM, and decreased by $17\% \pm 7\%$ in the IM, illustrating the physiological impact of contrast-induced acute kidney injury (AKI) and its effects on tissue perfusion and oxygenation throughout the 60-min observation period.

These findings demonstrate the comprehensive physiological insights provided by combining DDU-Net–based segmentation and T_2 mapping, enabling a deeper understanding of renal pathophysiology under diverse experimental conditions.

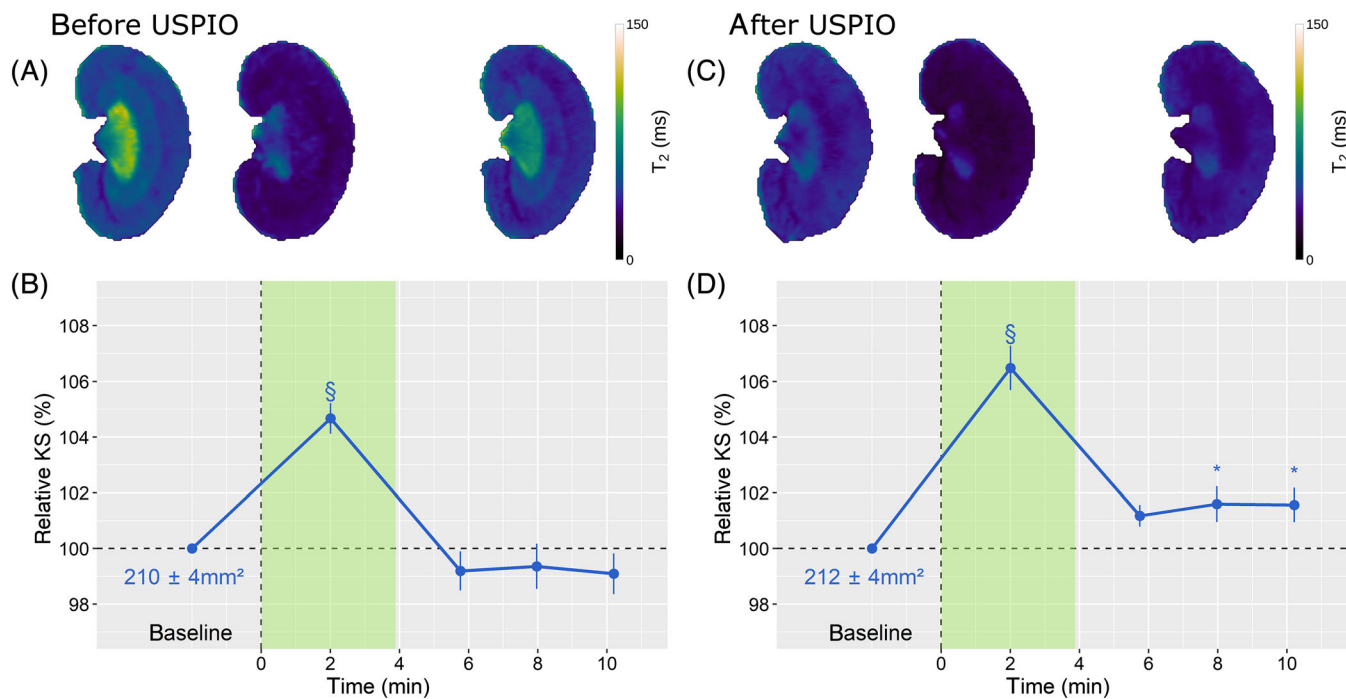


FIGURE 8 Relative kidney size (KS) over time for the renal venous occlusion. (A) Exemplary T_2 maps acquired before (left) and during (center) the intervention, as well as during recovery (right). (B) Development of relative KS changes upon the occlusion of the left renal vein before the administration of ultrasmall superparamagnetic iron oxide (USPIO). The blue line depicts the mean of each time point and the corresponding interval of the standard error of the mean (SEM) ($n = 12$ rats). The mean KS at baseline was $210 \pm 4 \text{ mm}^2$. The intervention is highlighted in green. (C) The analogous exemplary T_2 maps of the same rat, and (D) Time course of relative KS changes after administration of USPIO with a mean KS at baseline of $212 \pm 4 \text{ mm}^2$. §: $p < 0.001$; *: $p < 0.05$.

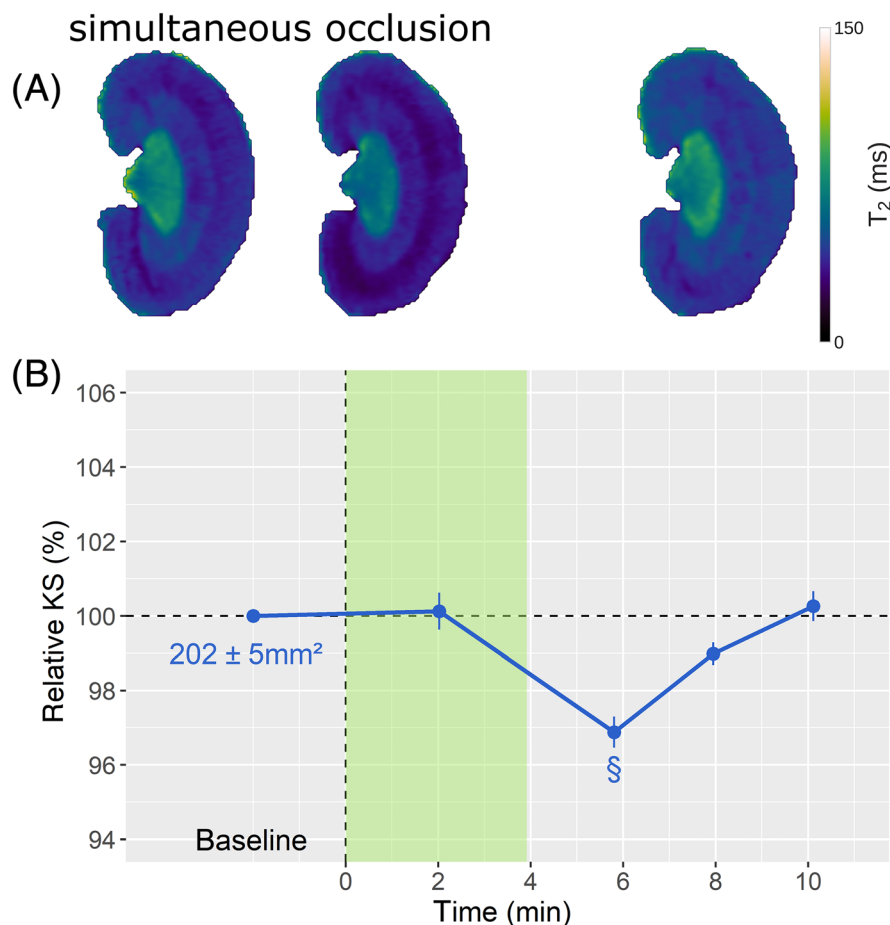


FIGURE 9 Relative kidney size (KS) over time for the simultaneous occlusion of the suprarenal aorta and the vena renalis. (A) Exemplanary T_2 maps derived before (left) and during (center) the intervention, as well as after recovery (right). (B) Development of relative KS changes upon the simultaneous occlusion of the suprarenal aorta and the left renal vein. The blue line depicts the mean of each time point and the corresponding interval of the standard error of the mean (SEM) ($n = 8$ rats). The mean KS at baseline was $202 \pm 5 \text{ mm}^2$. The intervention is highlighted in green. §: $p < 0.001$.

4 | DISCUSSION

This work demonstrates the feasibility of automated quantification of acute changes in KS of the rat kidney using segmentation of parametric MR images with a DDU-Net. The feasibility and reliability of DDU-Net-based longitudinal KS monitoring is demonstrated on in vivo T_2 mapping data, with acute pathophysiological interventions performed directly while the animal was inside the MRI scanner. Despite the limited number of subjects involved in the training of the network, data augmentation, ensemble training, and stochastic weight averaging permitted successful optimization of the neural network. When benchmarked against the ground truth, our DDU-Net achieved a highly accurate segmentation of the kidney (Dice = 0.98) and quantification of KS ($R^2 = 0.92$). With a MAPE of 1.1%, DDU-Net facilitated accurate and precise prediction of KS, and outperformed the state-of-the-art nnU-Net. By reproducing the observed KS changes before and after the administration of USPIO, DDU-Net demonstrated resilience towards severely reduced signal intensities, SNRs and image contrasts. The results from the subgroup that received the X-ray CM reinforce the ability of DDU-Net to overcome a data shift resulting from a different image acquisition protocol.

To date, the application of our DDU-Net has been focused on retrospective assessment of changes in KS during acute pathophysiological scenarios. Future steps include prospective and real-time application of DDU-Net, integrating our DDU-Net library directly into the image reconstruction and postprocessing pipeline of the MR scanner. While the CPU computation time was ~ 70 ms per segmentation, GPU implementation improved the computation time to 10 ms. These computation times align with the temporal resolution of preclinical renal T_2 mapping, in the range of 10–90 s per T_2 map; thus, DDU-Net is suitable for on-the-fly assessment of KS. While it might be conceptually appealing to extend the DDU-Net approach to three-dimensional (3D) datasets, whole kidney coverage T_2 mapping comes at the cost of increased acquisition time, which is a temporal resolution constraint for dynamic studies. In the preclinical experimental context, the temporal resolution required for longitudinal

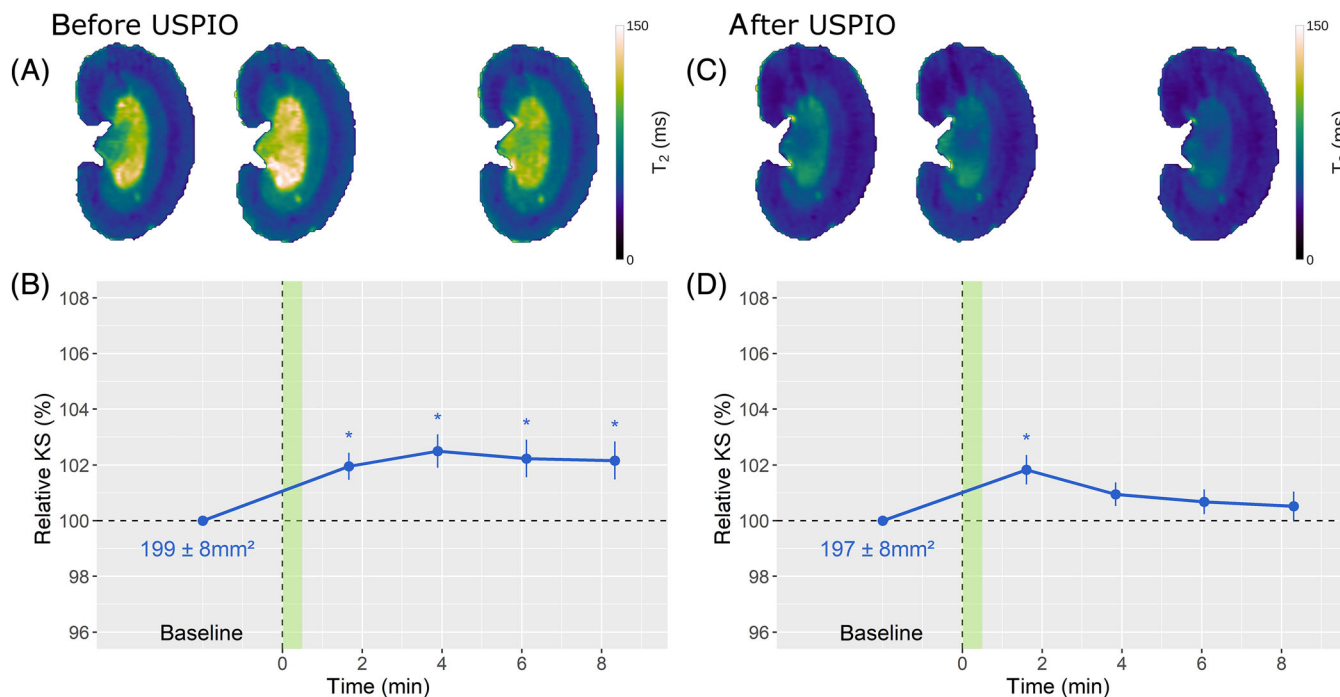


FIGURE 10 Relative kidney size (KS) over time upon the administration of furosemide. (A) Exemplary T₂ maps are shown before (left) and following the intervention (center and right). (B) Development of relative KS changes upon bolus injection of furosemide before the administration of ultrasmall superparamagnetic iron oxide (USPIO). The blue line depicts the mean of each time point and the corresponding interval of the standard error of the mean (SEM) ($n = 5$ rats). The mean KS at baseline was $199 \pm 8 \text{ mm}^2$. The time of injection is highlighted in green. (C) Analogous exemplary T₂ maps of the same rat, and (D) Development of relative KS changes after the administration of USPIO with a mean KS at baseline of $197 \pm 8 \text{ mm}^2$. *: $p < 0.05$.

experiments with physiological interventions such as those used in the current study would be severely constrained by the longer acquisition times that would be needed to obtain 3D data.

nnU-Net has been previously applied on longitudinal MRI data, including assessment of lesion size and counts in patients with multiple sclerosis, and on cardiac chamber quantification.⁷⁹ However, these cases generally do not include substantial reductions in image contrast or SNR. Our study illustrates the limitations of nnU-Net when dramatic changes do occur in the distribution of the data, signal intensities, or contrast of the ROI. We showed that nnU-Net failed to correctly segment the kidney in several cases with contrast changes, thereby compromising KS measurement. Conversely, our DDU-Net performed well in scenarios of reduced signal intensities, SNRs, and image contrasts. Furthermore, DDU-Net achieved a more precise determination of KS in terms of MAPE and R^2 . While residual connections and dilations are possible options in the self-configuration of nnU-Net, the principles of our DDU-Net are not yet implemented in a self-configuring approach.^{79,80} Several characteristics considered in our approach could be added to the portfolio of nnU-Net self-configuration: increased depth of the network, additional dilated convolutions, intra-level skip connections, concatenation of the layer output, and downsampling via strided convolution. Adding these configurations may reduce the number of channels needed in the network, and would decrease the memory needed for the model. In our case, the DDU-Net ensemble needed 135 MB, compared with 450 MB for nnU-Net, a reduction of 70%.

Fast, automated, and accurate kidney segmentation was facilitated by the DDU-Net architecture. This is of paramount relevance for monitoring KS to enable physiological interpretation of MRI-based assessments of renal oxygenation in acute pathophysiological scenarios. Renal tissue hypoperfusion and hypoxia are pivotal early events in the pathophysiology of AKI, and in progression to CKD.^{9,81–86} The MR relaxation parameters T_2^* , T_2 are surrogates of renal tissue oxygenation, because they are sensitive to the amount of deoxygenated hemoglobin (deoxyHb) per tissue volume. However, the relationship between renal tissue partial pressure of oxygen (pO_2) and T_2^* , T_2 is confounded by changes in hematocrit, the O_2 affinity of hemoglobin, and crucially by the blood and tubular volume fractions.^{16,28,87} Recognizing that events leading to acute renal hypoxia are often associated with changes in the blood and/or tubular volume fractions, and that these changes are mirrored by changes in KS, our recent preclinical study used simultaneous MRI-based measurements of changes in KS and T_2^* , T_2 for accurate physiological interpretation of acute renal oxygenation changes.⁴⁴ If a decrease in T_2^* , T_2 is observed without any change in KS, this indicates that renal tissue oxygenation is reduced. However, if T_2^* , T_2 decrease and there is a corresponding decrease in KS, the reduction in tissue oxygenation should be regarded as more severe than if KS was unchanged. Conversely, if a similar decrease in T_2^* , T_2 is observed together with an increase in KS, the reduction in tissue oxygenation should be considered less severe than if KS remained unchanged. Thus, MRI-based measurement of renal oxygenation by T_2^* , T_2

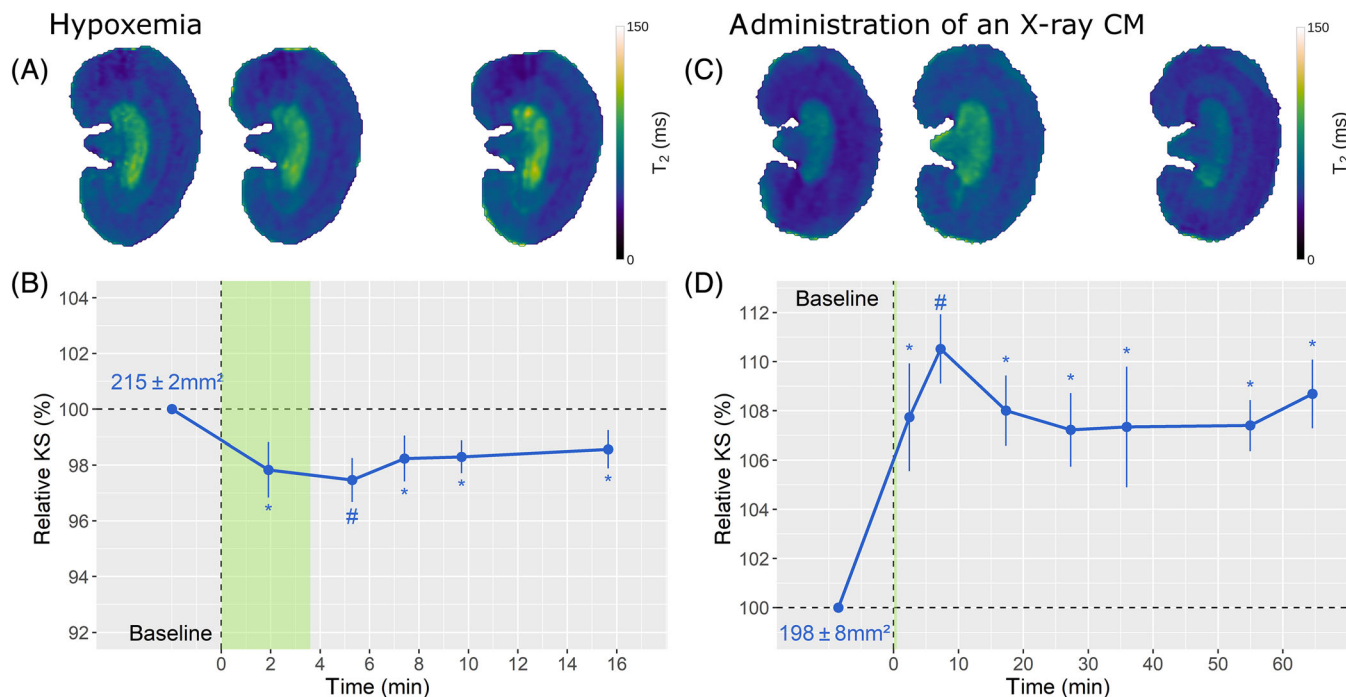


FIGURE 11 Relative kidney size (KS) over time during hypoxemia and upon administration of an X-ray contrast medium. (A) Exemplary T_2 maps are shown before (left) and during (center) hypoxemia, as well as during recovery (right). (B) Development of relative KS changes upon hypoxemia. The blue line depicts the mean of each time point and the corresponding interval of the standard error of the mean (SEM) ($n = 11$ rats). The mean KS at baseline was $215 \pm 2 \text{ mm}^2$. The period of hypoxemia is highlighted in green. (C) Exemplary T_2 maps are shown before (left) and upon (center and right) injection of an X-ray contrast medium. (D) Development of relative KS changes upon bolus injection of the X-ray contrast media (CM) ($n = 9$ rats). The mean KS at baseline was $198 \pm 8 \text{ mm}^2$. The time of the injection is highlighted in green. #: $p < 0.01$; *: $p < 0.05$.

is highly dependent on monitoring accompanying changes in KS. To enable correct interpretation of T_2^* , T_2 derived from parametric MRI of renal oxygenation, accurate KS assessment is especially relevant, because a biophysical model can then be used to estimate changes in O_2 saturation of hemoglobin from changes in T_2^* and KS.⁴⁴ The current study demonstrates that DDU-Net meets this need. As KS can be readily obtained from DDU-Net-based automated segmentation of T_2 maps without the need for additional MRI scans, this should always accompany assessment of MRI-derived oxygenation results.

The insights obtained from this work are not limited to experimental and preclinical studies. The basic principle of our approach can be adapted to the segmentation of human kidney MR images obtained from large-scale population MRI studies, like the German National Cohort,^{88–90} which include longitudinal MRI data with baseline and follow-up examinations. This wealth of data offers ample potential to train ML algorithms to establish real-time KS assessment in routine clinical practice. Leveraging MRI-based KS assessment from these large datasets would also support deciphering of the relationships between KS, molecular and biochemical markers, and data from physiological measurements, medical history and lifestyle data included in large cohort studies. These insights will advance our understanding of the determinants and associations of renal disease, and would promote extension of the US Food and Drug Administration (FDA) and European Medicines Agency (EMA) guidelines on KS assessment, which are currently restricted to the use of KS as a biomarker for autosomal-dominant PKD. An expanded perspective would advance nephrology, radiology, and physiology, involving clinicians and patients, and provide a foundation for new insights into renal physiology and pathology.

ACKNOWLEDGMENTS

The authors thank A. Pohlmann, J. Periquito (Max-Delbrueck Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany), A. Anger, B. Flemming (Institute of Translational Physiology, Charité – Universitätsmedizin, Berlin, Germany), and Fabian Isensee (Helmholtz Imaging, German Cancer Research Center [DKFZ], Heidelberg, Germany) for technical and other support. The authors further thank the iNAMES - MDC-Weizmann Helmholtz International Research School (HIRS) for Imaging and Data Science from the NANO to the MESO. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

ORCID

Tobias Klein  <https://orcid.org/0000-0003-2236-8490>

REFERENCES

- Hoste EAJ, Kellum JA, Selby NM, et al. Global epidemiology and outcomes of acute kidney injury. *Nat Rev Nephrol*. 2018;14(10):607-625. doi:10.1038/s41581-018-0052-0
- Luyckx VA, Tonelli M, Stanifer JW. The global burden of kidney disease and the sustainable development goals. *Bull World Health Organ*. 2018;96(6):414-422D. doi:10.2471/BLT.17.206441
- Levin A, Tonelli M, Bonventre J, et al. Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy. *Lancet*. 2017;390(10105):1888-1917. doi:10.1016/S0140-6736(17)30788-2
- Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. *Review Kidney Int Suppl*. 2022;12(1):7-11. doi:10.1016/j.kisu.2021.11.003
- GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2020;395(10225):709-733. doi:10.1016/S0140-6736(20)30045-3
- Sundström J, Bodegard J, Bollmann A, et al. Prevalence, outcomes, and cost of chronic kidney disease in a contemporary population of 2.4 million patients from 11 countries: the CaReMe CKD study. *Lancet Reg Health Eur*. 2022;20:100438. doi:10.1016/j.lanpe.2022.100438
- Kellum JA, Ronco C, Bellomo R. Conceptual advances and evolving terminology in acute kidney disease. *Nat Rev Nephrol*. 2021;17(7):493-502. doi:10.1038/s41581-021-00410-w
- Pickkers P, Ostermann M, Joannidis M, et al. The intensive care medicine agenda on acute kidney injury. *Intensive Care Med*. 2017;43(9):1198-1209. doi:10.1007/s00134-017-4687-2
- Fähling M, Seeliger E, Patzak A, Persson PB. Understanding and preventing contrast-induced acute kidney injury. *Nat Rev Nephrol*. 2017;13(3):169-180. doi:10.1038/nrneph.2016.196
- Selby NM, Taal MW. Long-term outcomes after AKI—a major unmet clinical need. *Kidney Int*. 2019;95(1):21-23. doi:10.1016/j.kint.2018.09.005
- Molitoris BA. Urinary biomarkers: alone are they enough? *J Am Soc Nephrol*. 2015;26(7):1485-1488. doi:10.1681/ASN.2014111145
- van Duijl TT, Soonawala D, de Fijter JW, Ruhaak LR, Cobbaert CM. Rational selection of a biomarker panel targeting unmet clinical needs in kidney injury. *Clin Proteomics*. 2021;18(1):10. doi:10.1186/s12014-021-09315-z
- Porrini E, Ruggenenti P, Luis-Lima S, et al. Estimated GFR: time for a critical appraisal. *Nat Rev Nephrol*. 2019;15(3):177-190. doi:10.1038/s41581-018-0080-9
- van Duijl TT, Ruhaak LR, de Fijter JW, Cobbaert CM. Kidney injury biomarkers in an academic hospital setting: where are we now? *Clin Biochem Rev*. 2019;40(2):79-97. doi:10.33176/AACB-18-00017
- Zhang WR, Parikh CR. Biomarkers of acute and chronic kidney disease. *Annu Rev Physiol*. 2019;81:309-333. doi:10.1146/annurev-physiol-020518-114605
- Niendorf T, Pohlmann A, Arakelyan K, et al. How bold is blood oxygenation level-dependent (BOLD) magnetic resonance imaging of the kidney? Opportunities, challenges and future directions. *Acta Physiol*. 2015;213(1):19-38. doi:10.1111/apha.12393
- Matejovic M, Ince C, Chawla LS, et al. Renal hemodynamics in AKI: in search of new treatment targets. *J Am Soc Nephrol*. 2016;27(1):49-58. doi:10.1681/ASN.2015030234
- Grenier N, Merville P, Combe C. Radiologic imaging of the renal parenchyma structure and function. *Nat Rev Nephrol*. 2016;12(6):348-359. doi:10.1038/nrneph.2016.44
- Selby NM, Blankestijn PJ, Boor P, et al. Magnetic resonance imaging biomarkers for chronic kidney disease: a position paper from the European Cooperation in Science and Technology Action PARENCHIMA. *Nephrol Dial Transplant*. 2018;33(suppl_2):ii4-ii14. doi:10.1093/ndt/gfy152
- Caroli A, Remuzzi A, Remuzzi G. Does MRI trump pathology? A new era for staging and monitoring of kidney fibrosis. *Kidney Int*. 2020;97(3):442-444. doi:10.1016/j.kint.2019.10.008
- Simms R, Sourbron S. Recent findings on the clinical utility of renal magnetic resonance imaging biomarkers. *Nephrol Dial Transplant*. 2020;35(6):915-919. doi:10.1093/ndt/gfaa125
- Puelles VG, Combes AN, Bertram JF. Clearly imaging and quantifying the kidney in 3D. *Kidney Int*. 2021;100(4):780-786. doi:10.1016/j.kint.2021.04.042
- Charlton JR, Xu Y, Wu T, et al. Magnetic resonance imaging accurately tracks kidney pathology and heterogeneity in the transition from acute kidney injury to chronic kidney disease. *Kidney Int*. 2021;99(1):173-185. doi:10.1016/j.kint.2020.08.021
- Selby NM, Duranteau J. New imaging techniques in AKI. *Curr Opin Crit Care*. 2020;26(6):543-548. doi:10.1097/MCC.0000000000000768
- Caroli A, Remuzzi A, Lerman LO. Basic principles and new advances in kidney imaging. *Kidney Int*. 2021;100(5):1001-1011. doi:10.1016/j.kint.2021.04.032
- Bane O, Mendichovszky IA, Milani B, et al. Consensus-based technical recommendations for clinical translation of renal BOLD MRI. *MAGMA*. 2020;33(1):199-215. doi:10.1007/s10334-019-00802-x
- Boehmert L, Kuehne A, Waiczies H, et al. Cardioresonant sodium MRI at 7.0 Tesla using a 4/4 channel (1) H/(23) Na radiofrequency antenna array. *Magn Reson Med*. 2019;82(6):2343-2356. doi:10.1002/mrm.27880
- Li LP, Hack B, Seeliger E, Prasad PV. MRI mapping of the blood oxygenation sensitive parameter T2* in the kidney: basic concept. *Methods Mol Biol*. 2021;2216:171-185. doi:10.1007/978-1-0716-0978-1_10
- Pohlmann A, Back SJ, Fekete A, et al. Recommendations for preclinical renal MRI: a comprehensive open-access protocol collection to improve training, reproducibility, and comparability of studies. *Methods Mol Biol*. 2021;2216:3-23. doi:10.1007/978-1-0716-0978-1_1
- Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron Clin Pract*. 2012;120(4):c179-c184. doi:10.1159/000339789

31. Zarjou A, Sanders PW, Mehta RL, Agarwal A. Enabling innovative translational research in acute kidney injury. *Clin Transl Sci.* 2012;5(1):93-101. doi:10.1111/j.1752-8062.2011.00302.x
32. Evans RG, O'Connor PM. Initiation and progression of chronic kidney disease: can we definitively test the chronic hypoxia hypothesis? *Hypertension.* 2013;62(5):827-828. doi:10.1161/HYPERTENSIONAHA.113.02047
33. Cantow K, Hummel L, Flemming B, Waiczies S, Niendorf T, Seeliger E. Imagine physiology without imaging. *Acta Physiol.* 2020;230(3):e13549. doi:10.1111/apha.13549
34. Niendorf T, Frydman L, Neeman M, Seeliger E. Google maps for tissues: multiscale imaging of biological systems and disease. *Acta Physiol.* 2020;228(2):e13392. doi:10.1111/apha.13392
35. Periquito JS, Gladysz T, Millward JM, et al. Continuous diffusion spectrum computation for diffusion-weighted magnetic resonance imaging of the kidney tubule system. *Quant Imaging Med Surg.* 2021;11(7):3098-3119. doi:10.21037/qims-20-1360
36. Grantham JJ, Torres VE. The importance of total kidney volume in evaluating progression of polycystic kidney disease. *Nat Rev Nephrol.* 2016;12(11):667-677. doi:10.1038/nrneph.2016.135
37. U.S. Food and Drug Administration. Qualification of biomarker total kidney volume in studies for treatment of autosomal dominant polycystic kidney disease draft guidance for industry. 2016. www.fda.gov/media/93105/download. Accessed: 18.11.2023.
38. European Medicine Agency. Qualification opinion: total kidney volume (TKV) as a prognostic biomarker for use in clinical trials evaluating patients with autosomal dominant polycystic kidney disease. 2015. www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-total-kidney-volume-tkv-prognostic-biomarker-use-clinical-trials-evaluating_en.pdf. Accessed: 18.11.2023.
39. Buchanan CE, Mahmoud H, Cox EF, et al. Quantitative assessment of renal structural and functional changes in chronic kidney disease using multiparametric magnetic resonance imaging. *Nephrol Dial Transplant.* 2020;35(6):955-964. doi:10.1093/ndt/gfz129
40. Gladysz T, Millward JM, Cantow K, et al. Reliable kidney size determination by magnetic resonance imaging in pathophysiological settings. *Acta Physiol.* 2021;233(2):e13701. doi:10.1111/apha.13701
41. Bak M, Thomsen K, Christiansen T, Flyvbjerg A. Renal enlargement precedes renal hyperfiltration in early experimental diabetes in rats. *J Am Soc Nephrol.* 2000;11(7):1287-1292. doi:10.1681/ASN.V1171287
42. Erokwu BO, Anderson CE, Flask CA, Dell KM. Quantitative magnetic resonance imaging assessments of autosomal recessive polycystic kidney disease progression and response to therapy in an animal model. *Pediatr Res.* 2018;83(5):1067-1074. doi:10.1038/pr.2018.24
43. Hueper K, Hensen B, Gutberlet M, et al. Kidney transplantation: multiparametric functional magnetic resonance imaging for assessment of renal allograft pathophysiology in mice. *Invest Radiol.* 2016;51(1):58-65. doi:10.1097/RLI.0000000000000205
44. Cantow K, Gladysz T, Millward JM, Waiczies S, Niendorf T, Seeliger E. Monitoring kidney size to interpret MRI-based assessment of renal oxygenation in acute pathophysiological scenarios. *Acta Physiol.* 2023;237(2):e13868. doi:10.1111/apha.13868
45. Bakker J, Olree M, Kaatee R, et al. Renal volume measurements: accuracy and repeatability of US compared with that of MR imaging. *Radiology.* 1999;211(3):623-628. doi:10.1148/radiology.211.3.r99jn19623
46. Seuss H, Janka R, Prümmer M, et al. Development and evaluation of a semi-automated segmentation tool and a modified ellipsoid formula for volumetric analysis of the kidney in non-contrast T2-weighted MR images. *J Digit Imaging.* 2017;30(2):244-254. doi:10.1007/s10278-016-9936-3
47. Irazabal MV, Rangel LJ, Bergstralh EJ, et al. CRISP Investigators Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials. *J Am Soc Nephrol.* 2015;26(1):160-172. doi:10.1681/ASN.2013101138
48. Sharma K, Caroli A, Quach LV, et al. Kidney volume measurement methods for clinical studies on autosomal dominant polycystic kidney disease. *PLoS ONE.* 2017;12(5):e0178488. doi:10.1371/journal.pone.0178488
49. Bae KT, Tao C, Wang J, et al. Novel approach to estimate kidney and cyst volumes using mid-slice magnetic resonance images in polycystic kidney disease. *Am J Nephrol.* 2013;38(4):333-341. doi:10.1159/000355375
50. Simms RJ, Doshi T, Metherall P, et al. A rapid high-performance semi-automated tool to measure total kidney volume from MRI in autosomal dominant polycystic kidney disease. *Eur Radiol.* 2019;29(8):4188-4197. doi:10.1007/s00330-018-5918-9
51. Bouteldja N, Klinkhammer BM, Bülow RD, et al. Deep learning-based segmentation and quantification in experimental kidney histopathology. *J Am Soc Nephrol.* 2021;32(1):52-68. doi:10.1681/ASN.2020050597
52. Kers J, Bülow RD, Klinkhammer BM, et al. Deep learning-based classification of kidney transplant pathology: a retrospective, multicentre, proof-of-concept study. *Lancet Digit Health.* 2022;4(1):e18-e26. doi:10.1016/S2589-7500(21)00211-9
53. Gloger O, Tonnie K, Laqua R, Volzke H. Fully automated renal tissue volumetry in MR volume data using prior-shape-based segmentation in subject-specific probability maps. *IEEE Trans Biomed Eng.* 2015;62(10):2338-2351. doi:10.1109/TBME.2015.2425935
54. van Gastel MDA, Edwards ME, Torres VE, Erickson BJ, Gansevoort RT, Kline TL. Automatic measurement of kidney and liver volumes from MR images of patients affected by autosomal dominant polycystic kidney disease. *J Am Soc Nephrol.* 2019;30(8):1514-1522. doi:10.1681/ASN.2018090902
55. Gloger O, Tonnie K, Mensel B, Volzke H. Fully automatized renal parenchyma volumetry using a support vector machine based recognition system for subject-specific probability map generation in native MR volume data. *Phys Med Biol.* 2015;60(22):8675-8693. doi:10.1088/0031-9155/60/22/8675
56. Edwards ME, Periyana S, Anaam D, Gregory AV, Kline TL. Automated total kidney volume measurements in pre-clinical magnetic resonance imaging for resourcing imaging data, annotations, and source code. *Kidney Int.* 2021;99(3):763-766. doi:10.1016/j.kint.2020.07.040
57. Kline TL, Korfiatis P, Edwards ME, et al. Performance of an artificial multi-observer deep neural network for fully automated segmentation of polycystic kidneys. *J Digit Imaging.* 2017;30(4):442-448. doi:10.1007/s10278-017-9978-1
58. Sharma K, Rupprecht C, Caroli A, et al. Automatic segmentation of kidneys using deep learning for total kidney volume quantification in autosomal dominant polycystic kidney disease. *Sci Rep.* 2017;7(1):2049. doi:10.1038/s41598-017-01779-0
59. Langner T, Östling A, Maldonis L, et al. Kidney segmentation in neck-to-knee body MRI of 40,000 UK Biobank participants. *Sci Rep.* 2020;10(1):20963. doi:10.1038/s41598-020-77981-4
60. Daniel AJ, Buchanan CE, Allcock T, et al. Automated renal segmentation in healthy and chronic kidney disease subjects using a convolutional neural network. *Magn Reson Med.* 2021;86(2):1125-1136. doi:10.1002/mrm.28768
61. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Presented at: MICCAI; 2015; <https://arxiv.org/abs/1505.04597>

62. Piao S, Liu J. Accuracy improvement of UNet based on dilated convolution. *J Phys Conf Ser.* 2019;1345(5):052066. doi:10.1088/1742-6596/1345/5/052066
63. Zhao W, Jiang D, Peña Queraltó J, Westerlund T. MSS U-Net: 3D segmentation of kidneys and tumors from CT images with a multi-scale supervised U-Net. *IMU.* 2020;19:100357. doi:10.1016/j.imu.2020.100357
64. Bones IK, Bos C, Moonen C, Hendrikse J, van Stralen M. Workflow for automatic renal perfusion quantification using ASL-MRI and machine learning. *Magn Reson Med.* 2022;87(2):800-809. doi:10.1002/mrm.29016
65. Wang S, Singh VK, Cheah E, et al. Stacked dilated convolutions and asymmetric architecture for U-Net-based medical image segmentation. *Comput Biol Med.* 2022;148:105891. doi:10.1016/j.combiomed.2022.105891
66. Sharbatdaran A, Romano D, Teichman K, et al. Deep learning automation of kidney, liver, and spleen segmentation for organ volume measurements in autosomal dominant polycystic kidney disease. *Tomography.* 2022;8(4):1804-1819. doi:10.3390/tomography8040152
67. Kim Y, Tao C, Kim H, Oh GY, Ko J, Bae KT. A deep learning approach for automated segmentation of kidneys and exophytic cysts in individuals with autosomal dominant polycystic kidney disease. *J Am Soc Nephrol.* 2022;33(8):1581-1589. doi:10.1681/ASN.2021111400
68. 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.* 2022;146:105431. doi:10.1016/j.combiomed.2022.105431
69. Raj A, Tollens F, Hansen L, et al. Deep learning-based total kidney volume segmentation in autosomal dominant polycystic kidney disease using attention, cosine loss, and sharpness aware minimization. *Diagnostics.* 2022;12(5), 1159. doi:10.3390/diagnostics12051159
70. Edwards ME, Blais JD, Czerwicz FS, Erickson BJ, Torres VE, Kline TL. Standardizing total kidney volume measurements for clinical trials of autosomal dominant polycystic kidney disease. *Clin Kidney J.* 2019;12(1):71-77. doi:10.1093/ckj/sfy078
71. Cantow K, Ladwig-Wiegand M, Flemming B, Fekete A, Hosszu A, Seeliger E. Reversible (patho)physiologically relevant test interventions: rationale and examples. *Methods Mol Biol.* 2021;2216:57-73. doi:10.1007/978-1-0716-0978-1_4
72. Cantow K, Ladwig-Wiegand M, Flemming B, Pohlmann A, Niendorf T, Seeliger E. Monitoring renal hemodynamics and oxygenation by invasive probes: experimental protocol. *Methods Mol Biol.* 2021;2216:327-347. doi:10.1007/978-1-0716-0978-1_19
73. Pohlmann A, Cantow K, Huelnhagen T, et al. Experimental MRI monitoring of renal blood volume fraction variations en route to renal magnetic resonance oximetry. *Tomography.* 2017;3(4):188-200. doi:10.18383/j.tom.2017.00012
74. Cantow K, Pohlmann A, Flemming B, et al. Acute effects of ferumoxytol on regulation of renal hemodynamics and oxygenation. *Sci Rep.* 2016;6:29965. doi:10.1038/srep29965
75. Cantow K, Evans RG, Grosenick D, et al. Quantitative assessment of renal perfusion and oxygenation by invasive probes: basic concepts. *Methods Mol Biol.* 2021;2216:89-107. doi:10.1007/978-1-0716-0978-1_6
76. Kavur AE, Gezer NS, Barış M, et al. CHAOS challenge—combined (CT-MR) healthy abdominal organ segmentation. *Med Image Anal.* 2021;69:101950. doi:10.1016/j.media.2020.101950
77. Izmailov P, Podoprikin D, Garipov T, Vetrov D, Wilson AG. Averaging weights leads to wider optima and better generalization. Presented at: UAI; 2018; <https://arxiv.org/abs/1803.05407>
78. Riazly L, Milani B, Periquito JS, et al. Subsegmentation of the kidney in experimental MR images using morphology-based regions-of-interest or multiple-layer concentric objects. *Methods Mol Biol.* 2021;2216:549-564. doi:10.1007/978-1-0716-0978-1_33
79. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods.* 2021;18(2):203-211. doi:10.1038/s41592-020-01008-z
80. Isensee F, Ulrich C, Wald T, Maier-Hein KH. Extending nnU-Net is all you need. In: Deserno TM, Handels H, Maier A, Maier-Hein K, Palm C, Tolxdorff T. *Bildverarbeitung für die Medizin 2023. BVM 2023. Informatik aktuell.* Springer Vieweg; 2023:12-17. doi:10.1007/978-3-658-41657-7_7
81. Brezis M, Rosen S. Hypoxia of the renal medulla—its implications for disease. *N Engl J Med.* 1995;332(10):647-655. doi:10.1056/NEJM199503093321006
82. Calzavacca P, Evans RG, Bailey M, Bellomo R, May CN. Cortical and medullary tissue perfusion and oxygenation in experimental septic acute kidney injury. *Crit Care Med.* 2015;43(10):e431-e439. doi:10.1097/CCM.0000000000001198
83. Evans RG, Ince C, Joles JA, et al. Haemodynamic influences on kidney oxygenation: clinical implications of integrative physiology. *Clin Exp Pharmacol Physiol.* 2013;40(2):106-122. doi:10.1111/1440-1681.12031
84. Evans RG, Ow CP, Bie P. The chronic hypoxia hypothesis: the search for the smoking gun goes on. *Am J Physiol Renal Physiol.* 2015;308(2):F101-F102. doi:10.1152/ajprenal.00587.2014
85. Ma S, Evans RG, Iguchi N, et al. Sepsis-induced acute kidney injury: a disease of the microcirculation. *Microcirculation.* 2019;26(2):e12483. doi:10.1111/micc.12483
86. Zuk A, Bonventre JV. Recent advances in acute kidney injury and its consequences and impact on chronic kidney disease. *Curr Opin Nephrol Hypertens.* 2019;28(4):397-405. doi:10.1097/MNH.0000000000000504
87. Niendorf T, Seeliger E, Cantow K, Flemming B, Waiczies S, Pohlmann A. Probing renal blood volume with magnetic resonance imaging. *Acta Physiol.* 2020;228(4):e13435. doi:10.1111/apha.13435
88. Bamberg F, Kauczor HU, Weckbach S, et al. Whole-body MR imaging in the German National Cohort: rationale, design, and technical background. *Radiology.* 2015;277(1):206-220. doi:10.1148/radiol.2015142272
89. Schuppert C, Krüchten R, Hirsch JG, et al. Whole-body magnetic resonance imaging in the large population-based German National Cohort study: predictive capability of automated image quality assessment for protocol repetitions. *Invest Radiol.* 2022;57(7):478-487. doi:10.1097/RLI.0000000000000861
90. Peters A, Greiser KH, Göttlicher S et al. Framework and baseline examination of the German National Cohort (NAKO). *Eur J Epidemiol.* 2022;37(10):1107-1124. doi:10.1007/s10654-022-00890-5

How to cite this article: Klein T, Gladysz T, Millward JM, et al. Dynamic parametric MRI and deep learning: Unveiling renal pathophysiology through accurate kidney size quantification. *NMR in Biomedicine.* 2024;37(4):e5075. doi:10.1002/nbm.5075