MRI Contrast Enhancement Kinetics World Model

Supplementary Material

A. The Derivation of Dense Smooth Constraint

We regularize the second-order temporal difference of the output sequence to suppress sudden jumps without excessively weakening real contrast changes. The first-order difference directly penalizes the magnitude of changes, making the model tend to generate nearly static sequences. In contrast, the second-order difference describes the rate of change of velocity, allowing for monotonic enhancement/attenuation when such changes are smooth. Since the sampling time points in this task are non-uniform, we further derive the corresponding three-point central difference formula on a non-uniform time grid, as shown below.

Let $\{\tau_{k-1}, \tau_k, \tau_{k+1}\}$ be three consecutive time stamps and denote adjacent gaps by $h_0^k = \tau_k - \tau_{k-1}$ and $h_1^k = \tau_{k+1} - \tau_k$. Expanding $y(\tau_{k-1})$ and $y(\tau_{k+1})$ at τ_k and keeping the remainder terms,

$$y(\tau_{k-1}) = y(\tau_k) - h_0^k y'(\tau_k) + \frac{(h_0^k)^2}{2} y''(\tau_k) - \frac{(h_0^k)^3}{6} y'''(\xi_0),$$
$$y(\tau_{k+1}) = y(\tau_k) + h_1^k y'(\tau_k) + \frac{(h_1^k)^2}{2} y''(\tau_k) + \frac{(h_1^k)^3}{6} y'''(\xi_1),$$

for some $\xi_0 \in (\tau_{k-1}, \tau_k)$ and $\xi_1 \in (\tau_k, \tau_{k+1})$. Multiplying the first expansion by h_1^k and the second by h_0^k cancels $y'(\tau_k)$; neglecting $O\big((h_0^k)^3 + (h_1^k)^3\big)$ then yields the nonuniform three-point central difference

$$\begin{aligned} y''(\tau_k) \approx{}& \frac{2}{h_0^k h_1^k (h_0^k + h_1^k)} \\ & \cdot \Big[h_1^k \big(y(\tau_k) - y(\tau_{k-1})\big) + h_0^k \big(y(\tau_{k+1}) - y(\tau_k)\big) \Big] \\ ={}& 2 \left(\frac{y(\tau_{k-1})}{h_0^k (h_0^k + h_1^k)} \right. \\ & \left. - \frac{y(\tau_k)}{h_0^k h_1^k} + \frac{y(\tau_{k+1})}{h_1^k (h_0^k + h_1^k)} \right). \end{aligned}$$

Notation by setting $y(\tau_j) = y_{\text{sort}}^j$ and adding a small stabilization δ to avoid zero denominators,

$$h_0^k = \tau_k - \tau_{k-1} + \delta, \qquad h_1^k = \tau_{k+1} - \tau_k + \delta, \quad (1)$$

we weight each stencil by $w^k = \frac{1}{1 + h_0^k + h_1^k}$ to reduce the penalty over large temporal gaps. The dense smoothness (center-difference) operator used in the loss at index k is therefore

$$\mathbf{D}_2^k = 2 \left(\frac{y_{\text{sort}}^{k-1}}{h_0^k (h_0^k + h_1^k)} - \frac{y_{\text{sort}}^{k}}{h_0^k h_1^k} + \frac{y_{\text{sort}}^{k+1}}{h_1^k (h_0^k + h_1^k)} \right) w^k,$$

which is a second-order accurate approximation of $y''(\tau_k)$ on a nonuniform grid and reduces to the standard centered stencil when $h_0^k = h_1^k$.

B. The Details of Experiments

B.1. Datasets

Abdominal DCE-MRI It is acquired via the CAIPIRINHA-Dixon-TWIST-VIBE technique[4], which fully covers multiple time phases before and after contrast agent injection to reflect tissues' hemodynamic characteristics accurately. Specifically, it comprises 1 pre-contrast phase before contrast agent injection, approximately 6 random arterial phases between 15 and 37 seconds post contrast agent injection for capturing the rapid filling process of contrast agent in the arterial system, approximately 6 random portal vein phases between 50 and 72 seconds post contrast agent injection for reflecting the distribution and perfusion of contrast agent into the portal vein system, and several delayed phases which are acquired nearby 90 s, 150 s, or 300 s post contrast agent injection, respectively, to observe the delayed enhancement or clearance of contrast agent in tissues. Since the raw data were not registered, all enhanced sequences in the dataset were registered to the non-contrast images. **Breast DCE-MRI** This dataset involves 2 manufacturers. Among them, equipment from GE Medical Systems accounts for a higher proportion, equipped with two magnetic field strengths: 1.5T and 3T. Equipment from Siemens mainly includes models such as Avanto, Skyra, and TrioTim, which also cover both 1.5T and 3T magnetic field strengths. A median of 131s passed between DCE sequences [9].

B.2. Visualization Results

B.2.1. Visualization of Different Tissues in Abdominal DCE-MRI

Different Contrast-enhanced results for the spleen, kidney, and liver demonstrate that CEKWorld is able to simulate different contrast enhancement kinetics laws, which shows the potential for accurate differentiation of normal tissues and lesions across multiple organs, supporting the improvement of diagnostic efficiency and precision in abdominal imaging.

Spleen As shown in Fig.1, CEKWorld conforms to the contrast agent kinetics of the spleen at all time points [3, 7, 11, 12]. In the early arterial phase, when t=6s and 9s, CEKWorld initiates enhancement rapidly and synchronously with the ground-truth, showing typical patchy heterogeneous hyperintensity. In contrast, ControlNet has an obvious delay with slow initiation of enhancement, and methods such as CCNet exhibit abnormal enhancement intensity. In the middle arterial phase, when t=12s and 15s, the enhancement of CEKWorld gradually transitions from

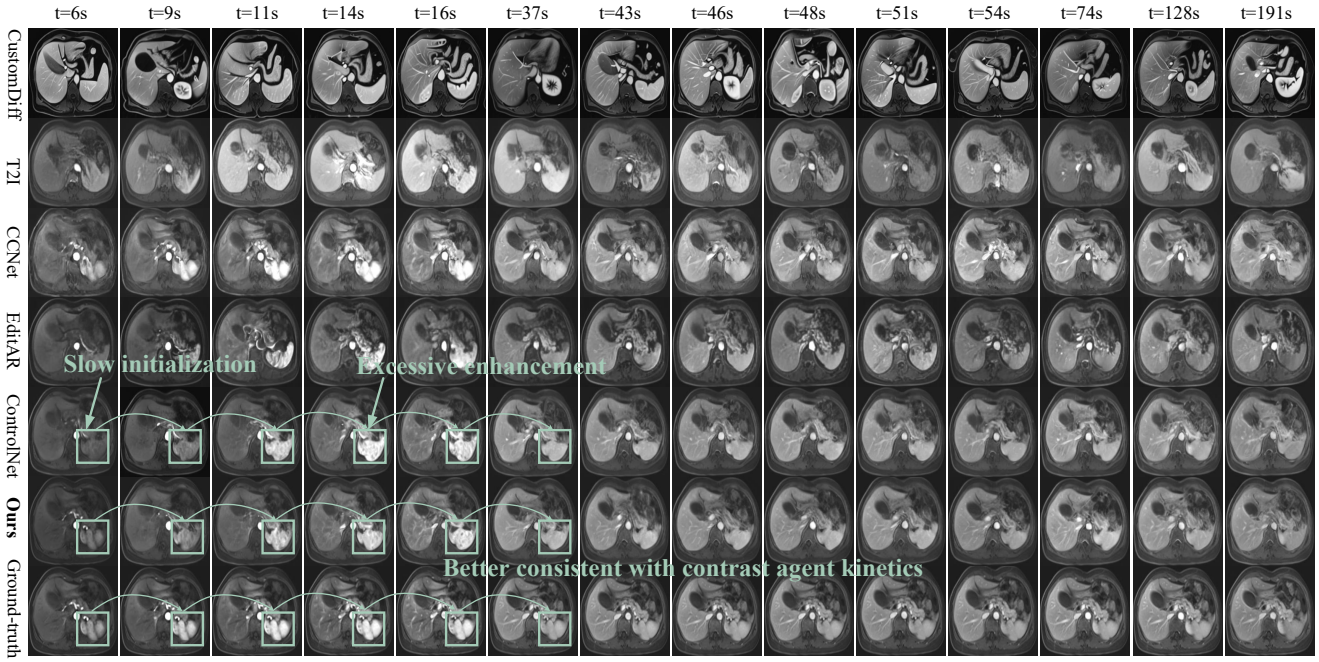

Figure 1. **Visualization results of spleen** in Abdominal DCE-MRI shows that our CEKWorld conforms to the spleen contrast agent kinetics, exhibiting the characteristics of rapid synchronous enhancement, progressive homogenization, and sustained homogeneous high signal. In contrast, methods such as ControlNet and CCNet have defects, including delayed start of enhancement, abnormal intensity, or content distortion.

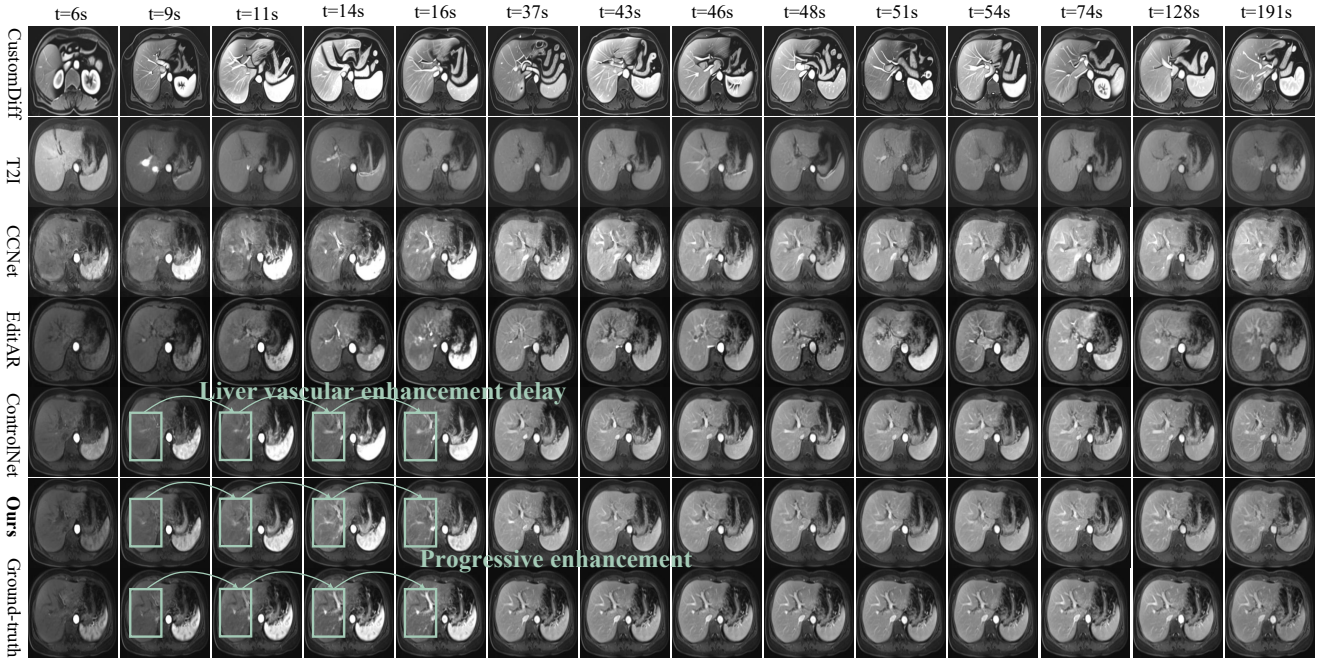

Figure 2. **Visualization results of liver** in Abdominal DCE-MRI shows that our CEKWorld conforms to the contrast agent kinetic laws of hepatic vessel progressive enhancement and liver parenchyma's regular enhancement-washout. In contrast, other methods exhibit distorted contents, while ControlNet suffers from non-physiological hepatic vessel enhancement delay, confirming our spatiotemporal consistency learning accurately captures liver contrast agent kinetics.

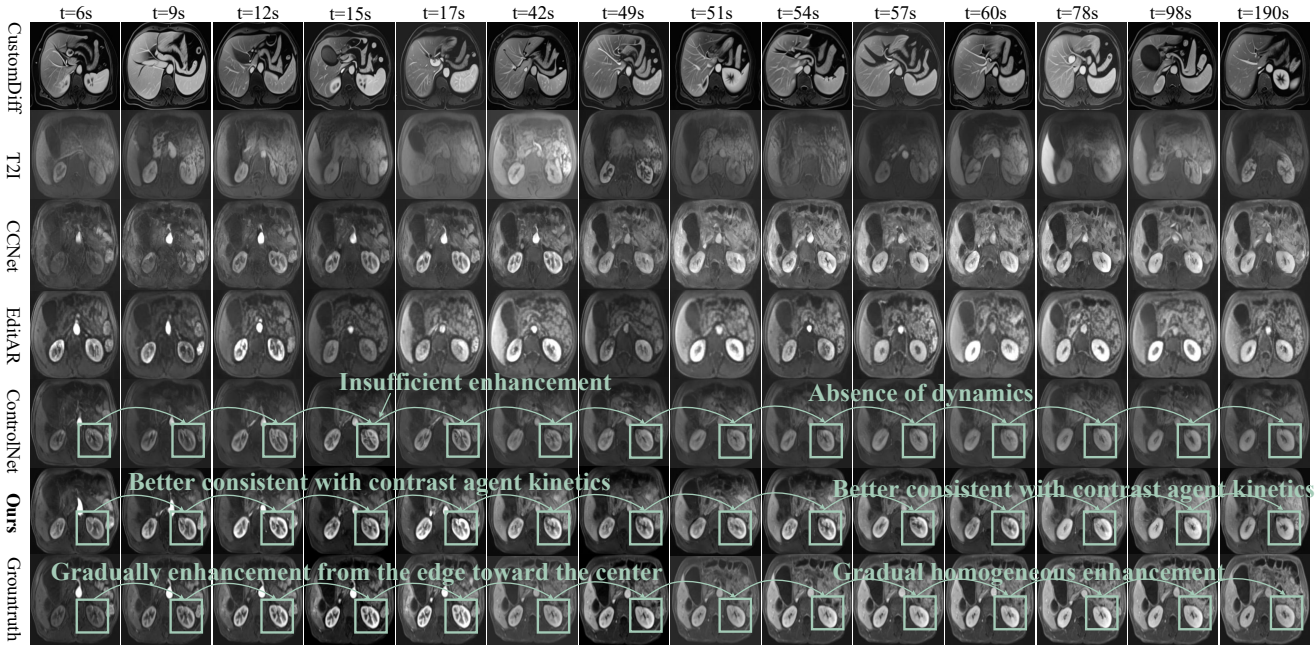

Figure 3. **Visualization results of kidney** in Abdominal DCE-MRI shows that our CEKWorld reproduces the typical renal perfusion pattern of cortical edge enhancement first, gradual diffusion to the medulla until homogeneity, while methods such as T2I and CCNet have problems of abnormal enhancement or blurred structure, and ControlNet lacks dynamics in the delay phase.

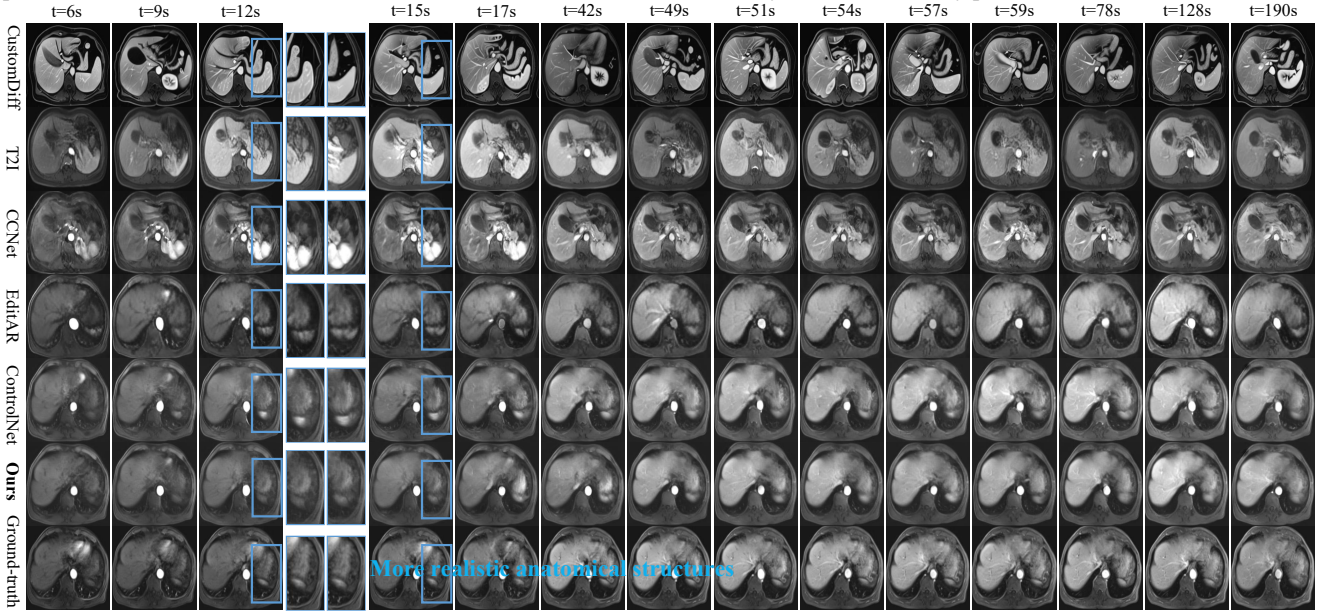

Figure 4. **Visualization results** in Abdominal DCE-MRI: The visualization results of our methods on different time points exhibit better spatial reality (zoom-in regions in blue boxes).

patchy to homogeneous, which is completely consistent with the physiological process of contrast agent gradual diffusion in the splenic red pulp sinuses. In contrast, other methods either have distorted enhancement morphology or disordered rhythm. In the late arterial phase and vein phase, when t=17s and 42s, CEKWorld continuously maintains homogeneous hyperintensity without abnormal washout, perfectly matching the imaging feature of the spleen that enhancement lasts long and becomes homogeneous in the later stage. However, other methods have already shown signal distortion or abrupt transition.

Liver In Fig. 2, we visualize the predicted dynamic contrast enhancement of the liver across multiple time points. CustomDiff, T2I, CCNet and EditAR all struggle to repro-

duce realistic liver kinetics [1, 10]: vascular structures either fail to enhance at the appropriate phases or exhibit unstable parenchymal signals, with noticeable frame-to-frame fluctuations and over-smoothed textures. ControlNet shows better anatomical fidelity but presents a clear liver vascular enhancement delay—hepatic vessels remain underenhanced in the early arterial and early portal phases and then suddenly become hyperintense several frames later, resulting in a phase-shifted and non-physiological enhancement pattern. In contrast, our method closely follows the ground truth: the hepatic vessels in the green ROI show progressive enhancement, with smooth, timely transition from arterial to portal venous and delayed phases, while the liver parenchyma brightens and washes out in a gradual, temporally coherent manner. This demonstrates that our spatiotemporal consistency learning not only preserves realistic liver anatomy but also captures the correct contrast-agent kinetics over time.

Kidney In Fig.3, the contrast agent kinetics in the renal cortex–medulla region highlight that our CEKWorld shows more similarities with the ground-truth sequence. In the ground-truth sequence, enhancement first appears along the cortical rim on the edge, gradually propagates toward the medulla in the center, and finally becomes spatially homogeneous at late phases, exhibiting a typical outside-in renal perfusion pattern [2, 8]. T2I and CCNet either under-enhance or severely corrupt textures, making the cortical–medullary layers indistinguishable; EditAR shows weak and unstable enhancement; and ControlNet suffers from a lack of smooth temporal evolution. In contrast, our CEKWorld reproduces early cortical enhancement, its gradual inward spread, and the final homogeneous enhancement in a structurally faithful and temporally continuous manner, yielding renal contrast dynamics that best match real DCE-MRI.

The virtual angiography visualization results of these tissues conform to the rules of contrast agents, which fully demonstrate that CEKWorld is able to simulate the physiological mechanisms of tissues, the reliability to meet clinical diagnostic needs, as well as the stability and universality across tissue scenarios, and thus has significant clinical application potential.

B.2.2. Visualization of Spatial Structures in Abdominal DCE-MRI

In Fig.4, our methods exhibit better anatomical fidelity. CCNet introduces strong, grainy noise and artifacts in the spleen and liver parenchyma, tearing and blurring organ boundaries, and severely distorting the underlying anatomy. EditAR preserves a roughly correct global outline, but local details fluctuate strongly between neighboring frames: the splenic hilum and adjacent parenchyma alternately swell and collapse, yielding unstable textures. ControlNet largely retains coarse organ shape, yet its internal structures gradually blur, and vessel–parenchyma interfaces drift slightly, indicating mild geometric shift and loss of structural sharpness. In contrast, our MRI CEKWorld consistently preserves spleen contours, liver parenchymal shape, and vascular trajectories that closely match the ground truth. The blue-highlighted regions reveal that our results are the closest to the reference in both organ geometry and texture distribution, demonstrating the effectiveness of our spatial regularization in enforcing patient-specific anatomical consistency.

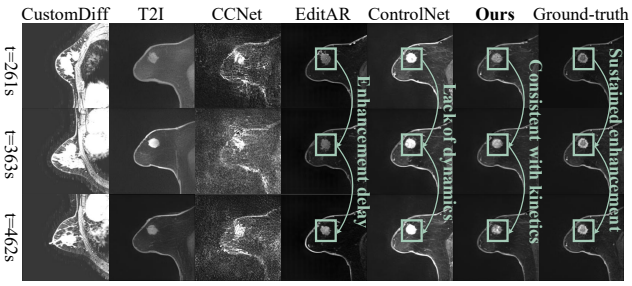

Figure 5. **Visualization results** in Breast DCE-MRI show that our method exhibits an enhancement pattern similar to the ground truth, characterized by the sustained enhancement, which is often associated with benign or low-risk lesions.

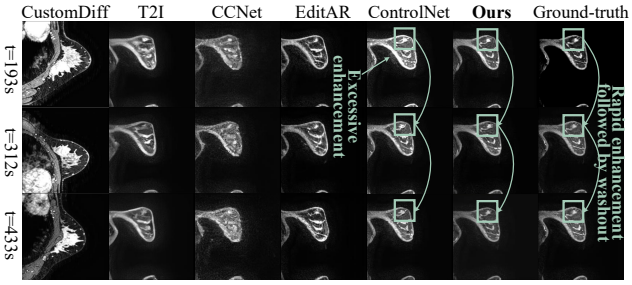

Figure 6. **Visualization results** in Breast DCE-MRI demonstrate that our method captures the dynamic pattern of rapid initial enhancement followed by clear washout, which is consistent with the ground truth and indicative of malignant or high-risk lesions.

B.2.3. Visualization Results for Different Enhancement Patterns in Breast DCE-MRI

The two breast DCE-MRI visualization cases shown in Fig.5 and Fig.6 demonstrate that our CEKWorld simulates clinically meaningful kinetic patterns, showing strong potential for downstream risk-stratification tasks since lesion malignancy is closely associated with characteristic enhancement patterns [5, 6]. In both figures, our model accurately reproduces the temporal trajectories observed in ground-truth DCE-MRI sequences, whereas competing methods exhibit severe issues such as incorrect enhancement magnitude, temporal shifts, or lack of temporal evolution. In Fig.5, the ground-truth exhibits a gradually rising and persistently elevated curve, typical of a persistent

enhancement pattern, often associated with benign or low-risk lesions. Our method produces consistent, sustained enhancement across all time points, faithfully matching the true kinetics. In contrast, EditAR shows enhancement delay, and ControlNet lacks any meaningful temporal progression, failing to reproduce the low-risk kinetic profile. This demonstrates our model's ability to reconstruct benign-like temporal behavior. In Fig.6, the lesion shows rapid initial enhancement followed by clear washout, a hallmark of malignant or high-risk lesions. Our method correctly captures this dynamic pattern: a sharp early rise followed by a decline consistent with the ground truth. EditAR suffers from excessive enhancement and fails to show the washout phase, while ControlNet exhibits collapsed or weakened dynamics, failing to reproduce the malignant-typical kinetic transition. Together, these two visualization examples show that our method not only preserves spatial fidelity but also faithfully models critical DCE-MRI temporal dynamics. Because low- and high-risk lesions exhibit fundamentally different enhancement trajectories, the ability of our model to reconstruct these dynamics suggests strong potential for downstream tasks such as risk stratification, benign-vs-malignant discrimination, and clinical subtype prediction.

References

[1] Nils Dahlström. *Quantitative evaluation of contrast agent dynamics in liver MRI*. Linkopings Universitet (Sweden), 2010. 4

[2] Eli Eikefjord, Erling Andersen, Erlend Hodneland, Frank Zöllner, Arvid Lundervold, Einar Svarstad, and Jarle Rørvik. Use of 3d dce-mri for the estimation of renal perfusion and glomerular filtration rate: an intrasubject comparison of flash and kwic with a comprehensive framework for evaluation. *American Journal of Roentgenology*, 204(3):W273–W281, 2015. 4

[3] Khaled M Elsayes, Vamsidhar R Narra, Govind Mukundan, James S Lewis Jr, Christine O Menias, and Jay P Heiken. Mr imaging of the spleen: spectrum of abnormalities. *Radiographics*, 25(4):967–982, 2005. 1

[4] Wen Hao, Weijun Peng, Cuiyan Wang, Bin Zhao, and Guangbin Wang. Image quality of the caipirinha-dixon-twist-vibe technique for ultra-fast breast dce-mri: Comparison with the conventional gre technique. *European Journal of Radiology*, 129:109108, 2020. 1

[5] Christiane Katharina Kuhl, Peter Mielcareck, Sven Klaschik, Claudia Leutner, Eva Wardelmann, Jurgen Gieseke, and Hans H Schild. Dynamic breast mr imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology*, 211(1):101–110, 1999. 4

[6] Katarzyna J Macura, Ronald Ouwerkerk, Michael A Jacobs, and David A Bluemke. Patterns of enhancement on breast mr images: interpretation and imaging pitfalls. *Radiographics*, 26(6):1719–1734, 2006. 4

[7] Scott A Mirowitz, Jeffrey J Brown, JK Lee, and Jay P Heiken. Dynamic gadolinium-enhanced mr imaging of the spleen: normal enhancement patterns and evaluation of splenic lesions. *Radiology*, 179(3):681–686, 1991. 1

[8] Michael Pedersen, Pietro Irrera, Walter Dastrù, Frank G Zöllner, Kevin M Bennett, Scott C Beeman, G Larry Bretthorst, Joel R Garbow, and Dario Livio Longo. Dynamic contrast enhancement (dce) mri–derived renal perfusion and filtration: Basic concepts. In *Preclinical MRI of the Kidney: Methods and Protocols*, pages 205–227. Springer US New York, NY, 2021. 4

[9] Ashirbani Saha, Michael R Harowicz, Lars J Grimm, Jingxi Weng, Elizabeth H Cain, Connie E Kim, Sujata V Ghate, Ryan Walsh, and Maciej A Mazurowski. Dynamic contrast-enhanced magnetic resonance images of breast cancer patients with tumor locations [data set]. *The Cancer Imaging Archive*, 10, 2021. 1

[10] Choon Hua Thng, Tong San Koh, David J Collins, and Dow Mu Koh. Perfusion magnetic resonance imaging of the liver. *World Journal of Gastroenterology: WJG*, 16(13): 1598, 2010. 4

[11] Thomas Vancauwenberghe, Annemiek Snoeckx, Dirk Vanbeckevoort, Steven Dymarkowski, and Filip M Vanhoenacker. Imaging of the spleen: what the clinician needs to know. *Singapore medical journal*, 56(3):133, 2015. 1

[12] Li Zhou, Tian-wu Chen, Xiao-ming Zhang, Cheng-jun Li, Zhen-feng Yang, Nan-lin Zeng, Li-ying Wang, Ting Li, Dan Wang, Jie Li, et al. Spleen dynamic contrast-enhanced magnetic resonance imaging as a new method for staging liver fibrosis in a piglet model. *PLoS One*, 8(12):e83697, 2013. 1