AbdomenGen: Sequential Volume-Conditioned Diffusion Framework for Abdominal Anatomy Generation

Yubraj Bhandari[1,3][0009−0004−7279−4097], Lavsen Dahal[1,2][0000−0002−8991−759X], Paul Segars[1][0000−0003−3687−5733], and Joseph Y. Lo[1,2][0000−0002−9540−5072]

[1]Center for Virtual Imaging Trials, Dept. of Radiology, Duke University School of Medicine
[2] Dept. of Electrical and Computer Engineering, Pratt School of Engineering, Duke University
[3] Dept. of Physics, Trinity School of Arts and Sciences, Duke University
{yubraj.bhandari,joseph.lo}@duke.edu

Abstract. Computational phantoms are widely used in medical imaging research, yet current systems to generate controlled, clinically meaningful anatomical variations remain limited. We present **AbdomenGen**, a sequential volume-conditioned diffusion framework for controllable abdominal anatomy generation. We introduce the **Volume Control Scalar (VCS)**, a standardized residual that decouples organ size from body habitus, enabling interpretable volume modulation. Organ masks are synthesized sequentially, conditioning on the body mask and previously generated structures to preserve global anatomical coherence while supporting independent, multi-organ control. Across 11 abdominal organs, the framework generates realistic anatomies, achieving highest fidelity for major solid organs (e.g., liver Dice 0.83 ± 0.05). VCS yields stable single-organ calibration over $[-3, +3]$ and disentangled multi-organ modulation. To showcase clinical utility on a hepatomegaly cohort selected from Merlin Dataset, Wasserstein-based VCS selection reduces distributional distance by 73.6% relative to the training distribution. These results demonstrate calibrated, distribution-aware anatomical generation suitable for controllable abdominal phantom construction.

Keywords: Computational Phantoms · Diffusion Models · Controllable Generation · Anatomical Modeling

1 Introduction

Computational anatomical phantoms underpin virtual imaging trials, radiation dosimetry, and surgical simulation [17, 22, 4]. As deep learning systems are deployed, retrospective data still undersample rare, extreme, and pathological anatomies, precisely where failures concentrate. Phantom populations enable controlled stress-testing by generating and enriching such regimes on demand [8].

A useful population therefore requires fidelity (realistic, spatially coherent organs), diversity (habitus/morphology/inter-organ variation), and targeted controllability (clinically grounded, calibrated enrichment).

To our knowledge, no existing method satisfies all three requirements for multi-organ abdominal shape synthesis with habitus-aware, organ-wise control. Model-based phantom frameworks, such as the XCAT and related extensions [17, 6], range from manual construction to scalable deep learning pipelines, but do not support deliberate or calibrated anatomical manipulation. Fully generative approaches offer more flexibility, yet remain limited to single-organ synthesis or uncalibrated latent interpolation [11, 15, 12]. MAISI targets paired CT image and mask synthesis with coarse volume conditioning, but shape controllability is secondary and does not support independent organ-wise manipulation within a fixed anatomical context [8]. Recent controllable abdominal and full-torso generators target CT/MRI image-volume synthesis [10, 21, 23, 1] or text-conditioned mask-and-CT generation [7], leaving habitus-calibrated, organ-wise multi-organ shape control unaddressed.

We propose **AbdomenGen**, a sequential, volume-conditioned diffusion framework for controllable abdominal phantom generation over 11 structures. We target the abdomen because its mix of deformable solid organs, filling-dependent hollow structures, and compact organs with tight margins imposes strong inter-organ constraints that stress controllable generation. Two challenges motivate our design: raw organ-volume conditioning produces contradictory signals against body habitus conditioning, and joint multi-organ generation struggles to maintain spatial consistency. Our main contributions are:

1. **AbdomenGen**: a diffusion framework generating 11 abdominal organs sequentially, each conditioned on the body mask and prior structures, for spatial coherence with independent organ-wise control.
2. **Volume Control Scalar (VCS)**: a habitus-decoupled conditioning variable resolving the contradictory-signal problem of raw-volume conditioning, giving a calibrated, interpretable size-control axis.
3. **Distribution-level controllability**: assigning VCS distributions matches a target population, reducing Wasserstein distance to a hepatomegaly cohort by 73.6% without retraining.

2 Method

Problem Formulation. Organ synthesis is modeled as conditional generation of an organ signed distance field (SDF) as shown in Fig. 1. For organ o, let S_o denote its SDF, B the body-mask SDF, S_o^{ctx} an aggregated context SDF from surrounding organs, and $v_o \in \mathbb{R}$ a scalar size control. We learn

$$p_{\theta_o}\big(S_o \mid B, S_o^{\mathrm{ctx}}, v_o\big)\,.$$

Shapes are represented as SDFs [19, 5], giving a smooth geometric signal on a fixed grid.

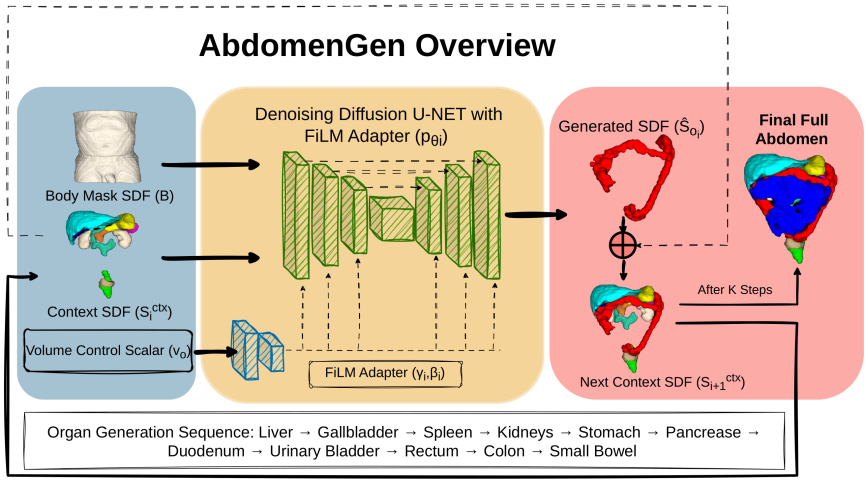


Fig. 1: **Overview of AbdomenGen.** sequential organ generation framework based on a denoising diffusion U-Net with FiLM conditioning. To generate organ o_i (shown left to right for colon), the model p_{θ_i} takes as input the body mask signed distance field (SDF) B, current context SDF S_i^{ctx}, and volume control scalar v_o. The scalar is injected through a FiLM adapter to modulate intermediate network features. The generated organ SDF $\hat{S}_{o_i}$ is obtained and incorporated into the context to form S_{i+1}^{ctx}. Organs are generated in a predefined order (bottom), and the accumulated outputs form the final full abdominal anatomy.

Sequential Generation Strategy. We synthesize the abdomen by generating organs sequentially in a fixed order, from large spatially defining structures to smaller, more dependent organs. For organ o_i, a fixed-size context SDF by aggregating previously generated organs is formed as: $S_i^{\text{ctx}} = \max_{j<i} S_{o_j}$. This representation encodes occupied space without increasing input channels as i grows. During training, S_i^{ctx} is built from reference mask; at inference it is constructed autoregressively from generated predictions.

Volume Control Scalar (VCS). We aim to explicitly control organ size, but organ volume is coupled to global body habitus. Prior work reports associations between liver/spleen volume and anthropometric measures [13, 16], and we observe similar trends: organ volume is partially predictable from body mask volume, with Pearson correlations of $r = 0.64$ (liver), $r = 0.52$ (kidneys), $r = 0.24$ (spleen), and $r = 0.28$ (stomach). Conditioning on raw organ volume would therefore partially re-encode information already provided by the body mask B and may request body–organ volume pairs not supported by the training distribution.

To decouple size control from habitus, we model the expected organ volume as a linear function of body volume,

$$\hat{V}_{o_i}(V_B) = a_i V_B + b_i,$$

with (a_i, b_i) fit on the training set. We define the residual and its standardized form (VCS) as

$$r_{o_i} = V_{o_i}^{\text{ref}} - \hat{V}_{o_i}(V_B), \qquad v_i = \frac{r_{o_i} - \mu_i}{\sigma_i}, \tag{1}$$

where (μ_i, σ_i) are the mean and standard deviation of residuals for organ i. We apply this uniformly across organs for consistency; even when correlations are weak (e.g., gallbladder), residualization keeps the control variable decorrelated from global body scale while preserving anatomical coherence.

Diffusion Architecture with FiLM. A denoising diffusion model is trained per organ [9]. At timestep t, the network input is $x_t = [S_{o,t}^{\text{noisy}}; B; S_o^{\text{ctx}}]$, i.e., the noisy organ SDF, the body-mask SDF B, and a context SDF encoding surrounding organs. We condition on the volume control scalar v_o using FiLM [14] and use x_0-parameterization to predict the clean SDF: $\widehat{S}_o = f_{\theta_o}(x_t, t, v_o)$.

Training Objective. Let $\widehat{S}_{o_i}$ denote the predicted SDF, $S_{o_i}^{\text{ref}}$ the reference mask, and soft occupancy map is obtained via sigmoid function as $\widehat{M}_{o_i} = \sigma(10 \odot \widehat{S}_{o_i})$ and $M_{o_i}^{\text{ref}}$ is obtained by simple thresholding. We use an ℓ_1 reconstruction loss $\mathcal{L}_{\text{SDF}}$ and binary cross-entropy loss $\mathcal{L}_{\text{BCE}}$ as,

$$\mathcal{L}_{\text{SDF}} = \|\widehat{S}_{o_i} - S_{o_i}^{\text{ref}}\|_1 \qquad \mathcal{L}_{\text{BCE}} = \text{BCE}\Big(\widehat{M}_{o_i}, M_{o_i}^{\text{ref}}\Big).$$

Additionally, we introduce a differentiable reverse Dice penalty $\mathcal{L}_{\text{ov}}$ to discourage context-organ overlap. After the initial training epochs, we activate a VCS calibration loss $\mathcal{L}_{\text{vcs}}$, where the predicted organ volume (computed from the generated mask) yields a standardized residual $\hat{v}_i$ that is matched to the target control v_i:

$$\mathcal{L}_{\text{ov}} = 1 - \frac{2\sum_x \hat{M}_i(x)\, M_i^{\text{ctx}}(x)}{\sum_x \hat{M}_i(x) + \sum_x M_i^{\text{ctx}}(x) + \epsilon}, \qquad \mathcal{L}_{\text{vcs}} = \|\hat{v}_i - v_i\|_2^2.$$

Thus, the total loss function is, $\mathcal{L} = \mathcal{L}_{\text{SDF}} + \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{ov}} + \mathcal{L}_{\text{vcs}}$

3 Experimental Setup

Data. We use 556 abdominal CT scans from Duke University Health System, split into 314 training, 42 internal validation, and 200 held-out test cases. Multi-organ masks were produced by the organ segmentation framework DukeSeg [6] and reviewed by a physician for quality control. These masks serve as reference anatomical supervision for training and evaluation. We also evaluate distribution matching on an external hepatomegaly cohort (from MERLIN) [2]. Volumes are cropped to the abdominal region (liver to rectum), resampled to a 128^3 grid, and converted to signed distance functions.

Table 1: **Anatomical fidelity: generated vs. reference masks**. Dice and ASSD/HD95/Chamfer (mm) reported as Mean ± Std

Organ	**Dice↑**	**ASSD↓**	**HD95↓**	**Chamfer↓**
Liver	0.83 ± 0.05	5.84 ± 2.35	17.55 ± 9.34	11.68 ± 4.69
Spleen	0.68 ± 0.13	6.22 ± 2.71	18.79 ± 9.23	12.44 ± 5.41
Kidneys	0.60 ± 0.23	18.34 ± 36.34	33.74 ± 58.26	36.68 ± 72.67
Pancreas	0.53 ± 0.19	6.79 ± 4.93	25.88 ± 21.76	13.58 ± 9.86
Stomach	0.68 ± 0.15	6.36 ± 4.15	21.48 ± 18.24	12.72 ± 8.30
Small bowel	0.48 ± 0.15	9.16 ± 6.30	31.49 ± 26.62	18.32 ± 12.60
Colon	0.33 ± 0.16	24.94 ± 17.76	78.57 ± 43.72	49.89 ± 35.51
Rectum	0.65 ± 0.17	5.87 ± 8.75	17.07 ± 14.98	11.73 ± 17.50
Duodenum	0.51 ± 0.22	6.52 ± 5.28	22.08 ± 19.62	13.03 ± 10.56
Gallbladder	0.60 ± 0.16	4.20 ± 1.33	12.39 ± 5.23	8.39 ± 2.66
Urinary bladder	0.65 ± 0.08	5.66 ± 6.68	16.04 ± 8.76	11.31 ± 13.37

Training and inference. We implement a 3D diffusion U-Net in MONAI [3] with channel widths $[32, 64, 64, 128, 256]$. Models are trained for 1000 linear-β diffusion steps to predict x_0 (clean SDF) using AdamW (10^{-4} learning rate, 10^{-4} weight decay). For robustness to missing controls, we drop conditioning signals with probability $p = 0.3$. A full 11-organ generation takes $\sim$7 min/patient on an RTX A5000 ($\sim$38 s/organ, 10-step DDIM [18]); sequential synthesis keeps peak memory at organ level.

Evaluation. After rigid ANTs registration [20], we report Dice, and extract surface points with PCA alignment to report ASSD, HD95, and symmetric Chamfer distance (mm). Manifold realism is assessed by nearest-training-neighbor distances (Chamfer) for both generated and test samples. Diversity is measured via pairwise Dice/Chamfer dispersion among generated livers, and distribution matching is quantified by the 1D Wasserstein-1 distance (W_1) between generated and target liver-volume distributions across VCS sweeps.

4 Results

4.1 Shape Fidelity and Diversity

Table 1 shows strongest agreement for the liver (Dice 0.83 ± 0.05), with performance decreasing for smaller or more variable organs. Two organs stand out: the kidneys, modeled as a single aggregate label despite being disconnected and often Left/Right-asymmetric (Dice spread 0.60 ± 0.23), and the colon, whose elongated tubular anatomy is poorly captured by an SDF, producing path discontinuities.

To rule out memorization, we compare each generated sample's nearest-training-neighbor distance against the corresponding reference-mask distance (Table 2). Across all 11 organs the difference Δ stays near zero within ± 1 mm Chamfer; kidneys and colon are the only exceptions, consistent with the representation limits above. Since differences are computed end-to-end over the full

Table 2: **Manifold realism** Chamfer distance (mm) for all organs reported as Mean ± std. **(A)** Generated-to-nearest-training sample distance **(B)** reference-test-to-nearest-training distance (empirical anatomical variability); **(C)** Difference Δ = (A)−(B), near zero indicating generated shapes lie within test-set variability.

Organ	(A) Gen-vs-NN	(B) Ref-vs-NN	(C) Difference (Δ)
Liver	11.28 ± 2.02	12.04 ± 2.41	-0.86 ± 2.18
Spleen	6.93 ± 1.58	7.19 ± 1.68	-0.27 ± 1.74
Kidneys	8.45 ± 11.20	5.82 ± 0.71	$+2.73 \pm 11.54$
Pancreas	7.63 ± 1.89	7.34 ± 1.03	$+0.24 \pm 2.15$
Stomach	8.53 ± 1.66	8.95 ± 1.93	-0.32 ± 2.35
Small bowel	14.74 ± 1.85	15.12 ± 3.40	-0.48 ± 3.44
Colon	33.53 ± 9.76	28.89 ± 7.23	$+4.51 \pm 9.80$
Rectum	7.07 ± 1.78	7.50 ± 1.12	-0.42 ± 1.89
Duodenum	7.09 ± 1.08	7.55 ± 1.25	-0.42 ± 1.49
Gallbladder	5.41 ± 1.38	4.70 ± 0.65	$+0.72 \pm 1.05$
Urinary bladder	4.98 ± 0.65	5.90 ± 1.06	-0.92 ± 1.20

Table 3: **Liver diversity: AbdomenGen vs. MAISI.** Pairwise dispersion over generated livers (mean ± std), with MAISI reported across its usable size range.

Model	Pairwise Dice↓	Pairwise Chamfer (mm)↑
MAISI (size=1.0)	0.957 ± 0.043	3.71 ± 3.43
MAISI (size=0.9)	0.803 ± 0.057	13.97 ± 4.31
MAISI (size=0.8)	0.802 ± 0.098	13.37 ± 6.69
Ours (ref masks VCS)	$\mathbf{0.755 \pm 0.096}$	$\mathbf{17.01 \pm 6.82}$

sequence, the small late-stage values indicate sequential error remains bounded rather than compounding. HD95 also showed similar trends across all organs.

Figure 2 illustrates representative samples.

Liver diversity vs. MAISI. Table 3 compares liver dispersion against MAISI across its size parameter. Small values (0.1–0.3) produce partially cropped livers and are not meaningful anatomical baselines, so we restrict comparison to high-size parameter for MAISI. Across all settings AbdomenGen attains the greatest inter-sample variability, consistent with patient-specific conditioning producing a broader range of plausible shapes.

4.2 Volume Controllability and Calibration

VCS ablation. We train the same architecture with raw-volume, z-score-volume, and VCS conditioning, then sweep each and measure liver-volume targeting (Table 4). Raw-volume targets poorly, as the request re-encodes information already in the body mask, while z-score improves it and VCS performs best (239.1 mL,

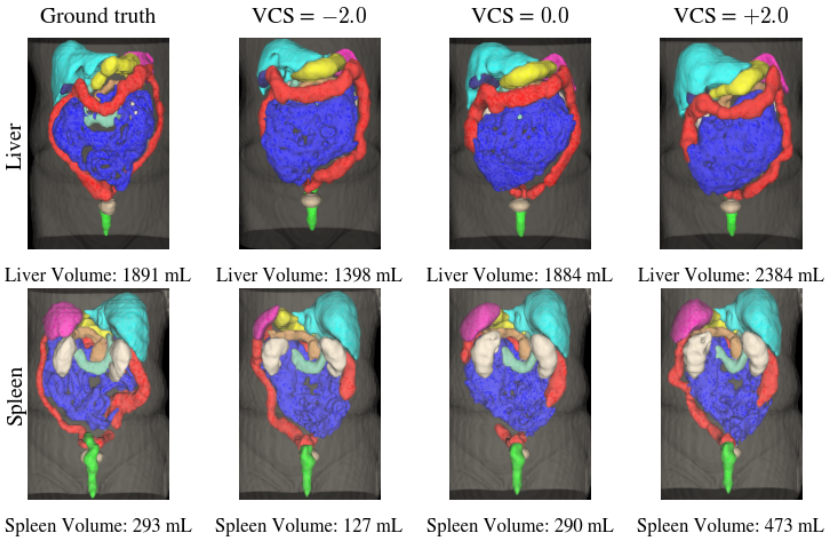


Fig. 2: **Qualitative evaluation:** Rows correspond to Liver (cyan, anteroposterior view) and Spleen (pink, posteroanterior view). Columns (left to right) show Reference Masks and VCS = −2, 0, and +2 respectively. As VCS varies, target organ volume (below each rendering) exhibits controlled changes while surrounding anatomical structures remain spatially coherent.

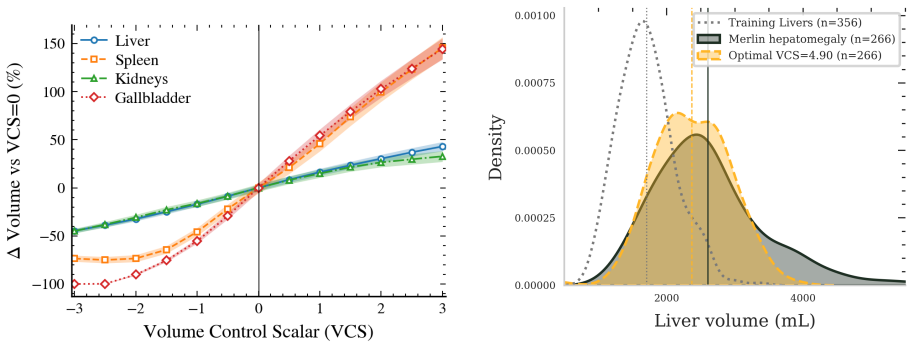
Table 4: **VCS conditioning ablation** (liver). Lower MAE indicates more accurate volume targeting; slope closer to 1 a stronger response to the requested control; higher R^2 more consistent calibration.

Conditioning	**MAE (mL)↓**	**Slope (→1)**	**R^2↑**
Raw volume	460.3	0.39	0.13
Z-score volume	286.3	0.59	0.50
VCS (ours)	**239.1**	**0.74**	**0.70**

slope 0.74, R^2 0.70). For weakly correlated organs the linear habitus model reduces VCS to a centered, scaled volume parameter while preserving a standardized control axis.

Single-organ calibration. Fig 3(a) shows all evaluated organs exhibit a stable scaling response to VCS. Normalization to baseline (VCS =0) generation enables cross-organ comparison. Liver and kidneys follow a near-linear trend, while spleen and gallbladder show amplified growth at positive VCS. The tight 95% CIs indicate consistent realization of the requested volumes.

Distribution matching to hepatomegaly. To match a pathological target distribution, we sweep liver VCS and compute the Wasserstein-1 distance between gen-



(a) Cross-organ VCS calibration. (b) Distribution-level calibration.

Fig. 3: **(a) Mean organ volume change (Δ%)**. Shown relative to VCS = 0 (mean ± 95% CI) as VCS is swept over [-3, 3]. **(b) Kernel density estimates of liver volumes**. Training cohort (dotted), Merlin hepatomegaly reference masks (solid), and Generated samples under the selected VCS setting (dashed). Vertical dashed lines denote cohort means.

erated liver volumes and the Merlin hepatomegaly cohort [2]. The optimum occurs at VCS= 4.9, reducing W_1 from 899.5 mL (training vs. Merlin) to 237.0 mL (73.6%). Figure 3(b) shows that, at this setting, generated volumes closely align with the Merlin distribution while remaining distinct from the training cohort.

Independent multi-organ control. Figure 4 evaluates joint liver–spleen conditioning. Changing one organ's VCS shifts samples primarily along the intended axis with minimal movement in the orthogonal dimension, indicating limited cross-organ coupling. This enables targeted exploration of underrepresented joint configurations beyond high-density regions of the training distribution.

5 Conclusion

We introduce a statistically grounded volume-conditioning strategy for controllable anatomical generation. Our Volume Control Scalar (VCS) parameterizes organ size as a standardized residual relative to body habitus, yielding an interpretable control axis. This enables calibrated volume modulation with approximately linear response and supports stable independent control across multiple organs. By selecting VCS operating points to match clinically defined target distributions, the framework can synthesize cohort-calibrated phantom populations without latent-space interpolation heuristics or post-hoc registration.

Limitations and future work. Sequential synthesis can accumulate small geometric errors across organs, volume control alone does not fully specify local morphology, and the SDF representation performs less reliably for tubular structures such as the colon. Future work will ablate the sequential generation order,

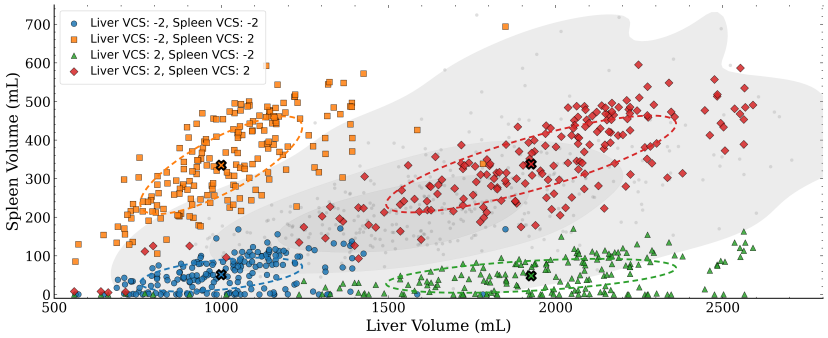


Fig. 4: **Joint liver–spleen volume distributions** under independent VCS conditioning (±2). Gray contours denote the empirical training distribution; colored clusters show diverse controlled generations.

extend the approach to additional anatomical regions, incorporate richer conditioning signals (e.g., shape and tissue descriptors), and quantify downstream impact in image synthesis, segmentation stress-testing, and simulation pipelines.

Acknowledgments. This work was funded by the Center for Virtual Imaging Trials, NIH grants P41EB028744, R01EB001838, and R01CA261457.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Belton, N., Joppin, V., Lawlor, A., Curran, K.M., Masson, C., Bege, T., Bendahan, D.: Dyabd: A dataset and technique for synthetically generating dynamic abdominal mris with dual class and anatomically conditioned diffusion models. In: Medical Imaging with Deep Learning (2024)
2. Blankemeier, L., Cohen, J.P., Kumar, A., Van Veen, D., Gardezi, S.J.S., Paschali, M., Chen, Z., Delbrouck, J.B., Reis, E., Truyts, C., et al.: Merlin: A vision language foundation model for 3d computed tomography. Research Square pp. rs–3 (2024)
3. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
4. Chheang, V., Saalfeld, P., Joeres, F., Boedecker, C., Huber, T., Huettl, F., Lang, H., Preim, B., Hansen, C.: A collaborative virtual reality environment for liver surgery planning. Computers & Graphics **99**, 234–246 (2021)
5. Chou, G., Bahat, Y., Heide, F.: Diffusion-sdf: Conditional generative modeling of signed distance functions. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 2262–2272 (2023)
6. Dahal, L., Ghojoghnejad, M., Vancoillie, L., Ghosh, D., Bhandari, Y., Kim, D., Ho, F.C., Tushar, F.I., Luo, S., Lafata, K.J., et al.: Xcat 3.0: A comprehensive library of personalized digital twins derived from ct scans. Medical Image Analysis p. 103636 (2025)

7. Dai, L., Zhang, R., Yu, Y., Zhang, X.: Guidegen: A text-guided framework for paired full-torso anatomy and ct volume generation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 3470–3478 (2026)
8. Guo, P., Zhao, C., Yang, D., Xu, Z., Nath, V., Tang, Y., Simon, B., Belue, M., Harmon, S., Turkbey, B., et al.: Maisi: Medical ai for synthetic imaging. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 4430–4441. IEEE (2025)
9. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
10. Konz, N., Chen, Y., Dong, H., Mazurowski, M.A.: Anatomically-controllable medical image generation with segmentation-guided diffusion models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 88–98. Springer (2024)
11. Li, J., Pepe, A., Luijten, G., Schwarz-Gsaxner, C., Kleesiek, J., Egger, J.: Anatomy completor: A multi-class completion framework for 3d anatomy reconstruction. In: Shape in Medical Imaging, Lecture Notes in Computer Science, vol. 14350, pp. 1–14. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-46914-5_1
12. Mouheb, K., Nejad, M.G., Dahal, L., Samei, E., Lafata, K.J., Segars, W.P., Lo, J.Y.: Large intestine 3d shape refinement using conditional latent point diffusion models. In: Shape in Medical Imaging, Lecture Notes in Computer Science, vol. 16171, pp. 103–116. Springer, Cham (2025). https://doi.org/10.1007/978-3-032-06774-6_8
13. Olthof, P.B., van Dam, R., Jovine, E., Campos, R.R., de Santibañes, E., Oldhafer, K., Malago, M., Abdalla, E.K., Schadde, E.: Accuracy of estimated total liver volume formulas before liver resection. Surgery **166**(3), 247–253 (2019)
14. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: Film: Visual reasoning with a general conditioning layer. In: Proceedings of the AAAI conference on artificial intelligence. vol. 32 (2018)
15. Podobnik, G., Balodi, N., Killeen, B.D., Vrtovec, T., Unberath, M.: Anatomygen: Generating anatomically plausible human phantoms at high resolution. In: Shape in Medical Imaging, Lecture Notes in Computer Science, vol. 16171, pp. 129–139. Springer, Cham (2025). https://doi.org/10.1007/978-3-032-06774-6_10
16. Prassopoulos, P., Daskalogiannaki, M., Raissaki, M., Hatjidakis, A., Gourtsoyiannis, N.: Determination of normal splenic volume on computed tomography in relation to age, gender and body habitus. European radiology **7**(2), 246–248 (1997)
17. Segars, W.P., Sturgeon, G., Mendonca, S., Grimes, J., Tsui, B.M.W.: 4d xcat phantom for multimodality imaging research. Medical Physics **37**(9), 4902–4915 (2010). https://doi.org/10.1118/1.3480985
18. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. arXiv preprint arXiv:2010.02502 (2020)
19. Takikawa, T., Litalien, J., Yin, K., Kreis, K., Loop, C., Nowrouzezahrai, D., Jacobson, A., McGuire, M., Fidler, S.: Neural geometric level of detail: Real-time rendering with implicit 3d shapes. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11358–11367 (2021)
20. Tustison, N.J., Cook, P.A., Holbrook, A.J., Johnson, H.J., Muschelli, J., Devenyi, G.A., Duda, J.T., Das, S.R., Cullen, N.C., Gillen, D.L., et al.: The antsx ecosystem for quantitative biological and medical imaging. Scientific reports **11**(1), 9068 (2021)
21. Wang, Y., Wan, Z., Qiu, Y., Wang, Z.: Devil is in details: Locality-aware 3d abdominal ct volume generation for self-supervised organ segmentation. In: Proceedings of the 32nd ACM International Conference on Multimedia. pp. 10640–10648 (2024)

22. Xu, X.G.: An exponential growth of computational phantom research in radiation protection, imaging, and radiotherapy: A review of the fifty-year history. Physics in Medicine & Biology **59**(18), R233–R302 (2014). https://doi.org/10.1088/0031-9155/59/18/R233
23. Zhuang, Y., Hou, B., Mathai, T.S., Mukherjee, P., Kim, B., s, R.M.: Semantic image synthesis for abdominal ct. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 214–224. Springer (2023)