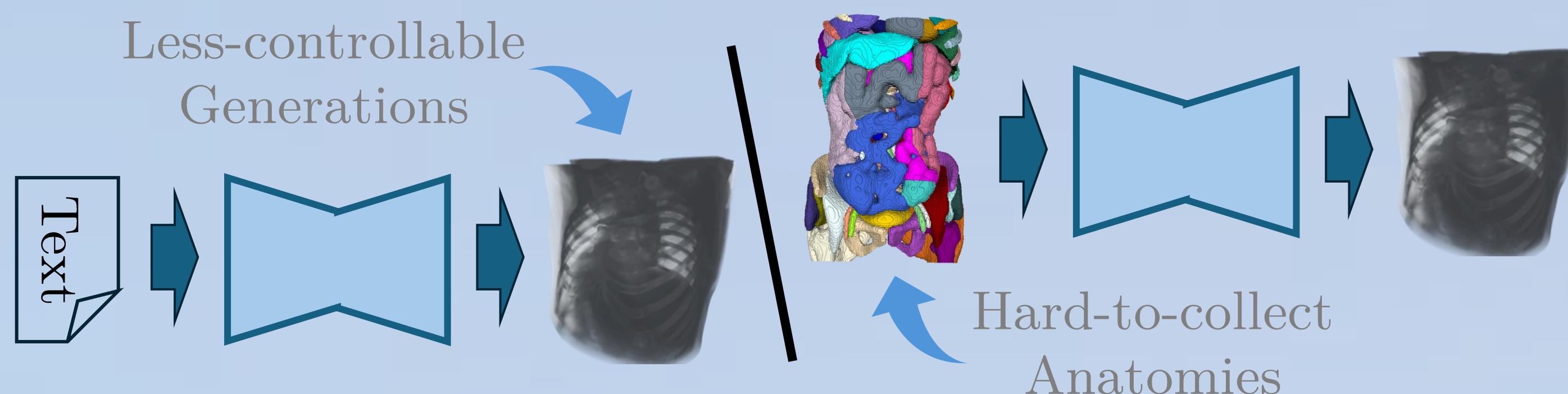


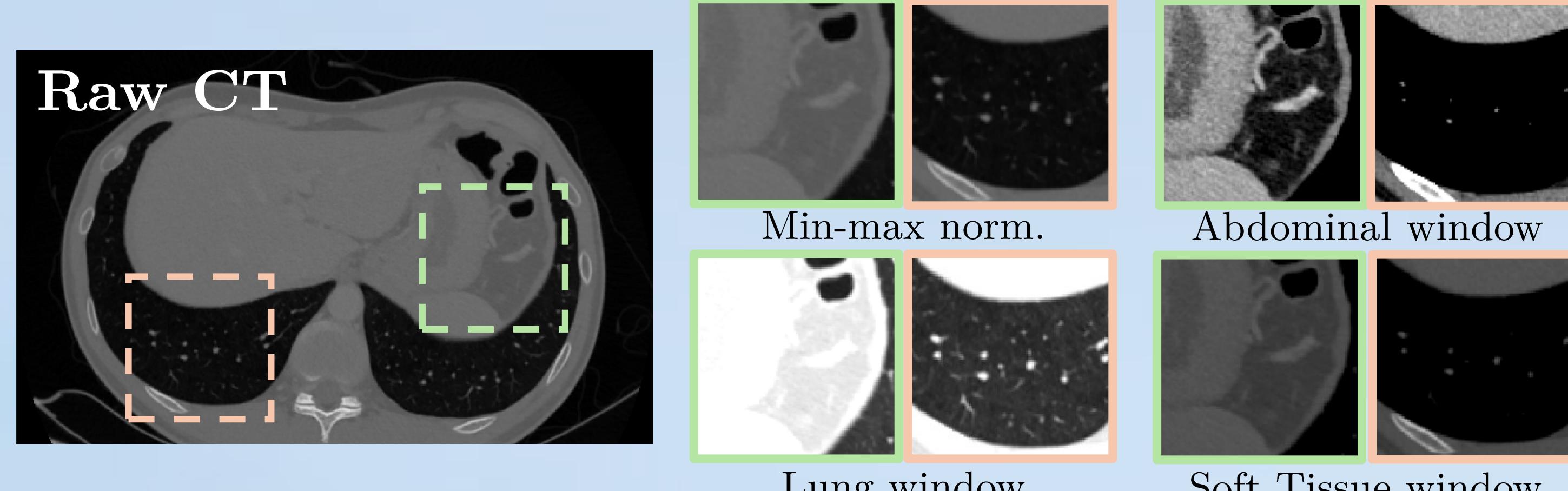


## I. Motivation

- Text-based generation frameworks often yields anatomies uncontrollable for segmentation purposes;
- Mask-based generation frameworks lack the flexibility for effortless sample synthesis at downstream;

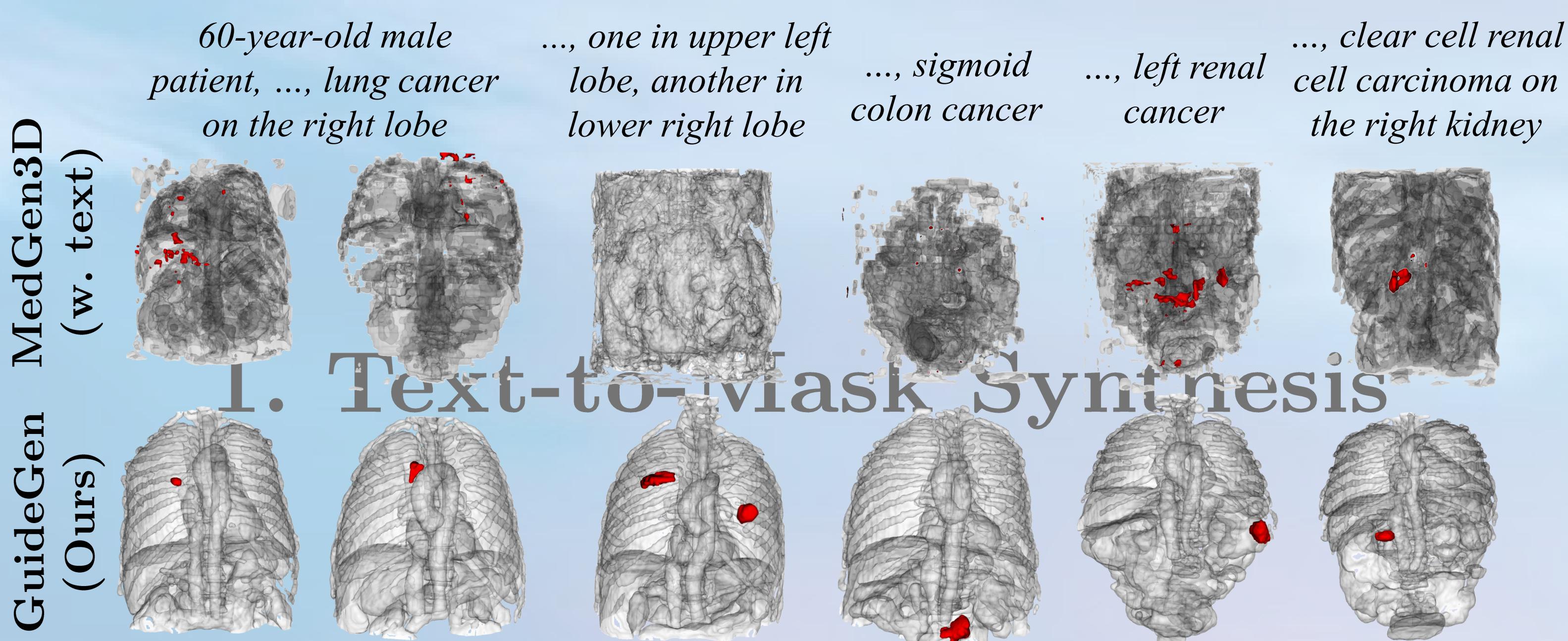


- By incorporating a separate mask synthesis module, our **GuideGen** relies solely on textual inputs to retrieve samples ready for a wide range of downstream tasks.
- Current generation frameworks often consider patches of local anatomy characterized by a normalization window;



- By extracting comprehensive anatomical features from multiple contrast levels, our **GuideGen** can generate patches throughout the torso without loss of details.

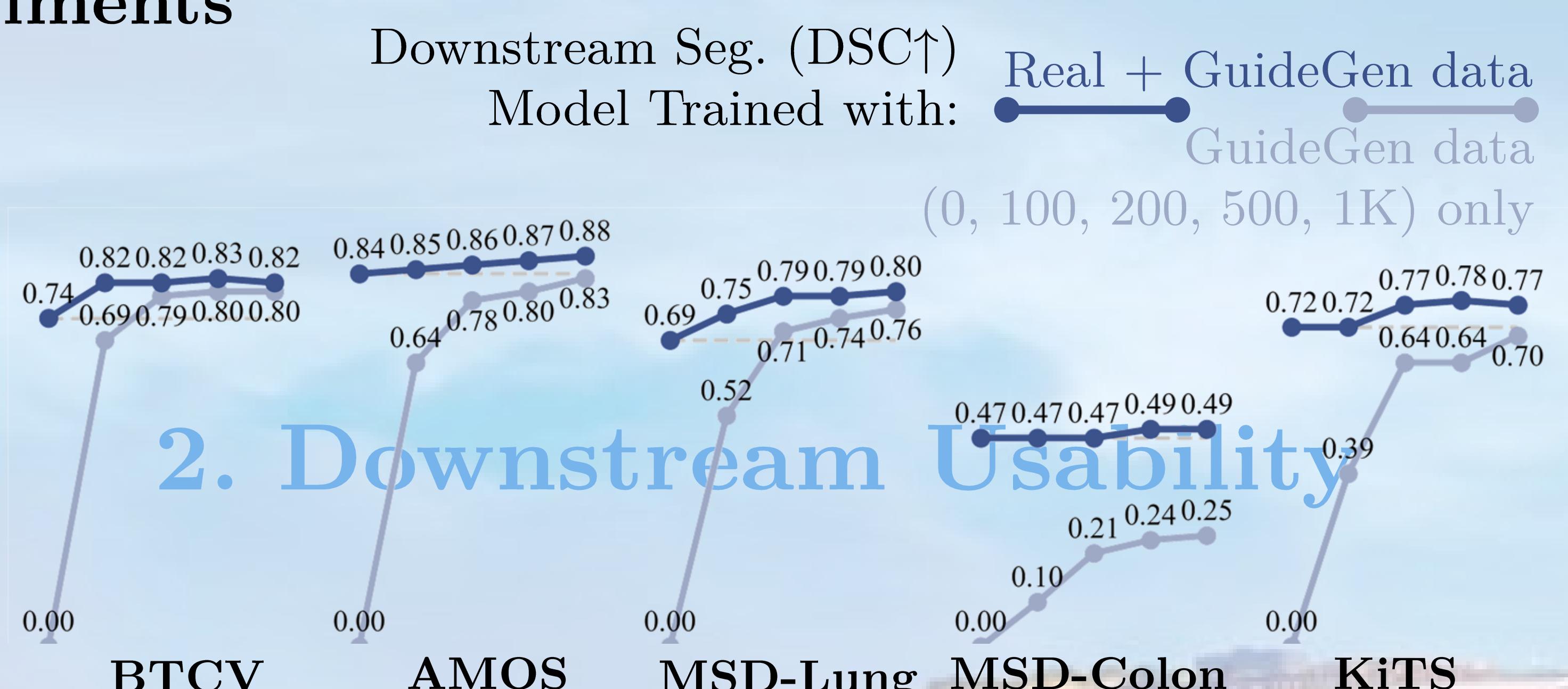
## III. Experiments



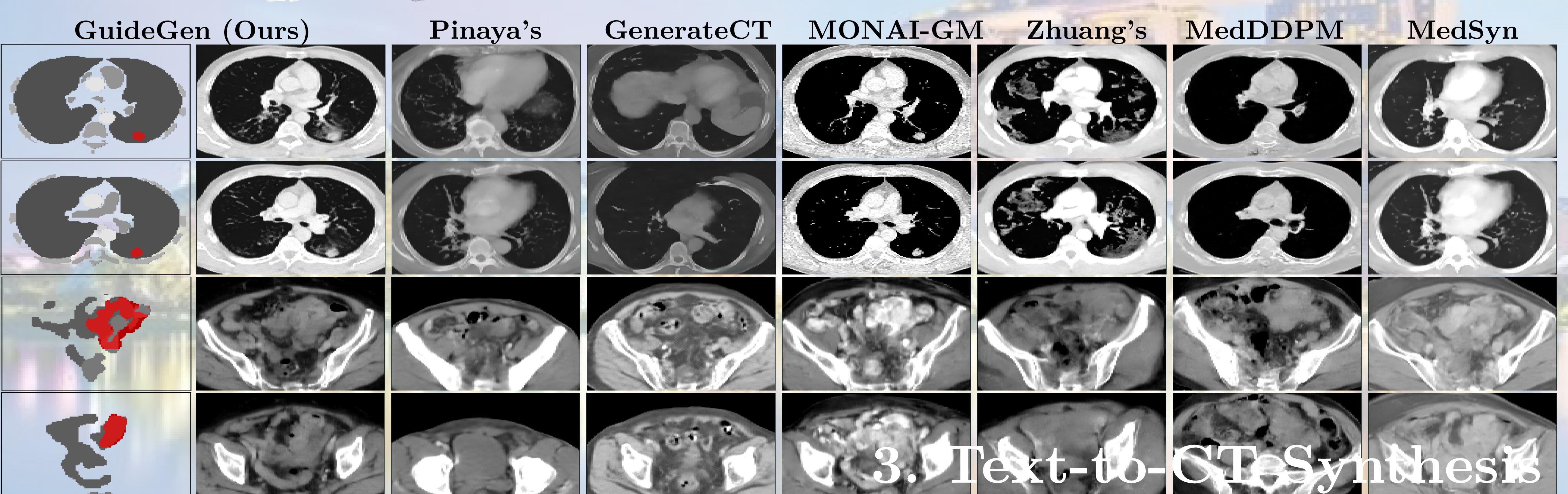
### 1. Text-to-Mask Synthesis

This patient is a white male. In this medical CT scan, his clinical status can be described as an advanced lung cancer in the upper left lobe, with a TNM staging of T3NMX.

This patient is a black African American of unknown age and non-Hispanic or Latino race. In this CT image, the patient's condition is described as huge tumor around the descendant colon, of stage T4NMX.



### 2. Downstream Usability



### 3. Text-to-CT Synthesis

## 4. Modality Alignment

(CT-Mask & CT-Prompt; Please refer to our paper for Mask-Prompt Alignment Analysis)

Methods	DSC↑										
	Spl.	Kid.	Liver	Sto.	Pan.	Lung	S.B.	Duo.	Colon	Heart	Avg.
MAISI	0.73	0.72	0.80	0.60	0.43	0.84	0.49	0.35	0.55	0.40	0.59
Zhuang's	0.43	0.43	0.62	0.36	0.21	0.69	0.37	0.23	0.26	0.24	0.38
MedDDPM	0.35	0.36	0.39	0.54	0.29	0.34	0.47	<b>0.43</b>	0.67	0.22	0.41
MedSyn	0.52	0.51	0.51	0.40	0.07	0.59	0.12	0.01	0.54	0.22	0.35
GuideGen	<b>0.75</b>	<b>0.72</b>	<b>0.90</b>	<b>0.63</b>	<b>0.46</b>	<b>0.84</b>	<b>0.51</b>	<b>0.41</b>	<b>0.70</b>	<b>0.53</b>	<b>0.65</b>

Methods	Age	Gender	Accuracy↑		
			Race	Tumor Loc.	Avg.
Pinaya's	0.06	0.35	0.10	0.17	0.17
GenerateCT	0.07	0.21	0.44	0.03	0.19
MedSyn	0.17	0.74	<b>0.51</b>	<b>0.47</b>	0.47
GuideGen	<b>0.39</b>	<b>0.90</b>	<b>0.60</b>	<b>0.89</b>	<b>0.69</b>