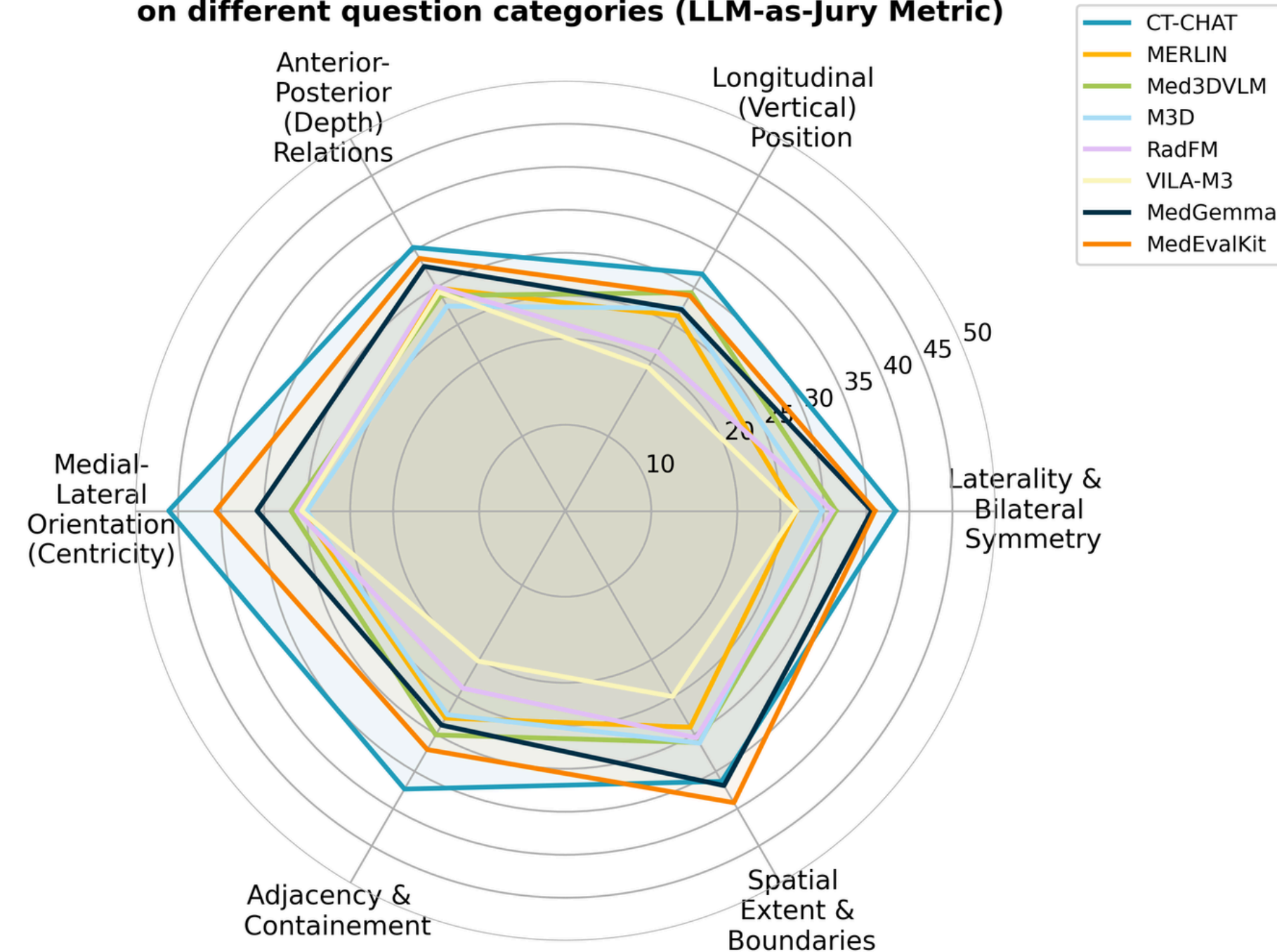


Performance of Medical 3D VLMs on the CT-SpatialVQA Benchmark
on different question categories (LLM-as-Jury Metric)



Category	Spatial Relevance
<ul style="list-style-type: none">• Laterality & Bilateral Symmetry	<ul style="list-style-type: none">• Consistent left-right anatomical grounding relative to the sagittal midline; prevents laterality inversion and side hallucination.
<ul style="list-style-type: none">• Longitudinal (Vertical) Position	<ul style="list-style-type: none">• Requires volumetric integration along the superior-inferior axis and enforces slice-level consistency across the 3D stack.
<ul style="list-style-type: none">• Anterior-Posterior (Depth) Relations	<ul style="list-style-type: none">• Preserves depth cues and front-back anatomical arrangement essential for compartment-level reasoning.
<ul style="list-style-type: none">• Medial-Lateral Orientation (Centricity)	<ul style="list-style-type: none">• Maintains stable central-peripheral localization and global anatomical reference frames.
<ul style="list-style-type: none">• Adjacency & Containment	<ul style="list-style-type: none">• Encodes topological relations, distinguishing touching from containment within anatomical boundaries.
<ul style="list-style-type: none">• Spatial Extent & Boundaries	<ul style="list-style-type: none">• Tracks spatial continuity, confinement, and compartment crossing to model disease spread accurately.