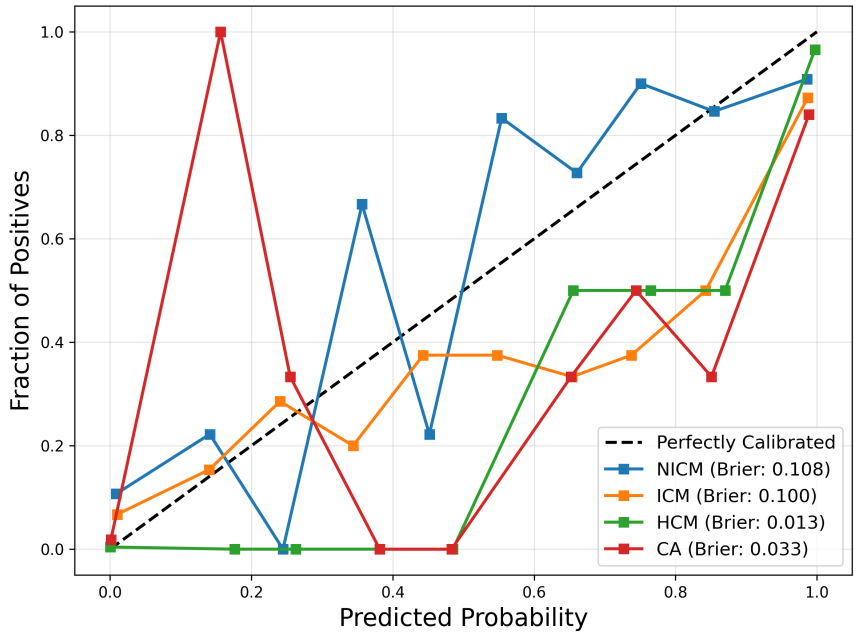
Supplementary Information

Title: Contrastive Language Image Pretraining for a Cardiac Magnetic Resonance Image Embedding with Zero-shot Capabilities

Authors:
[1,2]Makiya Nakashima MS, [3]Jielin Qiu PhD, [3]Peide Huang PhD, [1,2]Jihye Lee, [4]Po-hao Chen MD, MBA, [2,4]Richard Grimm DO, [1,2,4]Christopher Nguyen PhD, FSCMR, FACC, [5]Byung-Hak Kim PhD, [3,6]Ding Zhao PhD, [2,4]*Deborah Kwon MD, [1,4,7]*David Chen PhD

Author affiliation:
[1]Cardiovascular Innovation Research Center, Cleveland Clinic, Cleveland, OH, USA
[2]Heart Vascular Thoracic Institute, Cleveland Clinic, Cleveland, OH, USA
[3]Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA
[4]Diagnostics Institute, Cleveland Clinic, Cleveland, OH, USA
[5]Hyundai Card, Seoul, South Korea
[6]Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[7]Center for Diagnostics and Artificial Intelligence, Cleveland Clinic, Cleveland, OH, USA
*Joint supervision/Correspondence

Table of Contents:



Supplemental Figure 1: Calibration curves and Brier scores for supervised task. Abbreviations: NICM (non-ischemic cardiomyopathy), ICM (ischemic cardiomyopathy), HCM (hypertrophic cardiomyopathy), and CA (cardiac amyloidosis)

Label	Prompts
LVDys	Left ventricle with systolic dysfunction
	Depressed left ventricular function
RVDys	Right ventricle with systolic dysfunction
	Depressed right ventricular function
LVD	Left ventricle is dilated
	Left ventricle is enlarged in size
RVD	Right ventricle is dilated
	Right ventricle is enlarged in size
LVH	There is left ventricular hypertrophy
	There is septal hypertrophy
	LV is thickened
LGE	No enhancement and delayed imaging to suggest ischemic insult or infiltrate pathology
	There are no findings to suggest prior ischemic damage or an infiltrative process
AoD	Aortic root is ectatic
	Ascending aortic is dilated
	Descending aorta is dilated

Supplemental Table 1: Text prompts used for zero-shot classification.

Model	Label	Accuracy	F1-score	AUC
OPENAI CLIP	LVDys	0.537 (0.519-0.555)	0.393 (0.368-0.419)	0.564 (0.540-0.588)
	RVDys	0.520 (0.501-0.539)	0.211 (0.184-0.236)	0.550 (0.517-0.582)
	LVD	0.548 (0.529-0.566)	0.358 (0.331-0.385)	0.576 (0.550-0.600)
	RVD	0.532 (0.513-0.550)	0.140 (0.116-0.163)	0.565 (0.518-0.613)
	LVH	0.459 (0.439-0.475)	0.253 (0.228-0.279)	0.433 (0.407-0.458)
	LGE	0.509 (0.489-0.527)	0.571 (0.550-0.591)	0.501 (0.479-0.523)
	AoD	0.491 (0.471-0.508)	0.157 (0.135-0.181)	0.487 (0.448-0.526)
BiomedCLIP	LVDys	0.554 (0.536-0.573)	0.408 (0.383-0.435)	0.569 (0.545-0.593)
	RVDys	0.525 (0.507-0.545)	0.208 (0.181-0.235)	0.539 (0.507-0.574)
	LVD	0.538 (0.518-0.557)	0.349 (0.323-0.377)	0.553 (0.527-0.578)
	RVD	0.498 (0.479-0.516)	0.115 (0.094-0.137)	0.488 (0.442-0.533)
	LVH	0.482 (0.464-0.500)	0.296 (0.269-0.320)	0.475 (0.448-0.503)

	LGE	0.503 (0.484-0.521)	0.555 (0.531-0.576)	0.498 (0.476-0.519)
	AoD	0.459 (0.441-0.478)	0.152 (0.131-0.175)	0.447 (0.410-0.483)
VideoMAE-CLIP	LVDys	0.789 (0.774-0.803)	0.674 (0.648-0.697)	0.863 (0.850-0.878)
	RVDys	0.683 (0.666-0.699)	0.344 (0.312-0.376)	0.749 (0.725-0.772)
	LVD	0.745 (0.728-0.761)	0.583 (0.554-0.610)	0.825 (0.809-0.841)
	RVD	0.671 (0.654-0.689)	0.223 (0.192-0.251)	0.753 (0.721-0.784)
	LVH	0.860 (0.847-0.872)	0.731 (0.704-0.756)	0.928 (0.917-0.938)
	LGE	0.611 (0.594-0.628)	0.663 (0.645-0.680)	0.651 (0.631-0.672)
	AoD	0.585 (0.567-0.602)	0.208 (0.180-0.235)	0.608 (0.571-0.640)
CMR-CLIP	LVDys	0.733 (0.732-0.733)	0.593 (0.592-0.593)	0.794 (0.793-0.794)
	RVDys	0.687 (0.686-0.687)	0.323 (0.322-0.324)	0.734 (0.732-0.733)
	LVD	0.735 (0.735-0.736)	0.555 (0.554-0.556)	0.806 (0.806-0.807)
	RVD	0.699 (0.699-0.700)	0.251 (0.250-0.252)	0.775 (0.773-0.775)
	LVH	0.770 (0.770-0.771)	0.600 (0.599-0.601)	0.849 (0.848-0.849)
	LGE	0.636 (0.636-0.637)	0.692 (0.691-0.693)	0.680 (0.680-0.681)
	AoD	0.658 (0.658-0.659)	0.275 (0.276-0.278)	0.709 (0.709-0.711)

Supplemental Table 2: Zero-shot classification results with the internal cohort for four contrastive vision–language models (OPENAI CLIP, BiomedCLIP, VideoMAE-CLIP and CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Model	Label	Accuracy	F1-score	AUC
OPENAI CLIP	LVDys	0.500 (0.456-0.547)	0.448 (0.386-0.503)	0.539 (0.480-0.599)
	RVDys	0.329 (0.290-0.371)	0.243 (0.185-0.298)	0.488 (0.415-0.569)
	LVD	0.262 (0.224-0.301)	0.248 (0.196-0.303)	0.399 (0.331-0.473)
	RVD	0.222 (0.182-0.264)	0.170 (0.118-0.218)	0.499 (0.397-0.602)
	LVH	0.248 (0.208-0.287)	0.261 (0.201-0.315)	0.449 (0.372-0.523)
	LGE	0.633 (0.586-0.678)	0.760 (0.721-0.796)	0.502 (0.443-0.562)
	AoD	0.220 (0.182-0.264)	0.197 (0.150-0.251)	0.561 (0.475-0.651)
BiomedCLIP	LVDys	0.596 (0.551-0.643)	0.224 (0.154-0.298)	0.498 (0.438-0.560)
	RVDys	0.694 (0.647-0.736)	0.155 (0.084-0.229)	0.495 (0.426-0.568)
	LVD	0.423 (0.374-0.472)	0.245 (0.181-0.304)	0.478 (0.401-0.550)
	RVD	0.367 (0.322-0.411)	0.140 (0.090-0.196)	0.464 (0.376-0.550)
	LVH	0.472 (0.421-0.519)	0.226 (0.163-0.288)	0.486 (0.411-0.558)
	LGE	0.591 (0.542-0.633)	0.706 (0.662-0.746)	0.523 (0.465-0.578)
	AoD	0.519 (0.472-0.568)	0.169 (0.112-0.233)	0.486 (0.408-0.564)
VideoMAE-CLIP	LVDys	0.701 (0.567-0.602)	0.612 (0.553-0.674)	0.777 (0.733-0.819)
	RVDys	0.456 (0.411-0.502)	0.321 (0.255-0.380)	0.605 (0.539-0.671)
	LVD	0.364 (0.318-0.409)	0.333 (0.274-0.388)	0.674 (0.603-0.736)
	RVD	0.269 (0.229-0.306)	0.216 (0.163-0.266)	0.673 (0.600-0.743)
	LVH	0.664 (0.617-0.706)	0.450 (0.369-0.519)	0.839 (0.789-0.879)
	LGE	0.603 (0.556-0.650)	0.656 (0.608-0.702)	0.669 (0.614-0.724)
	AoD	0.269 (0.229-0.311)	0.208 (0.156-0.262)	0.602 (0.516-0.689)
CMR-CLIP	LVDys	0.715 (0.713-0.716)	0.647 (0.644-0.648)	0.799 (0.798-0.800)
	RVDys	0.355 (0.356-0.359)	0.320 (0.321-0.324)	0.726 (0.725-0.729)
	LVD	0.752 (0.751-0.754)	0.369 (0.364-0.370)	0.733 (0.731-0.735)
	RVD	0.453 (0.452-0.455)	0.259 (0.257-0.261)	0.809 (0.806-0.810)
	LVH	0.556 (0.555-0.558)	0.410 (0.409-0.413)	0.836 (0.835-0.838)
	LGE	0.645 (0.644-0.647)	0.722 (0.721-0.723)	0.665 (0.664-0.667)
	AoD	0.376 (0.375-0.378)	0.208 (0.206-0.209)	0.594 (0.589-0.595)

Supplemental Table 3: Zero-shot classification results with the external cohort for four contrastive vision–language models (OPENAI CLIP, BiomedCLIP, VideoMAE-CLIP and CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Model	Label	Accuracy	F1-score	AUC
OPENAI CLIP	ACDC DCM	0.625 (0.475-0.775)	0.400 (0.111-0.629)	0.639 (0.440-0.812)
	ACDC HCM	0.550 (0.400-0.700)	0.400 (0.160-0.605)	0.563 (0.348-0.779)
	ACDC MI	0.425 (0.275-0.600)	0.258 (0.067-0.452)	0.431 (0.201-0.676)
	ACDC Normal	0.550 (0.400-0.700)	0.308 (0.080-0.538)	0.472 (0.216-0.724)
	ACDC Abn RV	0.400 (0.250-0.550)	0.200 (0.000-0.375)	0.417 (0.158-0.663)
	NICM	0.431 (0.392-0.469)	0.464 (0.414-0.510)	0.397 (0.353-0.442)
	ICM	0.541 (0.502-0.582)	0.360 (0.304-0.419)	0.558 (0.499-0.615)
	CA	0.569 (0.527-0.608)	0.125 (0.077-0.179)	0.549 (0.457-0.649)
	HCM	0.553 (0.512-0.593)	0.274 (0.213-0.339)	0.606 (0.545-0.668)
BiomedCLIP	ACDC DCM	0.575 (0.425-0.725)	0.261 (0.000-0.483)	0.439 (0.230-0.664)
	ACDC HCM	0.300 (0.175-0.450)	0.222 (0.057-0.400)	0.323 (0.147-0.529)
	ACDC MI	0.425 (0.275-0.575)	0.343 (0.129-0.526)	0.745 (0.519-0.944)
	ACDC Normal	0.400 (0.250-0.550)	0.077 (0.000-0.231)	0.355 (0.177-0.540)
	ACDC Abn RV	0.500 (0.349-0.650)	0.286 (0.067-0.485)	0.544 (0.257-0.784)
	NICM	0.598 (0.560-0.637)	0.606 (0.559-0.650)	0.622 (0.580-0.668)
	ICM	0.507 (0.464-0.548)	0.298 (0.242-0.352)	0.464 (0.408-0.519)
	CA	0.481 (0.442-0.521)	0.096 (0.051-0.142)	0.424 (0.322-0.531)
	HCM	0.423 (0.383-0.462)	0.172 (0.122-0.222)	0.356 (0.302-0.406)
VideoMAE-CLIP	ACDC DCM	0.725 (0.575-0.850)	0.621 (0.370-0.800)	0.740 (0.563-0.893)
	ACDC HCM	0.750 (0.625-0.875)	0.615 (0.375-0.800)	0.793 (0.631-0.933)
	ACDC MI	0.450 (0.300-0.625)	0.083 (0.000-0.250)	0.304 (0.078-0.605)
	ACDC Normal	0.775 (0.650-0.900)	0.571 (0.266-0.800)	0.792 (0.617-0.941)
	ACDC Abn RV	0.725 (0.575-0.850)	0.353 (0.000-0.625)	0.686 (0.396-0.942)
	NICM	0.675 (0.636-0.713)	0.690 (0.646-0.733)	0.741 (0.695-0.779)
	ICM	0.567 (0.526-0.608)	0.357 (0.298-0.418)	0.605 (0.551-0.655)
	CA	0.751 (0.716-0.787)	0.233 (0.155-0.316)	0.753 (0.663-0.830)
	HCM	0.749 (0.713-0.785)	0.459 (0.383-0.532)	0.789 (0.747-0.825)
CMR-CLIP	ACDC DCM	0.850 (0.749-0.950)	0.769 (0.556-0.923)	0.981 (0.937-1.000)
	ACDC HCM	0.875 (0.775-0.975)	0.706 (0.400-0.917)	0.980 (0.934-1.000)
	ACDC MI	0.675 (0.525-0.800)	0.316 (0.000-0.571)	0.632 (0.414-0.820)
	ACDC Normal	0.750 (0.600-0.875)	0.500 (0.182-0.762)	0.866 (0.742-0.969)
	ACDC Abn RV	0.675 (0.525-0.825)	0.381 (0.105-0.632)	0.819 (0.625-0.960)
	NICM	0.655 (0.615-0.694)	0.674 (0.633-0.716)	0.682 (0.642-0.727)
	ICM	0.643 (0.603-0.680)	0.461 (0.397-0.519)	0.732 (0.683-0.775)
	CA	0.806 (0.773-0.837)	0.347 (0.250-0.434)	0.867 (0.794-0.926)
	HCM	0.852 (0.821-0.880)	0.629 (0.551-0.696)	0.896 (0.868-0.923)

Supplemental Table 4: Zero-shot classification results with disease cohort for four contrastive vision–language models (OPENAI CLIP, BiomedCLIP, VideoMAE-CLIP and CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Model	Label	Shot	Accuracy	F1-score	AUC
OPENAI CLIP	LVDys	1	0.526 (0.503-0.550)	0.326 (0.310-0.343)	0.511 (0.505-0.518)
		2	0.527 (0.508-0.546)	0.349 (0.337-0.359)	0.525 (0.518-0.531)
		4	0.547 (0.535-0.559)	0.358 (0.348-0.366)	0.535 (0.529-0.541)
		8	0.532 (0.521-0.543)	0.375 (0.368-0.382)	0.544 (0.537-0.551)
		16	0.547 (0.539-0.557)	0.390 (0.383-0.395)	0.564 (0.558-0.571)
		32	0.564 (0.558-0.571)	0.404 (0.398-0.410)	0.585 (0.578-0.591)
	RVDys	1	0.546 (0.514-0.579)	0.192 (0.185-0.199)	0.531 (0.521-0.540)
		2	0.524 (0.493-0.555)	0.201 (0.196-0.206)	0.541 (0.531-0.552)
		4	0.535 (0.510-0.559)	0.209 (0.205-0.214)	0.554 (0.544-0.564)
		8	0.546 (0.531-0.563)	0.218 (0.213-0.223)	0.571 (0.562-0.581)
		16	0.572 (0.560-0.585)	0.237 (0.232-0.242)	0.605 (0.596-0.614)
		32	0.592 (0.581-0.603)	0.251 (0.246-0.256)	0.630 (0.623-0.637)
	LVD	1	0.512 (0.488-0.536)	0.309 (0.289-0.325)	0.534 (0.524-0.544)
		2	0.523 (0.501-0.544)	0.320 (0.306-0.331)	0.538 (0.528-0.548)
		4	0.535 (0.519-0.551)	0.341 (0.332-0.350)	0.554 (0.543-0.564)
		8	0.538 (0.525-0.549)	0.355 (0.347-0.363)	0.569 (0.559-0.579)

		16	0.571 (0.561-0.581)	0.381 (0.375-0.388)	0.605 (0.597-0.614)
		32	0.586 (0.578-0.593)	0.403 (0.397-0.408)	0.637 (0.629-0.645)
	RVD	1	0.537 (0.502-0.573)	0.133 (0.127-0.138)	0.545 (0.532-0.558)
		2	0.518 (0.483-0.550)	0.140 (0.135-0.146)	0.560 (0.546-0.574)
		4	0.571 (0.546-0.596)	0.149 (0.143-0.154)	0.583 (0.570-0.596)
		8	0.581 (0.563-0.599)	0.159 (0.154-0.165)	0.607 (0.595-0.619)
		16	0.603 (0.589-0.618)	0.171 (0.166-0.176)	0.635 (0.624-0.646)
		32	0.618 (0.606-0.629)	0.186 (0.182-0.190)	0.670 (0.663-0.677)
	LVH	1	0.481 (0.453-0.509)	0.308 (0.294-0.319)	0.525 (0.515-0.535)
		2	0.509 (0.489-0.529)	0.320 (0.312-0.329)	0.539 (0.529-0.548)
		4	0.529 (0.512-0.545)	0.332 (0.322-0.343)	0.557 (0.547-0.568)
		8	0.548 (0.536-0.561)	0.347 (0.341-0.354)	0.575 (0.566-0.584)
		16	0.563 (0.554-0.572)	0.365 (0.359-0.370)	0.598 (0.590-0.606)
		32	0.585 (0.577-0.594)	0.380 (0.375-0.386)	0.621 (0.614-0.628)
	LGE	1	0.512 (0.500-0.525)	0.546 (0.511-0.575)	0.508 (0.502-0.514)
		2	0.505 (0.494-0.515)	0.535 (0.509-0.561)	0.513 (0.508-0.518)
		4	0.515 (0.507-0.524)	0.564 (0.544-0.584)	0.520 (0.514-0.526)
		8	0.516 (0.510-0.522)	0.571 (0.559-0.582)	0.524 (0.518-0.530)
		16	0.524 (0.518-0.530)	0.580 (0.569-0.591)	0.533 (0.528-0.539)
		32	0.536 (0.530-0.541)	0.595 (0.587-0.603)	0.549 (0.542-0.555)
	AoD	1	0.532 (0.493-0.571)	0.161 (0.155-0.166)	0.516 (0.508-0.524)
		2	0.527 (0.496-0.556)	0.172 (0.167-0.176)	0.524 (0.517-0.531)
		4	0.524 (0.501-0.547)	0.178 (0.174-0.181)	0.534 (0.528-0.540)
		8	0.545 (0.529-0.562)	0.183 (0.180-0.186)	0.543 (0.537-0.550)
		16	0.550 (0.536-0.563)	0.191 (0.188-0.195)	0.562 (0.555-0.569)
		32	0.566 (0.556-0.576)	0.197 (0.194-0.200)	0.576 (0.571-0.582)
BiomedCLIP	LVDys	1	0.527 (0.504-0.550)	0.338 (0.320-0.354)	0.524 (0.512-0.537)
		2	0.560 (0.539-0.580)	0.347 (0.331-0.362)	0.549 (0.538-0.562)
		4	0.562 (0.546-0.576)	0.386 (0.375-0.397)	0.570 (0.557-0.583)
		8	0.565 (0.553-0.577)	0.406 (0.397-0.414)	0.586 (0.575-0.598)
		16	0.588 (0.579-0.596)	0.432 (0.424-0.439)	0.621 (0.613-0.629)
		32	0.602 (0.595-0.608)	0.457 (0.451-0.462)	0.649 (0.643-0.655)
	RVDys	1	0.566 (0.532-0.601)	0.192 (0.184-0.200)	0.536 (0.523-0.551)
		2	0.560 (0.529-0.589)	0.207 (0.200-0.213)	0.558 (0.546-0.569)
		4	0.573 (0.550-0.597)	0.220 (0.213-0.225)	0.577 (0.566-0.587)
		8	0.562 (0.546-0.579)	0.227 (0.221-0.232)	0.592 (0.582-0.602)
		16	0.582 (0.569-0.593)	0.248 (0.244-0.253)	0.629 (0.620-0.637)
		32	0.615 (0.606-0.624)	0.268 (0.263-0.272)	0.663 (0.657-0.669)
	LVD	1	0.511 (0.485-0.538)	0.319 (0.301-0.336)	0.544 (0.531-0.559)
		2	0.517 (0.494-0.539)	0.333 (0.320-0.345)	0.547 (0.533-0.560)
		4	0.561 (0.544-0.577)	0.359 (0.348-0.370)	0.585 (0.572-0.598)
		8	0.567 (0.555-0.580)	0.378 (0.369-0.388)	0.608 (0.597-0.618)
		16	0.595 (0.585-0.605)	0.409 (0.401-0.417)	0.646 (0.636-0.657)
		32	0.612 (0.606-0.619)	0.430 (0.425-0.435)	0.672 (0.665-0.678)
	RVD	1	0.570 (0.532-0.608)	0.131 (0.126-0.137)	0.541 (0.529-0.553)
		2	0.523 (0.488-0.559)	0.135 (0.131-0.140)	0.549 (0.537-0.561)
		4	0.556 (0.531-0.580)	0.145 (0.141-0.150)	0.574 (0.561-0.586)
		8	0.573 (0.555-0.592)	0.153 (0.148-0.158)	0.590 (0.578-0.601)
		16	0.593 (0.578-0.608)	0.166 (0.162-0.171)	0.626 (0.615-0.636)
		32	0.608 (0.598-0.618)	0.180 (0.176-0.184)	0.661 (0.653-0.668)
	LVH	1	0.466 (0.441-0.492)	0.308 (0.297-0.319)	0.518 (0.511-0.525)
		2	0.497 (0.475-0.519)	0.315 (0.305-0.324)	0.528 (0.521-0.536)
		4	0.510 (0.494-0.527)	0.322 (0.315-0.330)	0.538 (0.531-0.545)
		8	0.529 (0.516-0.542)	0.331 (0.324-0.338)	0.548 (0.541-0.556)
		16	0.544 (0.535-0.552)	0.347 (0.341-0.352)	0.571 (0.564-0.578)
		32	0.564 (0.557-0.571)	0.363 (0.358-0.367)	0.595 (0.589-0.601)
	LGE	1	0.509 (0.495-0.522)	0.533 (0.498-0.566)	0.514 (0.505-0.522)
		2	0.518 (0.506-0.529)	0.554 (0.526-0.581)	0.525 (0.515-0.534)
		4	0.525 (0.516-0.535)	0.571 (0.554-0.589)	0.538 (0.530-0.547)
		8	0.530 (0.522-0.537)	0.577 (0.564-0.590)	0.549 (0.541-0.555)
		16	0.543 (0.537-0.549)	0.600 (0.590-0.609)	0.559 (0.553-0.565)
		32	0.551 (0.546-0.556)	0.607 (0.601-0.613)	0.575 (0.568-0.581)
	AoD	1	0.501 (0.462-0.539)	0.168 (0.161-0.174)	0.525 (0.515-0.535)
		2	0.532 (0.499-0.566)	0.172 (0.166-0.179)	0.532 (0.522-0.542)
		4	0.533 (0.510-0.555)	0.183 (0.178-0.188)	0.547 (0.537-0.557)
		8	0.539 (0.521-0.556)	0.191 (0.186-0.195)	0.561 (0.551-0.570)

		16	0.563 (0.552-0.576)	0.202 (0.197-0.206)	0.583 (0.574-0.592)
		32	0.577 (0.568-0.586)	0.212 (0.208-0.215)	0.602 (0.595-0.608)
VideoMAE-CLIP	LVDys	1	0.644 (0.614-0.673)	0.528 (0.507-0.548)	0.732 (0.703-0.758)
		2	0.694 (0.665-0.724)	0.591 (0.573-0.610)	0.819 (0.800-0.836)
		4	0.777 (0.759-0.793)	0.662 (0.647-0.676)	0.870 (0.860-0.880)
		8	0.827 (0.818-0.837)	0.715 (0.704-0.725)	0.899 (0.892-0.906)
		16	0.857 (0.852-0.862)	0.754 (0.748-0.759)	0.921 (0.918-0.924)
		32	0.866 (0.863-0.869)	0.768 (0.764-0.772)	0.929 (0.927-0.931)
	RVDys	1	0.640 (0.599-0.678)	0.307 (0.286-0.327)	0.668 (0.639-0.694)
		2	0.642 (0.608-0.675)	0.318 (0.301-0.336)	0.714 (0.690-0.734)
		4	0.697 (0.673-0.722)	0.353 (0.338-0.369)	0.758 (0.742-0.773)
		8	0.749 (0.732-0.764)	0.396 (0.384-0.406)	0.796 (0.786-0.806)
		16	0.788 (0.780-0.796)	0.429 (0.424-0.435)	0.821 (0.818-0.825)
		32	0.797 (0.792-0.802)	0.438 (0.434-0.442)	0.833 (0.831-0.836)
	LVD	1	0.623 (0.590-0.656)	0.448 (0.423-0.471)	0.701 (0.672-0.728)
		2	0.652 (0.626-0.674)	0.487 (0.469-0.504)	0.759 (0.738-0.779)
		4	0.746 (0.726-0.764)	0.567 (0.551-0.582)	0.816 (0.802-0.828)
		8	0.782 (0.770-0.793)	0.615 (0.605-0.624)	0.853 (0.846-0.859)
		16	0.808 (0.801-0.815)	0.648 (0.641-0.655)	0.875 (0.871-0.880)
		32	0.831 (0.827-0.835)	0.671 (0.667-0.676)	0.888 (0.885-0.891)
	RVD	1	0.571 (0.530-0.609)	0.178 (0.166-0.189)	0.618 (0.591-0.639)
		2	0.601 (0.565-0.642)	0.181 (0.170-0.194)	0.637 (0.614-0.660)
		4	0.650 (0.622-0.678)	0.206 (0.196-0.216)	0.688 (0.670-0.703)
		8	0.690 (0.670-0.710)	0.232 (0.224-0.241)	0.734 (0.724-0.745)
		16	0.725 (0.710-0.736)	0.255 (0.248-0.262)	0.766 (0.757-0.775)
		32	0.751 (0.743-0.758)	0.274 (0.270-0.278)	0.787 (0.783-0.791)
	LVH	1	0.655 (0.622-0.688)	0.475 (0.448-0.502)	0.743 (0.716-0.769)
		2	0.715 (0.689-0.741)	0.561 (0.538-0.582)	0.824 (0.805-0.842)
		4	0.777 (0.755-0.799)	0.622 (0.604-0.642)	0.868 (0.855-0.880)
		8	0.837 (0.827-0.847)	0.690 (0.678-0.701)	0.903 (0.898-0.908)
		16	0.863 (0.857-0.869)	0.725 (0.718-0.732)	0.917 (0.915-0.920)
		32	0.881 (0.878-0.883)	0.750 (0.747-0.753)	0.926 (0.925-0.927)
	LGE	1	0.553 (0.533-0.574)	0.593 (0.558-0.630)	0.571 (0.547-0.593)
		2	0.570 (0.550-0.589)	0.600 (0.565-0.631)	0.605 (0.584-0.626)
		4	0.608 (0.595-0.621)	0.655 (0.635-0.674)	0.654 (0.638-0.669)
		8	0.638 (0.627-0.648)	0.685 (0.672-0.698)	0.689 (0.676-0.700)
		16	0.657 (0.648-0.666)	0.710 (0.701-0.719)	0.707 (0.697-0.717)
		32	0.682 (0.677-0.686)	0.729 (0.724-0.734)	0.740 (0.736-0.745)
	AoD	1	0.452 (0.417-0.486)	0.181 (0.176-0.186)	0.547 (0.536-0.559)
		2	0.516 (0.485-0.544)	0.183 (0.177-0.190)	0.554 (0.542-0.567)
		4	0.534 (0.510-0.555)	0.195 (0.190-0.199)	0.580 (0.570-0.590)
		8	0.549 (0.530-0.568)	0.205 (0.199-0.211)	0.598 (0.587-0.608)
		16	0.597 (0.585-0.608)	0.227 (0.222-0.232)	0.637 (0.629-0.645)
		32	0.607 (0.597-0.616)	0.235 (0.230-0.240)	0.650 (0.643-0.657)
CMR-CLIP	LVDys	1	0.626 (0.604-0.646)	0.505 (0.485-0.523)	0.714 (0.692-0.736)
		2	0.673 (0.648-0.697)	0.553 (0.537-0.571)	0.782 (0.765-0.799)
		4	0.755 (0.739-0.772)	0.635 (0.616-0.652)	0.844 (0.829-0.857)
		8	0.807 (0.797-0.819)	0.697 (0.687-0.709)	0.887 (0.881-0.893)
		16	0.851 (0.845-0.857)	0.753 (0.747-0.760)	0.919 (0.916-0.922)
		32	0.872 (0.868-0.876)	0.782 (0.777-0.787)	0.934 (0.931-0.936)
	RVDys	1	0.576 (0.551-0.601)	0.258 (0.247-0.271)	0.639 (0.617-0.660)
		2	0.572 (0.541-0.602)	0.273 (0.262-0.284)	0.675 (0.658-0.690)
		4	0.653 (0.631-0.675)	0.311 (0.301-0.322)	0.719 (0.705-0.732)
		8	0.701 (0.685-0.714)	0.352 (0.342-0.361)	0.765 (0.754-0.776)
		16	0.751 (0.742-0.760)	0.393 (0.387-0.400)	0.799 (0.795-0.804)
		32	0.767 (0.760-0.773)	0.414 (0.410-0.418)	0.821 (0.818-0.824)
	LVD	1	0.584 (0.564-0.602)	0.443 (0.427-0.460)	0.698 (0.674-0.719)
		2	0.664 (0.640-0.688)	0.492 (0.473-0.511)	0.761 (0.741-0.780)
		4	0.741 (0.724-0.757)	0.569 (0.552-0.585)	0.818 (0.805-0.832)
		8	0.790 (0.777-0.802)	0.638 (0.627-0.649)	0.868 (0.860-0.875)
		16	0.829 (0.823-0.835)	0.686 (0.679-0.693)	0.897 (0.892-0.900)
		32	0.852 (0.848-0.856)	0.718 (0.714-0.723)	0.911 (0.908-0.913)
	RVD	1	0.572 (0.548-0.594)	0.175 (0.167-0.184)	0.642 (0.621-0.662)
		2	0.612 (0.581-0.645)	0.196 (0.186-0.207)	0.684 (0.667-0.701)
		4	0.663 (0.640-0.687)	0.220 (0.210-0.230)	0.722 (0.707-0.737)
		8	0.721 (0.704-0.735)	0.255 (0.248-0.263)	0.774 (0.766-0.783)

		16	0.750 (0.739-0.760)	0.287 (0.280-0.295)	0.813 (0.806-0.820)
		32	0.776 (0.769-0.783)	0.313 (0.308-0.319)	0.837 (0.833-0.840)
	LVH	1	0.627 (0.603-0.651)	0.456 (0.438-0.473)	0.720 (0.697-0.742)
		2	0.689 (0.664-0.714)	0.525 (0.507-0.542)	0.797 (0.781-0.812)
		4	0.752 (0.731-0.772)	0.590 (0.573-0.606)	0.847 (0.834-0.859)
		8	0.814 (0.803-0.824)	0.661 (0.651-0.672)	0.892 (0.887-0.898)
		16	0.846 (0.840-0.853)	0.702 (0.695-0.709)	0.909 (0.906-0.911)
		32	0.871 (0.867-0.874)	0.737 (0.732-0.742)	0.921 (0.920-0.923)
	LGE	1	0.561 (0.545-0.578)	0.603 (0.580-0.627)	0.589 (0.571-0.609)
		2	0.581 (0.565-0.598)	0.614 (0.586-0.643)	0.621 (0.603-0.639)
		4	0.618 (0.605-0.630)	0.662 (0.644-0.680)	0.668 (0.655-0.680)
		8	0.646 (0.637-0.655)	0.690 (0.678-0.702)	0.704 (0.695-0.713)
		16	0.677 (0.669-0.684)	0.725 (0.717-0.733)	0.732 (0.724-0.738)
		32	0.699 (0.694-0.703)	0.745 (0.739-0.750)	0.760 (0.756-0.764)
	AoD	1	0.510 (0.490-0.531)	0.197 (0.190-0.203)	0.574 (0.559-0.588)
		2	0.535 (0.506-0.563)	0.203 (0.196-0.209)	0.591 (0.575-0.606)
		4	0.559 (0.535-0.581)	0.215 (0.209-0.222)	0.617 (0.605-0.629)
		8	0.602 (0.586-0.621)	0.239 (0.233-0.245)	0.657 (0.648-0.668)
		16	0.646 (0.634-0.657)	0.272 (0.267-0.276)	0.707 (0.700-0.714)
		32	0.667 (0.658-0.675)	0.286 (0.282-0.290)	0.729 (0.724-0.735)

Supplemental Table 5: Few-shot classification results with the internal cohort for four contrastive vision–language models (OPENAI CLIP, BiomedCLIP, VideoMAE-CLIP and CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Model	Label	Shot	Accuracy	F1-score	AUC
OPENAI CLIP	LVDys	1	0.509 (0.492-0.526)	0.336 (0.310-0.361)	0.493 (0.489-0.496)
		2	0.501 (0.480-0.522)	0.346 (0.315-0.375)	0.501 (0.497-0.505)
		4	0.518 (0.498-0.539)	0.317 (0.284-0.347)	0.498 (0.494-0.502)
		8	0.498 (0.479-0.517)	0.357 (0.328-0.384)	0.500 (0.496-0.504)
		16	0.483 (0.465-0.502)	0.378 (0.350-0.403)	0.501 (0.498-0.505)
		32	0.473 (0.456-0.492)	0.396 (0.374-0.419)	0.499 (0.496-0.503)
	RVDys	1	0.537 (0.498-0.576)	0.202 (0.188-0.217)	0.499 (0.493-0.503)
		2	0.486 (0.448-0.527)	0.213 (0.198-0.226)	0.498 (0.492-0.504)
		4	0.500 (0.459-0.539)	0.214 (0.202-0.226)	0.495 (0.489-0.499)
		8	0.502 (0.463-0.541)	0.215 (0.203-0.227)	0.494 (0.490-0.498)
		16	0.484 (0.448-0.525)	0.220 (0.208-0.229)	0.491 (0.486-0.496)
		32	0.453 (0.417-0.491)	0.224 (0.214-0.233)	0.491 (0.486-0.496)
	LVD	1	0.537 (0.502-0.572)	0.231 (0.220-0.242)	0.503 (0.495-0.512)
		2	0.530 (0.487-0.571)	0.219 (0.203-0.234)	0.506 (0.499-0.514)
		4	0.529 (0.489-0.569)	0.229 (0.216-0.241)	0.502 (0.494-0.510)
		8	0.514 (0.477-0.552)	0.240 (0.227-0.251)	0.511 (0.503-0.518)
		16	0.557 (0.521-0.593)	0.239 (0.226-0.251)	0.517 (0.510-0.524)
		32	0.611 (0.575-0.644)	0.225 (0.214-0.236)	0.520 (0.513-0.526)
	RVD	1	0.568 (0.526-0.609)	0.152 (0.143-0.160)	0.500 (0.492-0.508)
		2	0.533 (0.486-0.580)	0.155 (0.145-0.164)	0.510 (0.503-0.517)
		4	0.562 (0.512-0.612)	0.149 (0.138-0.159)	0.512 (0.504-0.520)
		8	0.580 (0.532-0.627)	0.150 (0.140-0.158)	0.509 (0.506-0.517)
		16	0.627 (0.585-0.668)	0.151 (0.142-0.160)	0.513 (0.506-0.520)
		32	0.650 (0.612-0.688)	0.148 (0.139-0.156)	0.511 (0.504-0.518)
	LVH	1	0.483 (0.444-0.521)	0.231 (0.220-0.240)	0.499 (0.492-0.503)
		2	0.503 (0.453-0.548)	0.218 (0.204-0.229)	0.497 (0.490-0.504)
		4	0.496 (0.453-0.538)	0.215 (0.204-0.227)	0.491 (0.485-0.497)
		8	0.497 (0.459-0.536)	0.224 (0.215-0.233)	0.488 (0.482-0.493)
		16	0.525 (0.488-0.562)	0.224 (0.217-0.231)	0.491 (0.486-0.498)
		32	0.527 (0.493-0.562)	0.222 (0.214-0.229)	0.487 (0.482-0.492)
	LGE	1	0.495 (0.477-0.515)	0.505 (0.459-0.550)	0.498 (0.494-0.502)
		2	0.489 (0.469-0.511)	0.477 (0.429-0.527)	0.502 (0.499-0.505)
		4	0.501 (0.483-0.519)	0.520 (0.479-0.560)	0.505 (0.501-0.508)
		8	0.500 (0.483-0.519)	0.516 (0.477-0.557)	0.505 (0.502-0.509)
		16	0.500 (0.480-0.519)	0.508 (0.463-0.552)	0.508 (0.504-0.511)

		32	0.505 (0.487-0.524)	0.520 (0.482-0.562)	0.511 (0.508-0.514)
	AoD	1	0.522 (0.479-0.567)	0.158 (0.149-0.167)	0.495 (0.487-0.503)
		2	0.516 (0.471-0.562)	0.150 (0.138-0.162)	0.500 (0.492-0.508)
		4	0.521 (0.470-0.572)	0.150 (0.138-0.161)	0.497 (0.488-0.505)
		8	0.499 (0.455-0.546)	0.157 (0.148-0.166)	0.503 (0.496-0.511)
		16	0.507 (0.462-0.554)	0.156 (0.145-0.166)	0.505 (0.498-0.511)
		32	0.531 (0.486-0.575)	0.159 (0.149-0.168)	0.507 (0.500-0.513)
BiomedCLIP	LVDys	1	0.500 (0.482-0.518)	0.356 (0.331-0.384)	0.495 (0.491-0.500)
		2	0.517 (0.498-0.536)	0.339 (0.308-0.365)	0.498 (0.493-0.503)
		4	0.514 (0.496-0.532)	0.347 (0.323-0.371)	0.495 (0.491-0.499)
		8	0.507 (0.490-0.523)	0.360 (0.340-0.381)	0.497 (0.492-0.501)
		16	0.504 (0.490-0.517)	0.375 (0.358-0.390)	0.496 (0.492-0.499)
		32	0.505 (0.492-0.518)	0.376 (0.363-0.389)	0.494 (0.491-0.497)
	RVDys	1	0.548 (0.512-0.584)	0.204 (0.189-0.218)	0.496 (0.489-0.504)
		2	0.532 (0.496-0.566)	0.214 (0.201-0.225)	0.491 (0.485-0.498)
		4	0.536 (0.500-0.573)	0.195 (0.183-0.207)	0.481 (0.474-0.487)
		8	0.516 (0.482-0.552)	0.206 (0.193-0.216)	0.479 (0.473-0.486)
		16	0.524 (0.492-0.552)	0.206 (0.196-0.215)	0.477 (0.471-0.482)
		32	0.538 (0.513-0.563)	0.199 (0.191-0.207)	0.466 (0.462-0.471)
	LVD	1	0.501 (0.465-0.539)	0.234 (0.223-0.246)	0.500 (0.496-0.504)
		2	0.504 (0.471-0.538)	0.237 (0.224-0.248)	0.504 (0.500-0.508)
		4	0.542 (0.51-0.577)	0.234 (0.223-0.245)	0.506 (0.502-0.509)
		8	0.527 (0.496-0.558)	0.235 (0.224-0.245)	0.506 (0.502-0.510)
		16	0.537 (0.510-0.565)	0.242 (0.232-0.250)	0.507 (0.504-0.510)
		32	0.570 (0.547-0.592)	0.236 (0.229-0.244)	0.507 (0.505-0.510)
	RVD	1	0.563 (0.523-0.603)	0.153 (0.144-0.162)	0.503 (0.495-0.511)
		2	0.531 (0.489-0.572)	0.157 (0.148-0.166)	0.503 (0.495-0.511)
		4	0.517 (0.477-0.558)	0.163 (0.154-0.170)	0.510 (0.502-0.518)
		8	0.490 (0.451-0.528)	0.165 (0.158-0.171)	0.501 (0.495-0.508)
		16	0.518 (0.486-0.554)	0.165 (0.159-0.170)	0.506 (0.500-0.512)
		32	0.485 (0.448-0.522)	0.167 (0.162-0.171)	0.506 (0.500-0.512)
	LVH	1	0.485 (0.450-0.522)	0.213 (0.198-0.227)	0.489 (0.482-0.496)
		2	0.491 (0.458-0.524)	0.217 (0.204-0.228)	0.483 (0.476-0.491)
		4	0.511 (0.477-0.547)	0.210 (0.197-0.222)	0.484 (0.476-0.491)
		8	0.540 (0.507-0.575)	0.201 (0.187-0.213)	0.484 (0.476-0.492)
		16	0.560 (0.531-0.589)	0.191 (0.178-0.203)	0.479 (0.471-0.486)
		32	0.596 (0.571-0.621)	0.176 (0.164-0.188)	0.470 (0.464-0.475)
	LGE	1	0.503 (0.487-0.522)	0.536 (0.499-0.574)	0.497 (0.491-0.504)
		2	0.510 (0.492-0.526)	0.540 (0.500-0.581)	0.510 (0.504-0.516)
		4	0.509 (0.491-0.527)	0.542 (0.501-0.582)	0.510 (0.503-0.516)
		8	0.507 (0.489-0.524)	0.538 (0.498-0.575)	0.510 (0.505-0.516)
		16	0.522 (0.506-0.537)	0.575 (0.541-0.607)	0.514 (0.508-0.519)
		32	0.522 (0.508-0.536)	0.576 (0.545-0.607)	0.516 (0.511-0.522)
	AoD	1	0.498 (0.454-0.540)	0.157 (0.148-0.166)	0.495 (0.487-0.503)
		2	0.540 (0.494-0.585)	0.147 (0.137-0.157)	0.496 (0.488-0.503)
		4	0.535 (0.493-0.579)	0.157 (0.147-0.165)	0.496 (0.489-0.504)
		8	0.565 (0.529-0.599)	0.153 (0.145-0.162)	0.496 (0.489-0.504)
		16	0.623 (0.588-0.657)	0.144 (0.135-0.153)	0.500 (0.493-0.507)
		32	0.665 (0.633-0.694)	0.128 (0.119-0.137)	0.501 (0.494-0.508)
VideoMAE-CLIP	LVDys	1	0.591 (0.568-0.615)	0.511 (0.488-0.531)	0.682 (0.659-0.705)
		2	0.631 (0.604-0.658)	0.547 (0.527-0.566)	0.755 (0.737-0.772)
		4	0.686 (0.667-0.703)	0.587 (0.573-0.601)	0.789 (0.777-0.800)
		8	0.728 (0.715-0.739)	0.631 (0.619-0.642)	0.813 (0.806-0.821)
		16	0.739 (0.731-0.746)	0.647 (0.641-0.653)	0.822 (0.818-0.827)
		32	0.746 (0.738-0.753)	0.659 (0.654-0.663)	0.830 (0.827-0.833)
	RVDys	1	0.551 (0.515-0.587)	0.265 (0.250-0.278)	0.590 (0.573-0.604)
		2	0.561 (0.528-0.593)	0.281 (0.267-0.293)	0.621 (0.608-0.634)
		4	0.568 (0.537-0.600)	0.294 (0.281-0.306)	0.635 (0.623-0.647)
		8	0.593 (0.567-0.617)	0.317 (0.309-0.324)	0.653 (0.645-0.661)
		16	0.630 (0.615-0.644)	0.333 (0.327-0.339)	0.662 (0.656-0.668)
		32	0.636 (0.624-0.646)	0.339 (0.335-0.342)	0.665 (0.661-0.669)
	LVD	1	0.533 (0.493-0.570)	0.272 (0.256-0.287)	0.579 (0.564-0.593)
		2	0.535 (0.501-0.566)	0.301 (0.289-0.313)	0.611 (0.599-0.624)
		4	0.580 (0.557-0.603)	0.312 (0.303-0.320)	0.616 (0.606-0.625)
		8	0.579 (0.557-0.599)	0.330 (0.324-0.336)	0.624 (0.617-0.631)
		16	0.588 (0.570-0.603)	0.347 (0.342-0.352)	0.635(0.630-0.640)

		32	0.615 (0.602-0.627)	0.352 (0.348-0.356)	0.633 (0.629-0.637)
	RVD	1	0.479 (0.440-0.520)	0.169 (0.157-0.182)	0.542 (0.522-0.562)
		2	0.475 (0.433-0.518)	0.171 (0.159-0.182)	0.537 (0.518-0.555)
		4	0.484 (0.450-0.521)	0.183 (0.171-0.193)	0.554 (0.534-0.572)
		8	0.470 (0.440-0.500)	0.185 (0.177-0.193)	0.549 (0.534-0.565)
		16	0.471 (0.450-0.490)	0.193 (0.186-0.199)	0.555 (0.543-0.569)
		32	0.471 (0.457-0.484)	0.192 (0.188-0.197)	0.548 (0.539-0.557)
	LVH	1	0.614 (0.572-0.654)	0.360 (0.339-0.381)	0.706 (0.682-0.729)
		2	0.638 (0.603-0.674)	0.405 (0.387-0.422)	0.768 (0.753-0.782)
		4	0.696 (0.666-0.726)	0.444 (0.428-0.459)	0.800 (0.790-0.809)
		8	0.749 (0.731-0.765)	0.485 (0.476-0.494)	0.819 (0.815-0.823)
		16	0.777 (0.767-0.788)	0.505 (0.499-0.511)	0.827 (0.825-0.829)
		32	0.788 (0.782-0.794)	0.516 (0.512-0.520)	0.832 (0.831-0.834)
	LGE	1	0.540 (0.518-0.563)	0.568 (0.526-0.612)	0.553 (0.535-0.571)
		2	0.529 (0.506-0.551)	0.538 (0.493-0.580)	0.571 (0.553-0.588)
		4	0.552 (0.534-0.570)	0.582 (0.548-0.614)	0.608 (0.596-0.620)
		8	0.555 (0.541-0.571)	0.583 (0.555-0.610)	0.633 (0.623-0.642)
		16	0.571 (0.559-0.584)	0.610 (0.588-0.632)	0.643 (0.634-0.651)
		32	0.573 (0.563-0.584)	0.604 (0.586-0.623)	0.665 (0.660-0.669)
	AoD	1	0.402 (0.363-0.444)	0.173 (0.165-0.179)	0.510 (0.497-0.522)
		2	0.433 (0.391-0.475)	0.172 (0.164-0.180)	0.506 (0.494-0.518)
		4	0.463 (0.424-0.501)	0.175 (0.168-0.183)	0.515 (0.504-0.528)
		8	0.436 (0.401-0.470)	0.183 (0.177-0.189)	0.522 (0.511-0.533)
		16	0.485 (0.453-0.516)	0.187 (0.181-0.193)	0.527 (0.517-0.537)
		32	0.480 (0.454-0.506)	0.186 (0.181-0.191)	0.528 (0.519-0.537)
CMR-CLIP	LVDys	1	0.630 (0.606-0.654)	0.528 (0.507-0.549)	0.688 (0.663-0.712)
		2	0.672 (0.649-0.694)	0.560 (0.539-0.579)	0.755 (0.736-0.773)
		4	0.728 (0.712-0.742)	0.596 (0.579-0.612)	0.793 (0.779-0.806)
		8	0.771 (0.762-0.780)	0.640 (0.629-0.652)	0.826 (0.819-0.834)
		16	0.792 (0.787-0.798)	0.660 (0.653-0.668)	0.843 (0.837-0.848)
		32	0.806 (0.802-0.809)	0.677 (0.671-0.683)	0.858 (0.855-0.862)
	RVDys	1	0.553 (0.526-0.579)	0.298 (0.287-0.310)	0.619 (0.599-0.636)
		2	0.545 (0.529-0.579)	0.302 (0.291-0.312)	0.644 (0.628-0.659)
		4	0.598 (0.569-0.629)	0.321 (0.309-0.333)	0.675 (0.661-0.688)
		8	0.636 (0.613-0.657)	0.358 (0.348-0.368)	0.712 (0.701-0.723)
		16	0.680 (0.667-0.692)	0.377 (0.369-0.386)	0.733 (0.726-0.739)
		32	0.691 (0.682-0.699)	0.399 (0.392-0.406)	0.746 (0.741-0.750)
	LVD	1	0.581 (0.557-0.605)	0.310 (0.296-0.324)	0.603 (0.584-0.622)
		2	0.650 (0.620-0.676)	0.328 (0.312-0.342)	0.641 (0.625-0.657)
		4	0.697 (0.678-0.715)	0.336 (0.321-0.351)	0.644 (0.629-0.658)
		8	0.719 (0.702-0.735)	0.364 (0.356-0.373)	0.656 (0.645-0.667)
		16	0.747 (0.739-0.753)	0.386 (0.379-0.393)	0.669 (0.662-0.677)
		32	0.758 (0.754-0.762)	0.390 (0.386-0.395)	0.673 (0.667-0.679)
	RVD	1	0.527 (0.501-0.552)	0.210 (0.197-0.221)	0.620 (0.595-0.645)
		2	0.552 (0.519-0.588)	0.217 (0.204-0.228)	0.645 (0.624-0.664)
		4	0.614 (0.587-0.640)	0.238 (0.224-0.250)	0.675 (0.657-0.691)
		8	0.646 (0.625-0.667)	0.249 (0.237-0.263)	0.699 (0.685-0.713)
		16	0.657 (0.639-0.675)	0.279 (0.269-0.288)	0.733 (0.723-0.745)
		32	0.678 (0.668-0.689)	0.303 (0.296-0.309)	0.743 (0.735-0.750)
	LVH	1	0.622 (0.593-0.652)	0.363 (0.347-0.379)	0.694 (0.673-0.713)
		2	0.662 (0.633-0.693)	0.399 (0.379-0.416)	0.757 (0.742-0.772)
		4	0.727 (0.702-0.750)	0.459 (0.445-0.473)	0.797 (0.787-0.807)
		8	0.769 (0.756-0.781)	0.495 (0.485-0.505)	0.826 (0.821-0.831)
		16	0.801 (0.794-0.808)	0.527 (0.520-0.533)	0.838 (0.834-0.841)
		32	0.816 (0.811-0.820)	0.548 (0.542-0.553)	0.847 (0.845-0.849)
	LGE	1	0.576 (0.557-0.596)	0.642 (0.613-0.669)	0.582 (0.561-0.603)
		2	0.594 (0.570-0.617)	0.635 (0.594-0.677)	0.619 (0.601-0.637)
		4	0.631 (0.613-0.647)	0.696 (0.669-0.720)	0.657 (0.644-0.670)
		8	0.650 (0.638-0.662)	0.719 (0.700-0.736)	0.680 (0.671-0.691)
		16	0.679 (0.669-0.689)	0.750 (0.737-0.761)	0.708 (0.700-0.715)
		32	0.695 (0.688-0.703)	0.765 (0.756-0.774)	0.726 (0.720-0.731)
	AoD	1	0.486 (0.461-0.514)	0.190 (0.182-0.197)	0.538 (0.526-0.550)
		2	0.479 (0.442-0.514)	0.186 (0.178-0.194)	0.534 (0.521-0.548)
		4	0.487 (0.449-0.522)	0.190 (0.184-0.196)	0.549 (0.538-0.560)
		8	0.505 (0.476-0.532)	0.205 (0.199-0.210)	0.575 (0.565-0.585)
		16	0.539 (0.515-0.562)	0.216 (0.211-0.222)	0.594 (0.586-0.602)

		32	0.560 (0.539-0.584)	0.223 (0.218-0.228)	0.601 (0.594-0.609)

Supplemental Table 6: Few-shot classification results with the external cohort for four contrastive vision–language models (OPENAI CLIP, BiomedCLIP, VideoMAE-CLIP and CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Model	Text-to-Video Retrieval			Video-to-Text Retrieval			RSUM
	R@5	R@10	R@50	R@5	R@10	R@50	
OPENAI CLIP	0.2	0.4	1.9	0.2	0.4	1.8	4.9
BiomedCLIP	0.3	0.5	2.1	0.4	0.5	2.0	5.8
VideoMAE-CLIP	11.7	19.1	49.1	11.7	18.3	48.4	158.3
CMR-CLIP $_{\text{Cine}}$	18.2	26.8	50.9	18.4	26.5	51.1	191.9
CMR-CLIP $_{\text{LGE}}$	18.0	26.3	53.3	17.8	26.7	53.1	195.1
CMR-CLIP $_{\text{Cine-LGE}}$ (1)	8.1	12.4	30.1	7.8	12.4	29.6	100.4
CMR-CLIP $_{\text{Cine-LGE}}$ (4)	15.1	22.9	48.0	15.3	22.6	47.5	171.4
CMR-CLIP $_{\text{Cine-LGE}}$ (8)	17.3	25.6	51.0	17.7	25.5	50.7	187.7
CMR-CLIP $_{\text{Cine-LGE}}$ (16)	19.5	29.2	54.8	20.3	29.2	55.1	208.1
CMR-CLIP $_{\text{Cine-LGE}}$ (32)	20.6	29.6	56.2	19.9	29.6	55.9	211.7
CMR-CLIP $_{\text{Cine-LGE}}$ (64)	20.7	30.9	63.4	20.2	30.2	61.9	227.3

Supplemental Table 7: Comparison of zero-shot retrieval performance (Recall@5, Recall@10, Recall@50) in both text-to-video and video-to-text directions, plus the cumulative RSUM score with the external cohort. CMR-CLIP sequences are labeled by their image type followed by the number of sampled frames in parentheses. For example, $_{\text{Cine-LGE}}$ (64) indicates that both the Cine and LGE sequences were acquired with 64 frames in total. Other CLIP models likewise use the $_{\text{Cine-LGE}}$ (64) input; the cine-only model was run with 36 frames, and the LGE-only model with 28 frames.

Model	Text-to-Video Retrieval			Video-to-Text Retrieval			RSUM
	R@5	R@10	R@50	R@5	R@10	R@50	
OPENAI CLIP	0.9	1.9	10.5	0.9	1.9	12.1	28.2
BiomedCLIP	0.7	2.6	13.3	0.9	2.6	13.3	33.4
VideoMAE-CLIP	4.7	9.6	37.9	6.5	11.0	32.5	102.2
CMR-CLIP	15.4	23.4	56.1	12.9	22.9	56.5	187.2

Supplemental Table 8: Comparison of zero-shot retrieval performance (Recall@5, Recall@10, Recall@50) in both text-to-video and video-to-text directions, plus the cumulative RSUM score with the external cohort.

Model	Label	Accuracy	F1-score	AUC
OPENAI CLIP	ACDC DCM	0.725 (0.575-0.850)	0.000 (0.000-0.000)	0.611 (0.413-0.818)
	ACDC HCM	0.725 (0.599-0.850)	0.154 (0.000-0.462)	0.730 (0.550-0.878)
	ACDC MI	0.725 (0.575-0.850)	0.000 (0.000-0.000)	0.510 (0.276-0.730)
	ACDC Normal	0.775 (0.625-0.900)	0.182 (0.000-0.501)	0.636 (0.404-0.843)
	ACDC Abn RV	0.825 (0.700-0.925)	0.364 (0.000-0.714)	0.652 (0.342-0.928)
	NICM	0.730 (0.692-0.765)	0.753 (0.713-0.790)	0.780 (0.741-0.815)
	ICM	0.768 (0.732-0.801)	0.426 (0.338-0.491)	0.717 (0.662-0.762)
	CA	0.930 (0.907-0.948)	0.349 (0.192-0.483)	0.768 (0.673-0.856)
	HCM	0.919 (0.897-0.940)	0.712 (0.628-0.781)	0.957 (0.933-0.975)
BiomedCLIP	ACDC DCM	0.725 (0.575-0.850)	0.000 (0.000-0.000)	0.803 (0.631-0.936)
	ACDC HCM	0.725 (0.575-0.850)	0.154 (0.000-0.429)	0.703 (0.507-0.858)
	ACDC MI	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.490 (0.191-0.776)
	ACDC Normal	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.662 (0.434-0.874)
	ACDC Abn RV	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.809 (0.647-0.934)

	NICM	0.711 (0.674-0.749)	0.736 (0.696-0.775)	0.773 (0.735-0.811)
	ICM	0.761 (0.725-0.796)	0.371 (0.281-0.441)	0.669 (0.612-0.718)
	CA	0.943 (0.923-0.962)	0.441 (0.274-0.600)	0.800 (0.713-0.881)
	HCM	0.926 (0.905-0.947)	0.743 (0.662-0.812)	0.963 (0.946-0.977)
Frozen in Time	ACDC DCM	0.725 (0.575-0.850)	0.000 (0.000-0.000)	0.589 (0.384-0.798)
	ACDC HCM	0.750 (0.624-0.875)	0.000 (0.000-0.000)	0.513 (0.305-0.713)
	ACDC MI	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.750 (0.567-0.901)
	ACDC Normal	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.563 (0.315-0.817)
	ACDC Abn RV	0.850 (0.725-0.950)	0.000 (0.000-0.000)	0.735 (0.486-0.974)
	NICM	0.687 (0.651-0.723)	0.723 (0.684-0.761)	0.757 (0.719-0.791)
	ICM	0.746 (0.710-0.780)	0.229 (0.154-0.308)	0.665 (0.613-0.717)
	CA	0.931 (0.909-0.950)	0.130 (0.000-0.275)	0.666 (0.567-0.756)
	HCM	0.907 (0.883-0.931)	0.667 (0.581-0.746)	0.948 (0.926-0.966)
VideoMAE-CLIP	ACDC DCM	0.825 (0.700-0.925)	0.632 (0.316-0.842)	0.865 (0.741-0.964)
	ACDC HCM	0.850 (0.725-0.950)	0.667 (0.375-0.889)	0.907 (0.793-0.994)
	ACDC MI	0.800 (0.675-0.925)	0.000 (0.000-0.000)	0.667 (0.430-0.879)
	ACDC Normal	0.825 (0.700-0.925)	0.000 (0.000-0.000)	0.814 (0.660-0.944)
	ACDC Abn RV	0.850 (0.725-0.950)	0.400 (0.000-0.727)	0.838 (0.684-0.972)
	NICM	0.754 (0.720-0.787)	0.786 (0.752-0.818)	0.834 (0.801-0.865)
	ICM	0.768 (0.734-0.801)	0.440 (0.362-0.509)	0.779 (0.735-0.820)
	CA	0.966 (0.950-0.979)	0.677 (0.519-0.793)	0.920 (0.857-0.964)
	HCM	0.988 (0.978-0.997)	0.961 (0.929-0.988)	0.999 (0.997-1.000)
Amyloid-Targeted	ACDC DCM	0.900 (0.800-0.975)	0.778 (0.500-0.952)	0.950 (0.857-1.000)
	ACDC HCM	0.925 (0.850-1.000)	0.824 (0.545-1.000)	0.973 (0.909-1.000)
	ACDC MI	0.950 (0.875-1.000)	0.833 (0.500-1.000)	0.961 (0.869-1.000)
	ACDC Normal	0.850 (0.725-0.950)	0.400 (0.000-0.750)	0.844 (0.697-0.96)
	ACDC Abn RV	0.875 (0.750-0.975)	0.444 (0.000-0.800)	0.843 (0.624-0.983)
	NICM	0.826 (0.797-0.856)	0.845 (0.814-0.874)	0.900 (0.874-0.925)
	ICM	0.832 (0.801-0.861)	0.566 (0.486-0.638)	0.848 (0.809-0.887)
	CA	0.954 (0.933-0.969)	0.597 (0.435-0.719)	0.930 (0.875-0.970)
	HCM	0.981 (0.969-0.991)	0.938 (0.898-0.971)	0.993 (0.984-0.999)
CMR-CLIP	ACDC DCM	0.950 (0.875-1.000)	0.900 (0.727-1.000)	0.934 (0.806-1.000)
	ACDC HCM	0.950 (0.875-1.000)	0.889 (0.667-1.000)	0.997 (0.983-1.000)
	ACDC MI	0.925 (0.825-1.000)	0.667 (0.000-1.000)	0.985 (0.939-1.000)
	ACDC Normal	0.950 (0.875-1.000)	0.833 (0.500-1.000)	0.987 (0.944-1.000)
	ACDC Abn RV	0.950 (0.875-1.000)	0.857 (0.571-1.000)	0.995 (0.972-1.000)
	NICM	0.885 (0.859-0.912)	0.896 (0.870-0.921)	0.917 (0.893-0.941)
	ICM	0.880 (0.852-0.905)	0.733 (0.675-0.791)	0.892 (0.858-0.925)
	CA	0.962 (0.947-0.976)	0.686 (0.541-0.800)	0.940 (0.884-0.980)
	HCM	0.986 (0.976-0.995)	0.956 (0.920-0.983)	0.998 (0.996-1.000)

Supplemental Table 9: Supervised classification results with the disease cohort for six vision encoders (OPENAI CLIP, BiomedCLIP, Frozen in Time, VideoMAE-CLIP, Amyloid-Targeted, CMR-CLIP) across seven cardiac findings. Metrics reported are accuracy, F1-score, and area under the curve (AUC). Values in parentheses represent 95% confidence intervals obtained via bootstrapping with 1,000 random resamples.

Diagnostic confidence	CA	HCM	Other
Incorrect diagnosis	0.500 (10/20)	1.000 (1/1)	1.000 (1/1)
No diagnosis	1.000 (2/2)	0.500 (1/2)	0.722 (13/18)
Low confidence	0.714 (10/14)	0.833 (5/6)	0.909 (10/11)
Moderate confidence	0.900 (18/20)	0.000 (0/1)	0.750 (6/8)
High confidence	0.976 (40/41)	0.933 (70/75)	1.000 (26/26)

Supplemental Table 10: Proportion of correct diagnoses (and raw counts) stratified by reported confidence level, for patients with cardiac amyloidosis, hypertrophic cardiomyopathy, or other infiltrative cardiomyopathies.

NLG metrics	Retrieval-based	CMR-TARGET
BLEU-1	0.295	0.461
BLEU-2	0.168	0.350
BLEU-3	0.106	0.283
BLEU-4	0.072	0.236
METEOR	0.124	0.217
ROUGE-L	0.215	0.390
CIDEr	0.031	0.258
BERTScore	0.629	0.734

Supplemental Table 11: Evaluation results for the Retrieval-based model versus our CMR-TARGET model, measured across standard NLG metrics (BLEU-1–4, METEOR, ROUGE-L, CIDEr, and BERTScore) on the image-based report generation task.

Label	Accuracy	F1-score
LVDys	0.892	0.799
RVDys	0.931	0.403
LVD	0.896	0.763
RVD	0.931	0.258
LVH	0.896	0.668
LGE	0.715	0.801
AoD	0.832	0.253

Supplemental Table 12: CMR-TARGET classification results. Entities were extracted from generated impressions to identify findings and to evaluate the overall concordance between generated and clinical reports.

Manufacturer	Label	Accuracy	F1-score	AUC
Philips (N=2,726)	LVDys	0.734	0.593	0.795
	RVDys	0.689	0.323	0.734
	LVD	0.734	0.554	0.805
	RVD	0.701	0.253	0.777
	LVH	0.771	0.601	0.850
	LGE	0.637	0.692	0.681
	AoD	0.657	0.272	0.706
Siemens (N=32)	LVDys	0.656	0.560	0.691
	RVDys	0.500	0.333	0.705
	LVD	0.813	0.625	0.926
	RVD	0.563	0.125	0.617
	LVH	0.688	0.545	0.783
	LGE	0.563	0.650	0.674
	AoD	0.750	0.500	0.911

Supplemental Table 13: Zero-shot classification performance of CMR-CLIP model for detecting cardiac abnormalities on Philips (N=2,726) and Siemens (N=32) MRI scans on the internal cohort.

Manufacturer	Label	Accuracy	F1-score	AUC
Siemens (N=428)	LVDys	0.715	0.647	0.799
	RVDys	0.355	0.320	0.726
	LVD	0.752	0.369	0.733
	RVD	0.453	0.259	0.809
	LVH	0.556	0.410	0.836
	LGE	0.645	0.722	0.665
	AoD	0.376	0.208	0.594

Supplemental Table 14: Zero-shot classification performance of CMR-CLIP model for detecting cardiac abnormalities on Siemens (N=428) MRI scans on the external cohort.

Magnetic Field Strength	Label	Accuracy	F1-score	AUC
1.5T (N=1,844)	LVDys	0.743	0.579	0.787
	RVDys	0.735	0.328	0.732
	LVD	0.729	0.572	0.806
	RVD	0.721	0.282	0.795
	LVH	0.783	0.600	0.851
	LGE	0.642	0.697	0.694
	AoD	0.675	0.277	0.701
3.0T (N=914)	LVDys	0.713	0.616	0.811
	RVDys	0.590	0.317	0.751
	LVD	0.748	0.513	0.805
	RVD	0.655	0.194	0.732
	LVH	0.745	0.599	0.854
	LGE	0.624	0.682	0.652
	AoD	0.625	0.272	0.729

Supplemental Table 15: Zero-shot classification performance of CMR-CLIP model for detecting cardiac abnormalities on 1.5 T (N=1,844) and 3 T (N=914) MRI scans across the internal cohort.

Magnetic Field Strength	Label	Accuracy	F1-score	AUC
1.5T (N=421)	LVDys	0.717	0.651	0.799
	RVDys	0.361	0.326	0.726
	LVD	0.753	0.373	0.733
	RVD	0.451	0.252	0.819
	LVH	0.556	0.403	0.830
	LGE	0.639	0.716	0.661
	AoD	0.378	0.206	0.605
3.0T (N=7)	LVDys	0.571	0.400	0.500
	RVDys	0.000	0.000	0.000
	LVD	0.714	0.000	0.900
	RVD	0.571	0.571	0.700
	LVH	0.571	0.667	1.000
	LGE	1.000	1.000	1.000
	AoD	0.286	0.286	0.200

Supplemental Table 16: Zero-shot classification performance of CMR-CLIP model for detecting cardiac abnormalities on 1.5 T (N=421) and 3.0T (N=7) MRI scans on the internal cohort.

	Parameter	Value
Hyperparameter	number of gpu	4
	optimizer	AdamW
	learning rate	3e-5
	loss	NormSoftmaxLoss
	epoch	100
	batch size	16
	extraction res	256
	input res	224
	projection dim	512
	stride	1
Video Params	model	SpaceTimeTransformer
	arch config	base patch16 224
	num frames	64
	pretrained	true
	time init	Zeros
Text Params	model	Bio_ClinicalBERT
	pretrained	true

Supplemental Table 17: The training hyperparameters and the video- and text-encoder configurations used in CMR-CLIP.