Supplementary Table 1: Groups of keywords used in the search string.

Group	Keywords
AI	AI Artificial Intelligence Deep Learning Machine learning
Self-supervised Learning / Fondation Models	Foundation Model* Large Language Model* Vision Language Model* Unified Vision Model* LLM VLM UVM representation learning self-supervis* self supervis* contrastive learning contrastive-learning contrastive loss contrastive-loss contrastive training contrastive-training noise contrastive estimation NCE pretext learning pretext-learning pre-text learning pretext task pretext-task pre-text task autoencoder* auto-encoder auto encoder diffusion model generative model generative ai
Medical Images	endoscop* colonoscop* gastroscop* whole slide imaging

	microscope slide, microscope slides pathology imaging pathology slid* microscope microscopy ophthalmoscope ophthalmoscopy fundus photography fundus imaging retinal image dermatoscope dermatoscopy medical imag* diagnostic imag* radiolog* imag* CT scan* CT-scan* computed tomography computer-assisted tomography CAT scan* CAT-scan* PET scan* positron emission tomography FDG-PET PET-CT PET/CT MRI magnetic resonance imaging MR scan* MR-scan* MRI scan* MRI-scan* x-ray* radiograph* mammogram mammography fluoroscopy ultrasound ultrasonograph* sonograph* echocardio* panorama* spect-ct spect scan Scintigraphy gamma scan perfusion scan Endoscopy Colonoscopy Gastroscopy

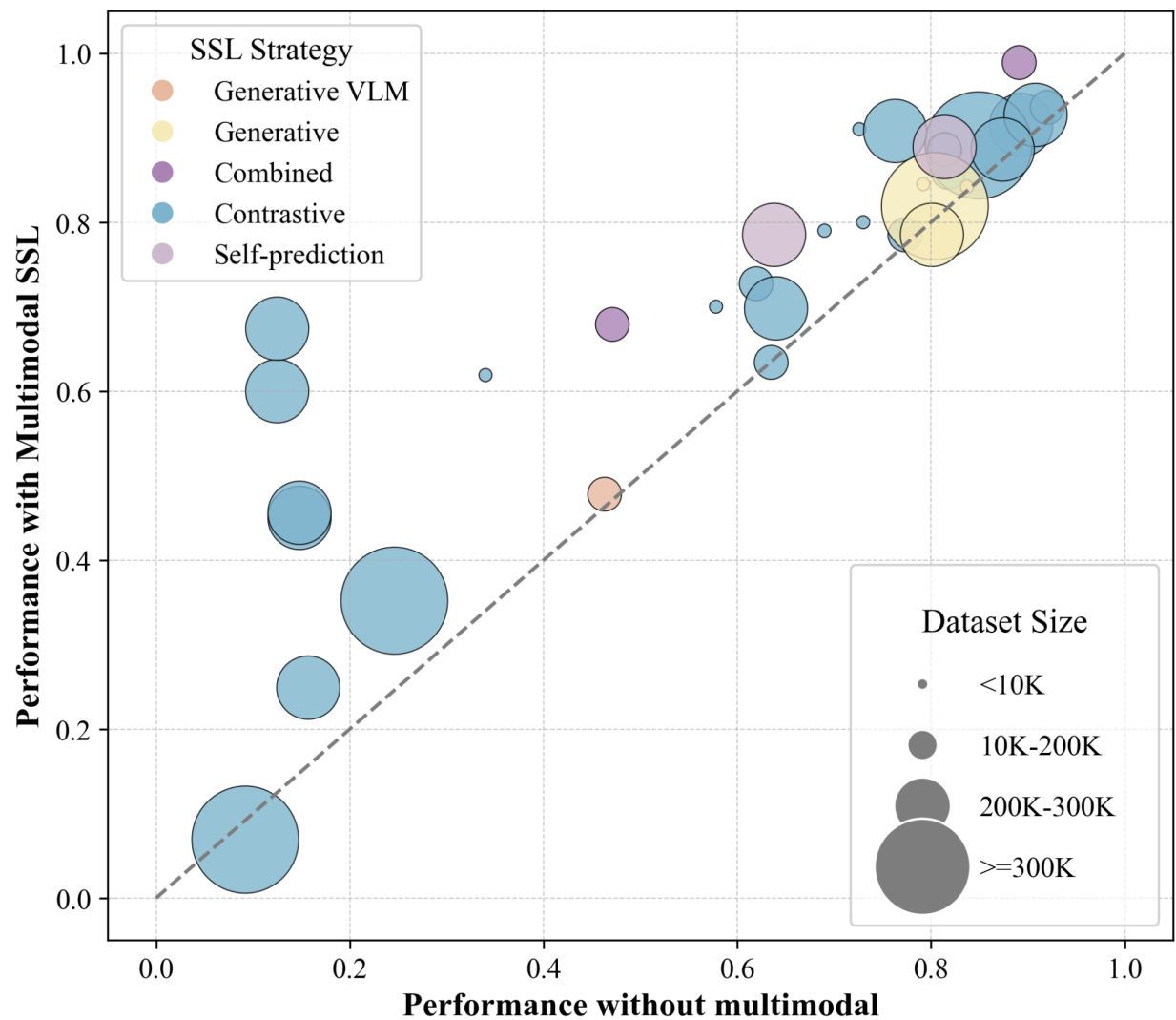|	radiography diagnostic imaging tomography, x-ray computed positron-emission tomography Positron Emission Tomography Computed Tomography Magnetic Resonance Imaging Radiography X-Rays mammography fluoroscopy Ultrasonography Echocardiography Single Photon Emission Computed Tomography Computed Tomography radionuclide imaging Ventilation-Perfusion Scan
Non-imaging Modalities	genomic* gene genes snp* single nucleotide polymorphism proteomic* dna rna Multimodal Health records Medical records Medication Vitals Lab test radiology report clinical note physician* note doctor* note

Supplementary Table 2: Complete search string for Scopus and PubMed.

Database	Search Strings
Scopus	((TITLE-ABS ("Foundation Model*") OR TITLE-ABS ("Large Language Model*") OR TITLE-ABS ("Vision Language Model*") OR TITLE-ABS ("Unified Vision Model*") OR TITLE-ABS ("LLM") OR TITLE-ABS ("VLM") OR TITLE-ABS ("UVM") OR TITLE-ABS ("representation learning") OR TITLE-ABS ("self-supervis*") OR TITLE-ABS ("self supervis*") OR TITLE-ABS ("contrastive learning") OR TITLE-ABS ("contrastive-learning") OR TITLE-ABS ("contrastive

loss") OR TITLE-ABS ("contrastive-loss") OR TITLE-ABS ("contrastive training") OR TITLE-ABS ("contrastive-training") OR TITLE-ABS ("noise contrastive estimation") OR TITLE-ABS ("NCE") OR TITLE-ABS ("pretext learning") OR TITLE-ABS ("pretext-learning") OR TITLE-ABS ("pre-text learning") OR TITLE-ABS ("pretext task") OR TITLE-ABS ("pretext-task") OR TITLE-ABS ("pre-text task") OR TITLE-ABS ("autoencoder*") OR TITLE-ABS ("auto-encoder") OR TITLE-ABS ("auto encoder") OR TITLE-ABS ("diffusion model") OR TITLE-ABS ("generative model") OR TITLE-ABS ("generative ai")) AND (TITLE-ABS ("endoscop*") OR TITLE-ABS ("colonoscop*") OR TITLE-ABS ("gastroscop*") OR TITLE-ABS ("whole slide imaging") OR TITLE-ABS ("microscope slide, microscope slides") OR TITLE-ABS ("pathology imaging") OR TITLE-ABS ("pathology slid*") OR TITLE-ABS ("microscope") OR TITLE-ABS ("microscopy") OR TITLE-ABS ("ophthalmoscope") OR TITLE-ABS ("ophthalmoscopy") OR TITLE-ABS ("fundus photography") OR TITLE-ABS ("fundus imaging") OR TITLE-ABS ("retinal image") OR TITLE-ABS ("dermatoscope") OR TITLE-ABS ("dermatoscopy") OR TITLE-ABS ("medical imag*") OR TITLE-ABS ("diagnostic imag*") OR TITLE-ABS ("radiolog* imag*") OR TITLE-ABS ("CT scan*") OR TITLE-ABS ("CT-scan*") OR TITLE-ABS ("computed tomography") OR TITLE-ABS ("computer-assisted tomography") OR TITLE-ABS ("CAT scan*") OR TITLE-ABS ("CAT-scan*") OR TITLE-ABS ("PET scan*") OR TITLE-ABS ("positron emission tomography") OR TITLE-ABS ("FDG-PET") OR TITLE-ABS ("PET-CT") OR TITLE-ABS ("PET/CT") OR TITLE-ABS ("MRI") OR TITLE-ABS ("magnetic resonance imaging") OR TITLE-ABS ("MR scan*") OR TITLE-ABS ("MR-scan*") OR TITLE-ABS ("MRI scan*") OR TITLE-ABS ("MRI-scan*") OR TITLE-ABS ("x-ray*") OR TITLE-ABS ("radiograph*") OR TITLE-ABS ("mammogram") OR TITLE-ABS ("mammography") OR TITLE-ABS ("fluoroscopy") OR TITLE-ABS ("ultrasound") OR TITLE-ABS ("ultrasonograph*") OR TITLE-ABS ("sonograph*") OR TITLE-ABS ("echocardio*") OR TITLE-ABS ("panorama*") OR TITLE-ABS ("spect-ct") OR TITLE-ABS ("spect scan") OR TITLE-ABS ("scintigraphy") OR TITLE-ABS ("gamma scan") OR TITLE-ABS ("perfusion scan") OR TITLE-ABS ("Endoscopy") OR TITLE-ABS ("Colonoscopy") OR TITLE-ABS ("Gastroscopy") OR TITLE-ABS ("radiography") OR TITLE-ABS ("diagnostic imaging") OR TITLE-ABS ("tomography, x-ray computed") OR TITLE-ABS ("positron-emission tomography") OR TITLE-ABS ("Positron Emission Tomography Computed Tomography") OR TITLE-ABS ("Magnetic Resonance Imaging") OR TITLE-ABS ("Radiography") OR TITLE-ABS ("X-Rays") OR TITLE-ABS ("mammography") OR TITLE-ABS ("fluoroscopy") OR TITLE-ABS ("Ultrasonography") OR TITLE-ABS ("Echocardiography") OR TITLE-ABS ("Single Photon Emission Computed Tomography Computed Tomography") OR TITLE-ABS ("radionuclide imaging") OR TITLE-ABS ("Ventilation-Perfusion Scan")) AND (TITLE-ABS ("genomic*") OR TITLE-ABS ("gene") OR TITLE-ABS ("genes") OR TITLE-ABS ("snp*") OR TITLE-ABS ("single nucleotide polymorphism ") OR TITLE-ABS ("proteomic*") OR TITLE-ABS ("dna") OR TITLE-ABS ("rna") OR TITLE-ABS ("Multimodal") OR TITLE-ABS ("Health records") OR TITLE-ABS ("Medical records") OR TITLE-ABS ("Medication") OR TITLE-ABS ("Vitals") OR TITLE-ABS ("Lab test") OR TITLE-ABS ("radiology report") OR TITLE-ABS ("clinical note") OR TITLE-ABS ("physician* note") OR TITLE-ABS ("doctor* note")) AND (LIMIT-TO (PUBYEAR , 2023) OR LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (

	PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015) OR LIMIT-TO (PUBYEAR , 2014) OR LIMIT-TO (PUBYEAR , 2013) OR LIMIT-TO (PUBYEAR , 2012)))
PubMed	("Foundation Model*"[Title/Abstract] OR "Large Language Model*"[Title/Abstract] OR "Vision Language Model*"[Title/Abstract] OR "Unified Vision Model*"[Title/Abstract] OR "LLM"[Title/Abstract] OR "VLM"[Title/Abstract] OR "UVM"[Title/Abstract] OR "representation learning"[Title/Abstract] OR "self-supervis*"[Title/Abstract] OR "self supervis*"[Title/Abstract] OR "contrastive learning"[Title/Abstract] OR "contrastive-learning"[Title/Abstract] OR "contrastive loss"[Title/Abstract] OR "contrastive-loss"[Title/Abstract] OR "contrastive training"[Title/Abstract] OR "contrastive-training"[Title/Abstract] OR "noise contrastive estimation"[Title/Abstract] OR "NCE"[Title/Abstract] OR "pretext learning"[Title/Abstract] OR "pretext-learning"[Title/Abstract] OR "pre-text learning"[Title/Abstract] OR "pretext task"[Title/Abstract] OR "pretext-task"[Title/Abstract] OR "pre-text task"[Title/Abstract] OR "autoencoder*"[Title/Abstract] OR "auto-encoder"[Title/Abstract] OR "auto encoder"[Title/Abstract] OR "diffusion model"[Title/Abstract] OR "generative model"[Title/Abstract] OR "generative ai"[Title/Abstract]) AND ("endoscop*"[Title/Abstract] OR "colonoscop*"[Title/Abstract] OR "gastroscop*"[Title/Abstract] OR "whole slide imaging"[Title/Abstract] OR "microscope slide, microscope slides"[Title/Abstract] OR "pathology imaging"[Title/Abstract] OR "pathology slid*"[Title/Abstract] OR "microscope"[Title/Abstract] OR "microscopy"[Title/Abstract] OR "ophthalmoscope"[Title/Abstract] OR "ophthalmoscopy"[Title/Abstract] OR "fundus photography"[Title/Abstract] OR "fundus imaging"[Title/Abstract] OR "retinal image"[Title/Abstract] OR "dermatoscope"[Title/Abstract] OR "dermatoscopy"[Title/Abstract] OR "medical imag*"[Title/Abstract] OR "diagnostic imag*"[Title/Abstract] OR "radiolog* imag*"[Title/Abstract] OR "CT scan*"[Title/Abstract] OR "CT-scan*"[Title/Abstract] OR "computed tomography"[Title/Abstract] OR "computer-assisted tomography"[Title/Abstract] OR "CAT scan*"[Title/Abstract] OR "CAT-scan*"[Title/Abstract] OR "PET scan*"[Title/Abstract] OR "positron emission tomography"[Title/Abstract] OR "FDG-PET"[Title/Abstract] OR "PET-CT"[Title/Abstract] OR "PET/CT"[Title/Abstract] OR "MRI"[Title/Abstract] OR "magnetic resonance imaging"[Title/Abstract] OR "MR scan*"[Title/Abstract] OR "MR-scan*"[Title/Abstract] OR "MRI scan*"[Title/Abstract] OR "MRI-scan*"[Title/Abstract] OR "x-ray*"[Title/Abstract] OR "radiograph*"[Title/Abstract] OR "mammogram"[Title/Abstract] OR "mammography"[Title/Abstract] OR "fluoroscopy"[Title/Abstract] OR "ultrasound"[Title/Abstract] OR "ultrasonograph*"[Title/Abstract] OR "sonograph*"[Title/Abstract] OR "echocardio*"[Title/Abstract] OR "panorama*"[Title/Abstract] OR "spect-ct"[Title/Abstract] OR "spect scan"[Title/Abstract] OR "Scintigraphy"[Title/Abstract] OR "gamma scan"[Title/Abstract] OR "perfusion scan"[Title/Abstract] OR "Endoscopy"[MeSH Terms] OR "Colonoscopy"[MeSH Terms] OR "Gastroscopy"[MeSH Terms] OR

| | "radiography"[MeSH Terms] OR "diagnostic imaging"[MeSH Terms] OR "tomography, x-ray computed"[MeSH Terms] OR "positron-emission tomography"[MeSH Terms] OR "Positron Emission Tomography Computed Tomography"[MeSH Terms] OR "Magnetic Resonance Imaging"[MeSH Terms] OR "Radiography"[MeSH Terms] OR "X-Rays"[MeSH Terms] OR "mammography"[MeSH Terms] OR "fluoroscopy"[MeSH Terms] OR "Ultrasonography"[MeSH Terms] OR "Echocardiography"[MeSH Terms] OR "Single Photon Emission Computed Tomography Computed Tomography"[MeSH Terms] OR "radionuclide imaging"[MeSH Terms] OR "Ventilation-Perfusion Scan"[MeSH Terms]) AND ("genomic*"[Title/Abstract] OR "gene"[Title/Abstract] OR "genes"[Title/Abstract] OR "snp*"[Title/Abstract] OR "single nucleotide polymorphism "[Title/Abstract] OR "proteomic*"[Title/Abstract] OR "dna"[Title/Abstract] OR "rna"[Title/Abstract] OR "Multimodal"[Title/Abstract] OR "Health records"[Title/Abstract] OR "Medical records"[Title/Abstract] OR "Medication"[Title/Abstract] OR "Vitals"[Title/Abstract] OR "Lab test"[Title/Abstract] OR "radiology report"[Title/Abstract] OR "clinical note"[Title/Abstract] OR "physician* note"[Title/Abstract] OR "doctor* note"[Title/Abstract] OR "Genomics"[MeSH Terms] OR "Genome"[MeSH Terms] OR "Genes"[MeSH Terms] OR "Polymorphism, Single Nucleotide"[MeSH Terms] OR "Proteomics"[MeSH Terms] OR "DNA"[MeSH Terms] OR "RNA"[MeSH Terms]) AND ("2012/01/01"[PDAT] : "2024/01/01"[PDAT]) |

Supplementary Figure 1: **Downstream task performance comparison of multimodal and single-modal models across varying dataset sizes for SSL pretraining**. The dashed, black line indicates equality in downstream performance between single modal and multimodal pretraining. The color of the points indicates the SSL pretraining methodxx½. The size of the points indicates the size of the pretraining dataset.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	1-3
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	1-3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	17-19
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	17-19
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	17-19, supplementary
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	17-19
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	17-19
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	17-19
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	17-19
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	17-19
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	17-19
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	17-19
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	17-19
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	17-19
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	17-19
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	17-19
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	17-19
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	17-19
Certainty	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	17-19

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	3-14, Figure 4
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	3-14
Study characteristics	17	Cite each included study and present its characteristics.	3-14
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	3-14, 19
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	3-14
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	3-14
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	3-14, Figure 5
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	3-14
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	3-14
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	3-14
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	3-14
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	14-17
	23b	Discuss any limitations of the evidence included in the review.	14-17
	23c	Discuss any limitations of the review processes used.	14-17
	23d	Discuss implications of the results for practice, policy, and future research.	14-17
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	17-19
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	17-19
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	17-19
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	20
Competing interests	26	Declare any competing interests of review authors.	20
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	17-19