

# Improving Accessibility for Health AI with Embeddings


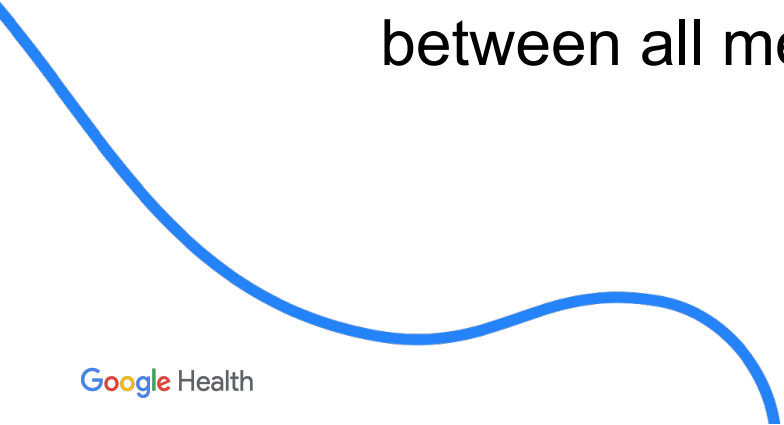
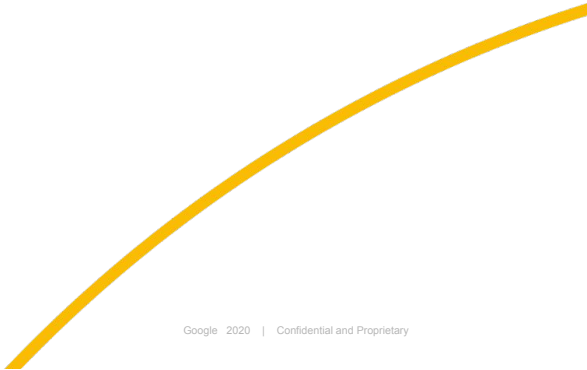
Andrew Sellergren

August 20, 2024



**Andrew Sellergren**

Software Engineer  
Health AI

- 
- Google Health's mission is to improve accessibility
  - Vector embeddings reduce barriers to entry
  - Vector embeddings are a potential bridge between all medical modalities
- 
- 



# Reducing barriers to AI development for medical imaging



# Developing AI for chest x-ray (CXR) is challenging

## CXRs are accessible and available

~1B CXRs are taken to detect and manage many health conditions

Accessible and inexpensive imaging modality

## AI can bridge gaps with interpretation

High quality interpretation is a challenge


Short supply of radiologists and variability between experts and sites

Long tail of rare conditions

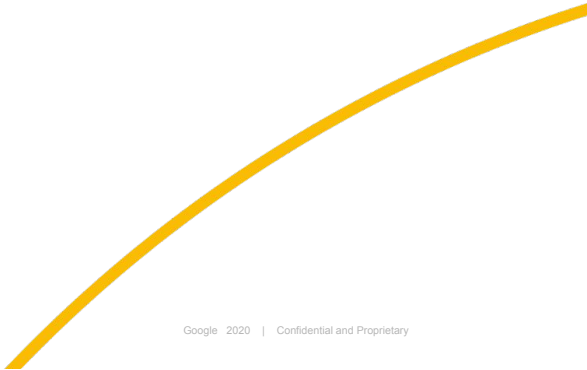
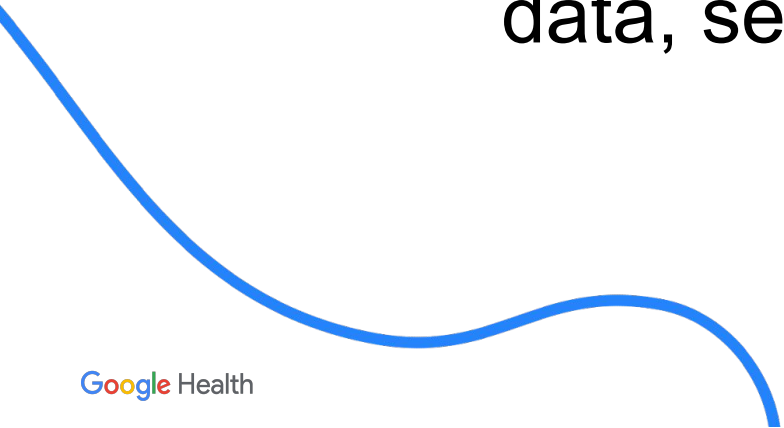
## CXR AI is difficult and expensive to build

Building robust CXR models is challenging and time consuming

Requires large, curated datasets and extensive fine tuning



Our **mission** is to enable others to train better custom medical imaging models with less data, setup, and compute



# Reducing the barrier to entry for training custom models to read CXRs

Decreasing training  
time




Improving label efficiency

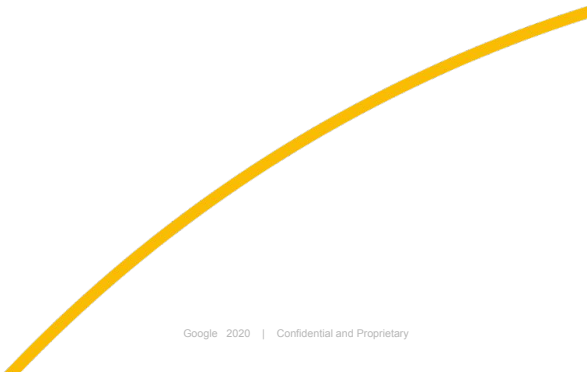
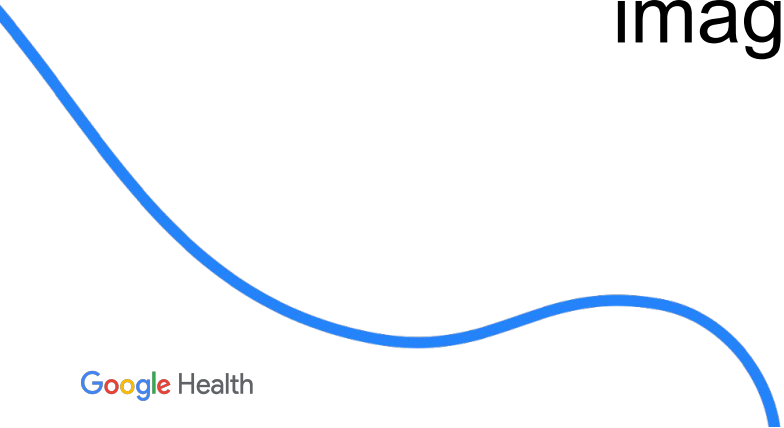


Reducing model  
complexity



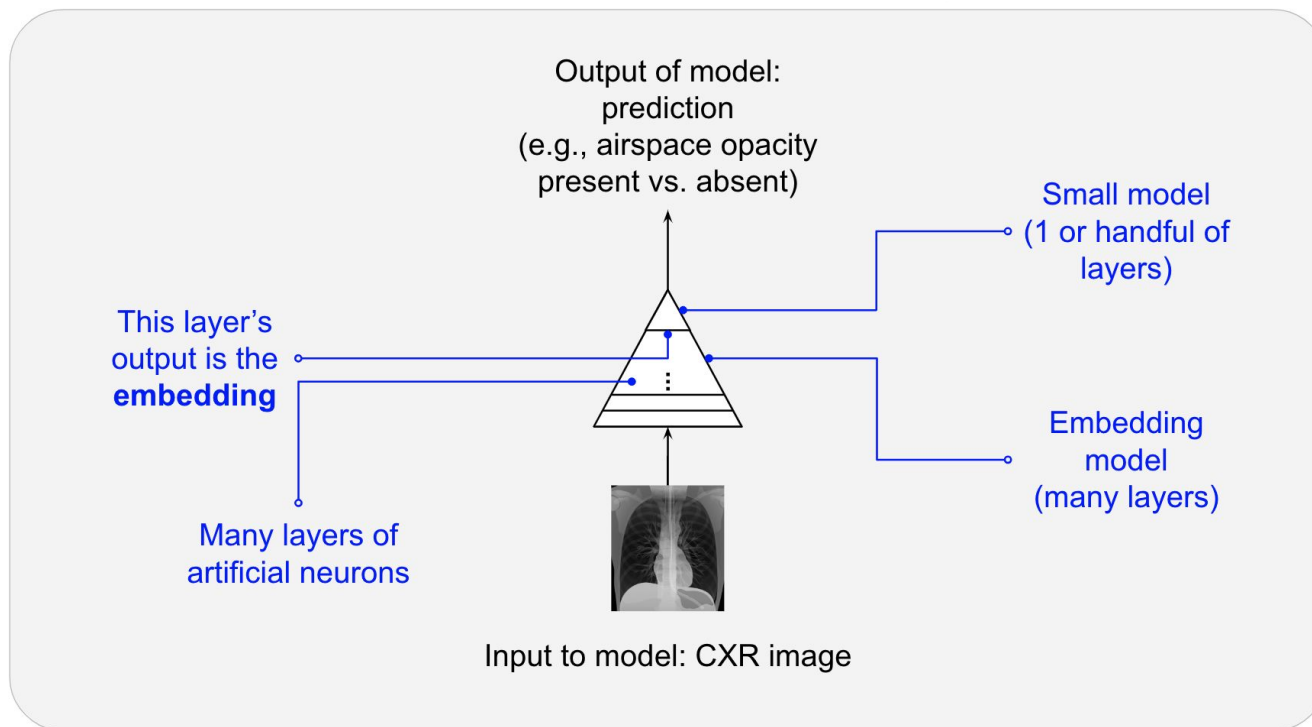


Our **method** is to offload as much of the heavy lifting as possible via large-scale medical imaging pretraining

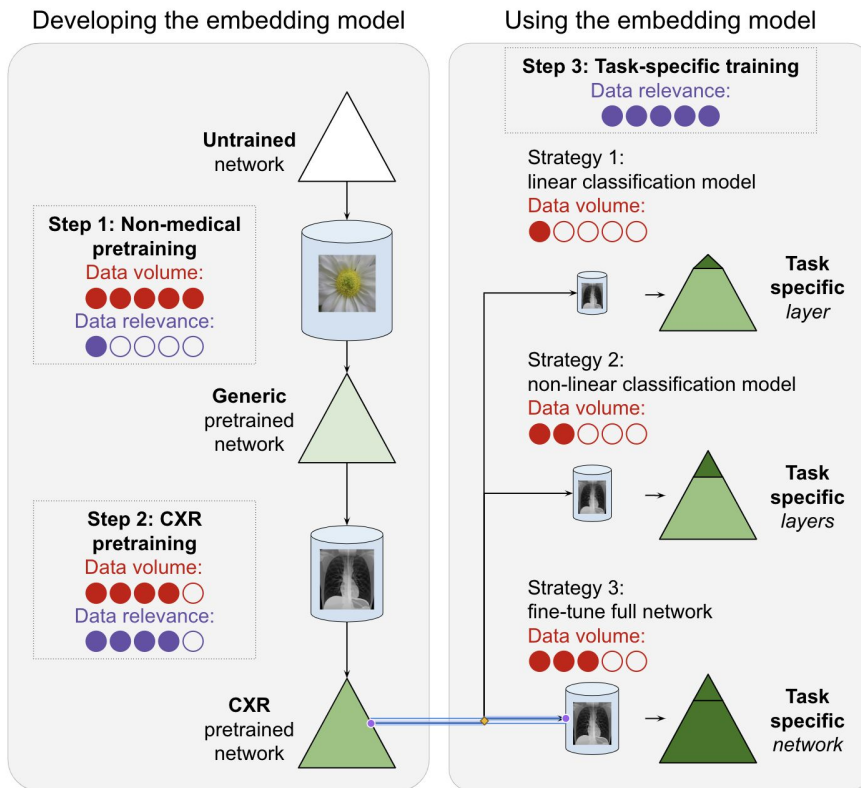




# Typical setup



# CXR-specific networks with a 2nd stage of pretraining



**Traditional**

**CXR Foundation**

# How do we optimize for learning the best embeddings?

Upstream and downstream task performance are not necessarily correlated<sup>1</sup>

Standard measure of representation quality is fewshot linear probe performance on a variety of downstream tasks

Contrastive losses more directly optimize the latent space of our representations

Cross-entropy loss may have shortcomings such as poor margins<sup>2</sup>

Bigger is better

Larger networks learn better representations when given sufficient data<sup>3</sup> (and more data is better)

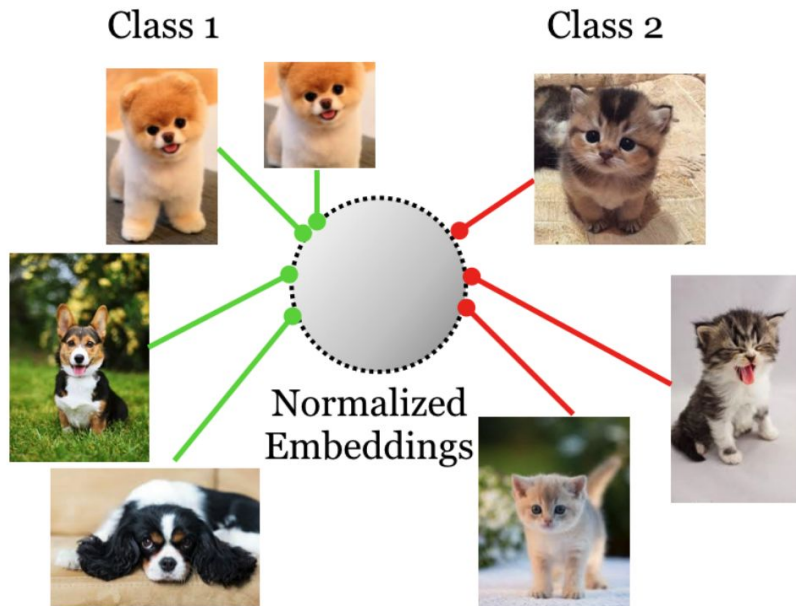
Requires large, curated datasets and extensive fine tuning

<sup>1</sup> ["Why Do Better Loss Functions Lead to Less Transferable Features?"](#)

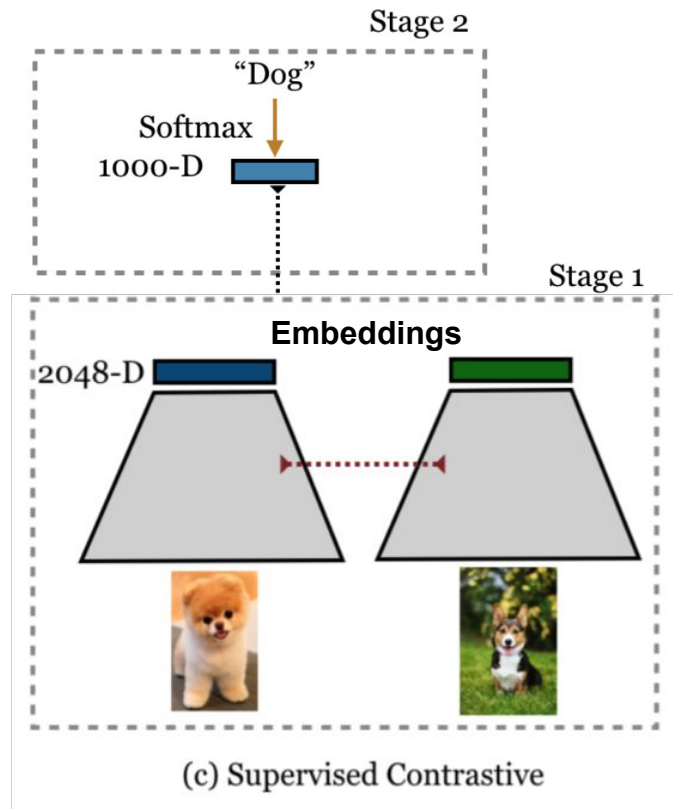
<sup>2</sup> from [SupCon paper](#): "lack of robustness to noisy labels [59, 44] and the possibility of poor margins [14, 30], leading to reduced generalization performance"

<sup>3</sup> see [Big Transfer \(BiT\) results](#)

# Supervised Contrastive Learning



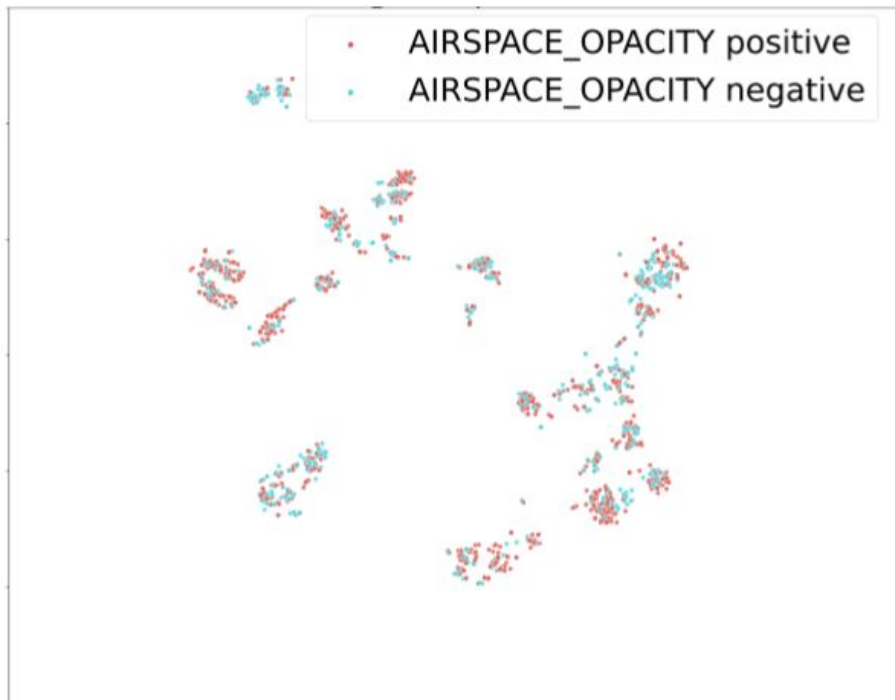
Supervised Contrastive



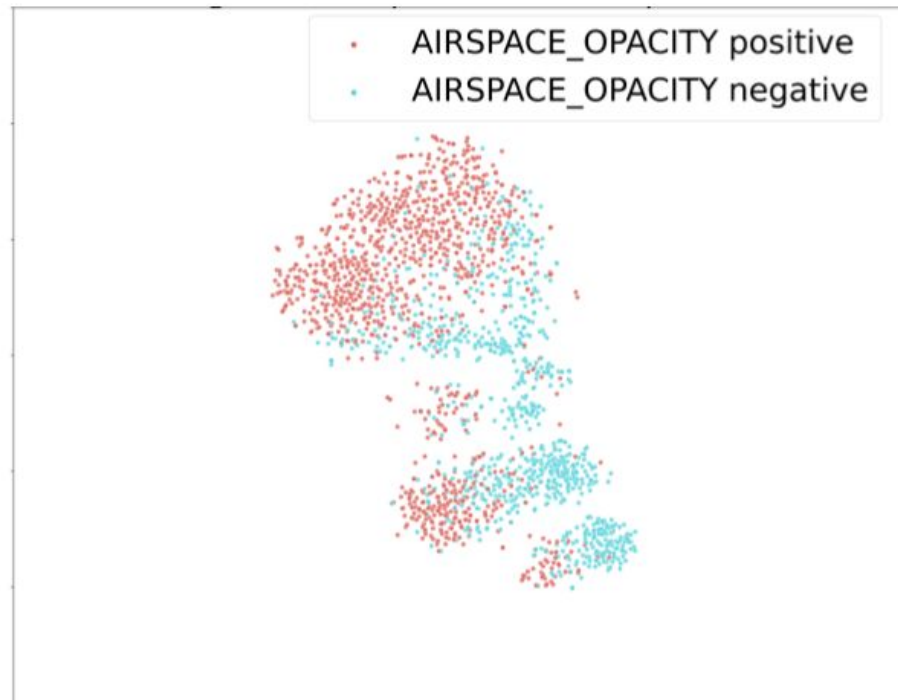
(c) Supervised Contrastive

# Embeddings to distinguish classes from each other

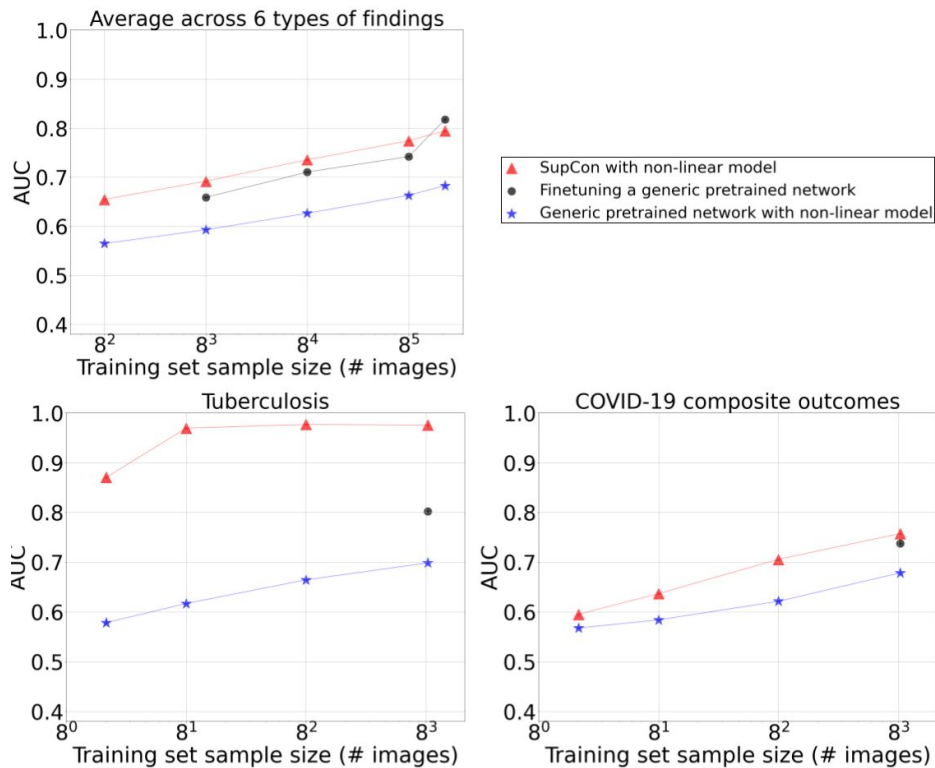
## Generic network embeddings



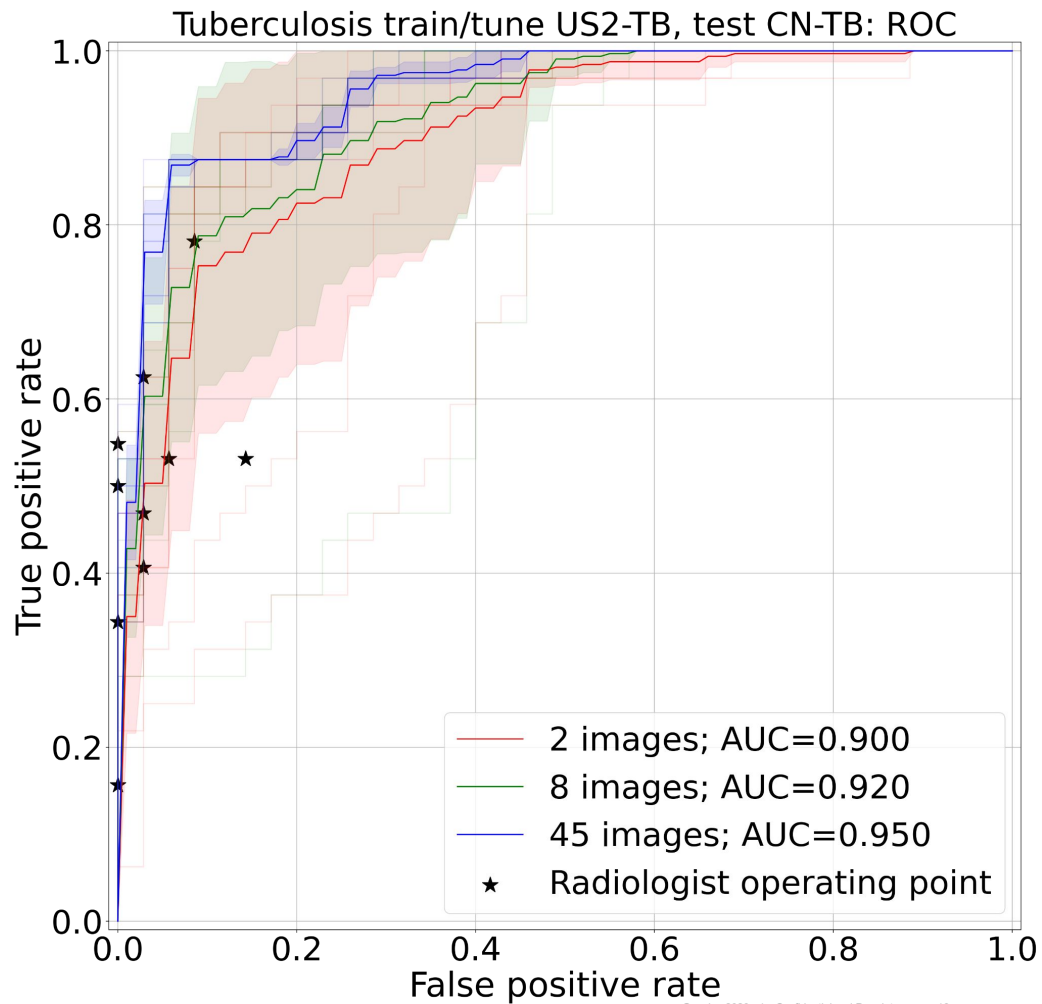
## CXR network embeddings



# Same performance with 100x to 1000x less data

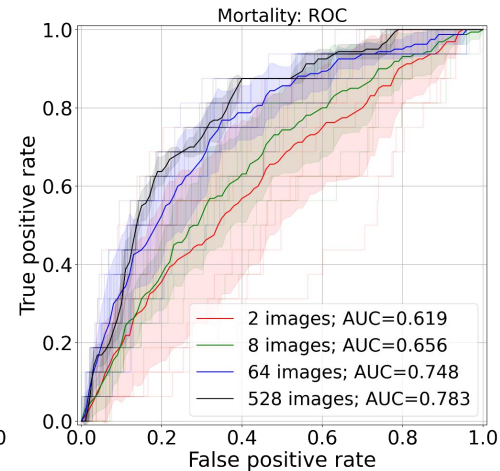
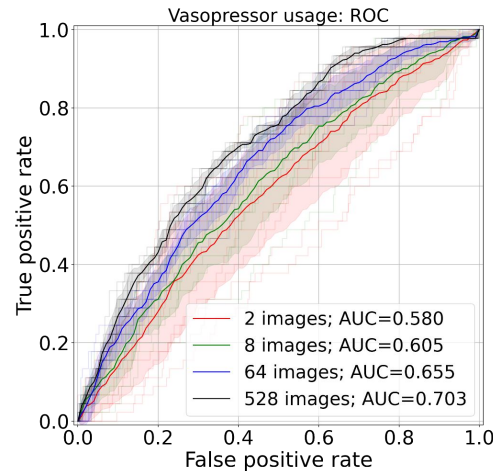
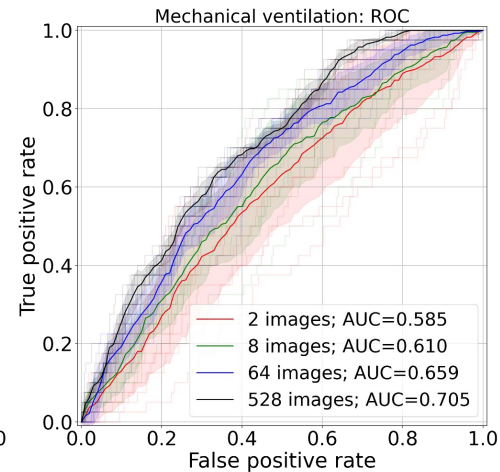
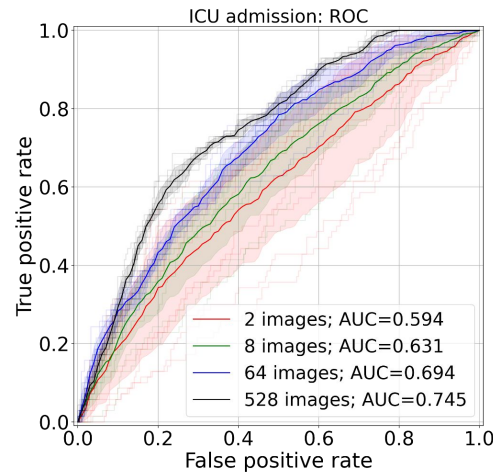


Example: TB model trained on <100 images was non-inferior to 10 radiologists

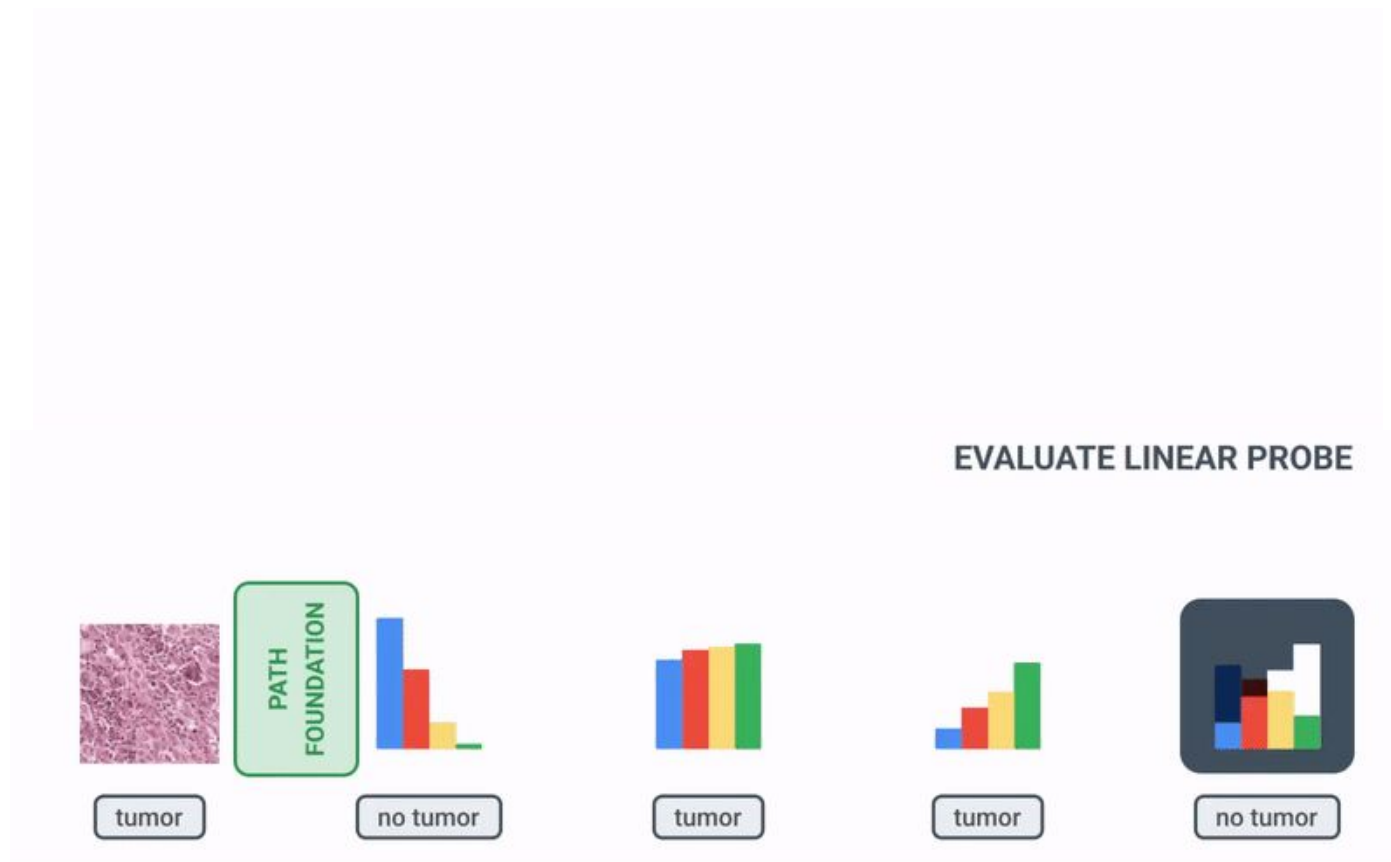




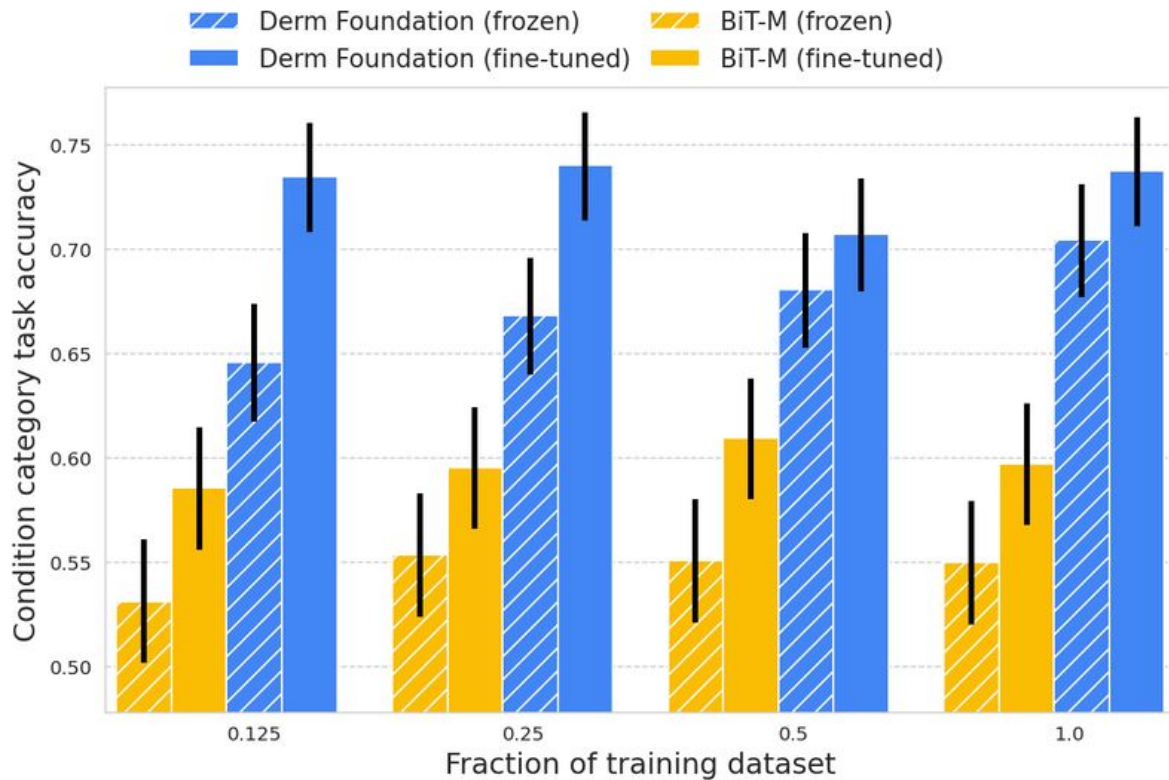
# Example: COVID-19 severity model trained on ~500 images

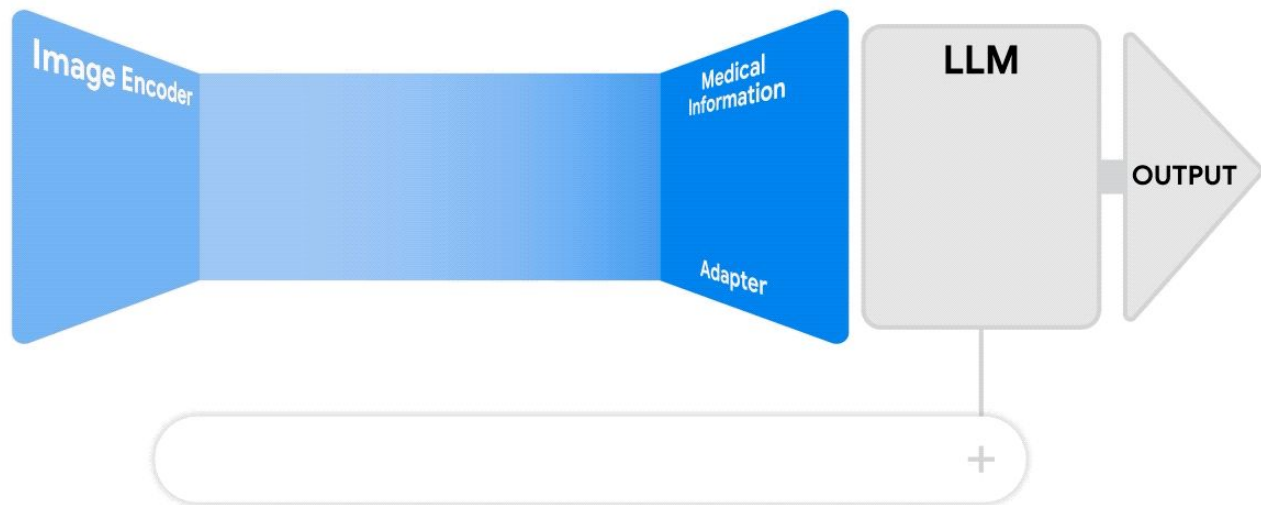


# Extending to other modalities (Pathology)



# Extending to other modalities (Dermatology)





Present results across the following tasks:

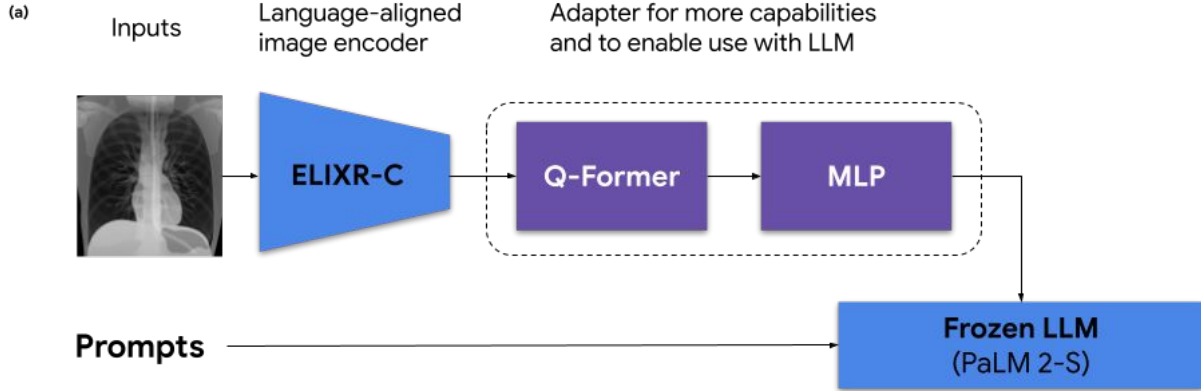
Zero-shot

Data-efficient  
classification

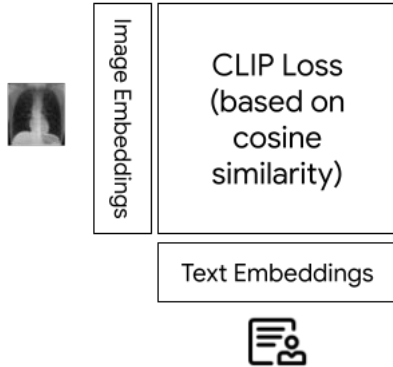
Semantic search

Visual question  
answering

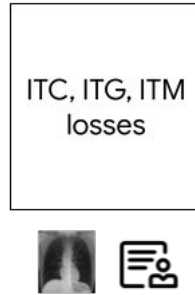
Radiology report quality  
assurance



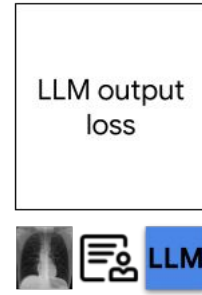
(b) **Step 1:** initial alignment of image embeddings to language.  
**Enables:** text-based examples (“zero-shot” classification).



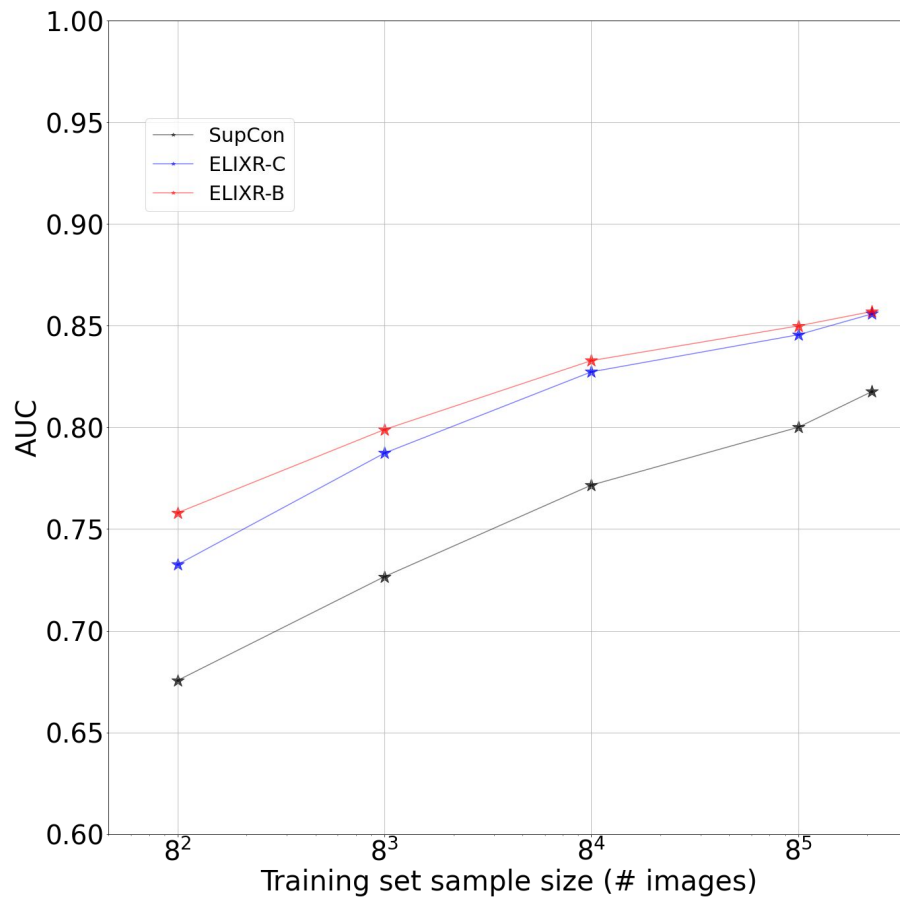
**Step 2:** train Q-Former to extract information relevant to text prompts, such as spatial locality.  
**Enables:** spatial reasoning and limited language output (captioning).



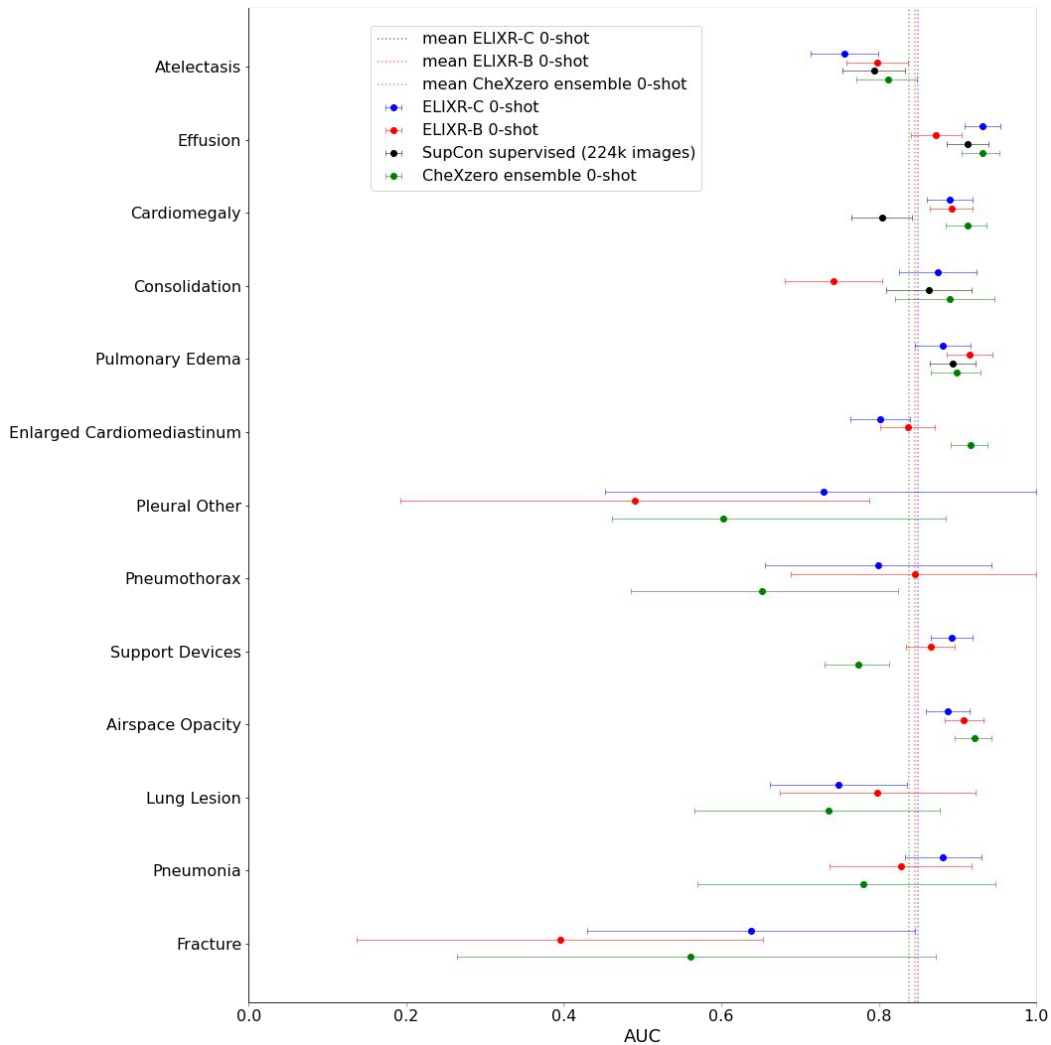
**Step 3:** train Q-Former and MLP to adapt information for LLM.  
**Enables:** complex language interaction/output such as visual question answering, report quality assurance.



# Data-efficiency improvements by 2 orders of magnitude



# Zero-shot comparable to fully supervised



Show me examples of small right pleural effusion, no left pleural effusion



Is pleural effusion present in this image?



yes

What is the location of the pleural effusion, if present?



left

What is the size of pleural effusion, if present?



moderate



IMPRESSION: Stable appearance of right-sided pleural effusion. Pneumothorax has resolved.

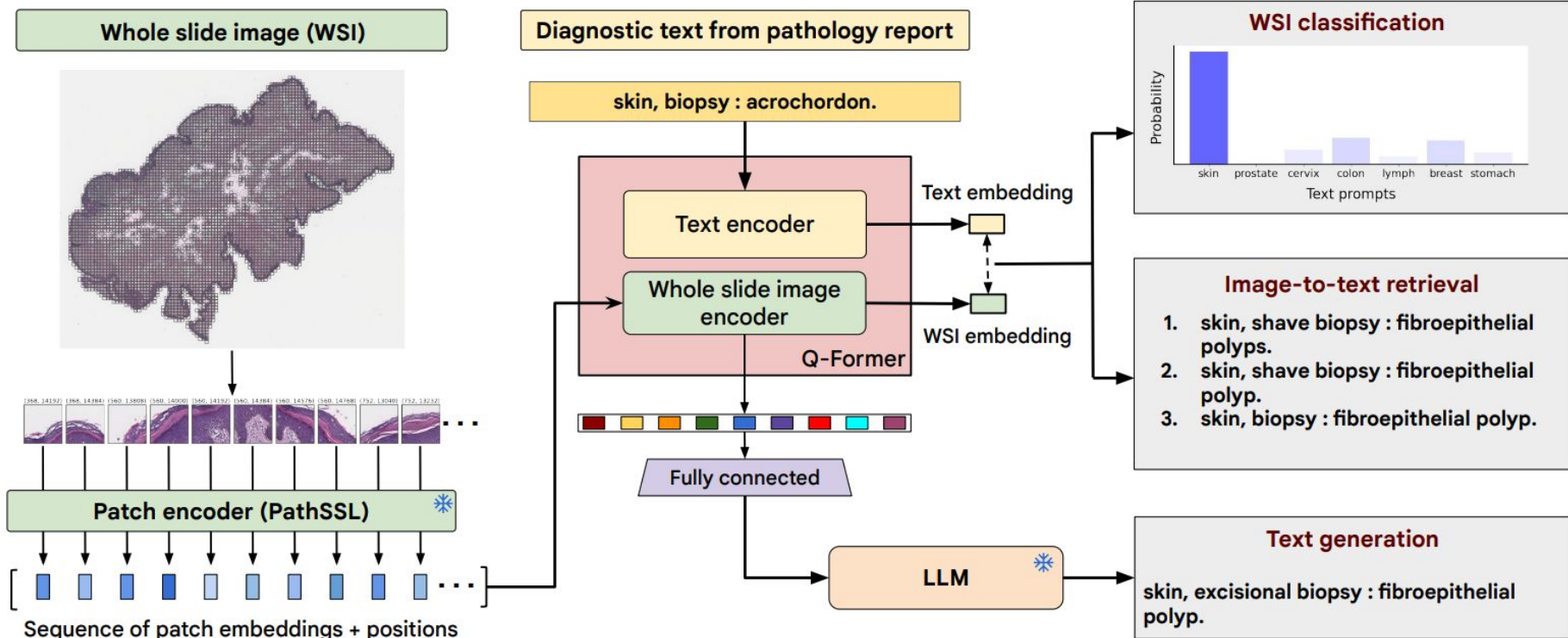
Is the report inaccurate and need to be edited before being signed off?






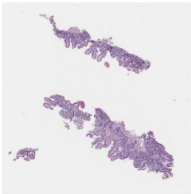
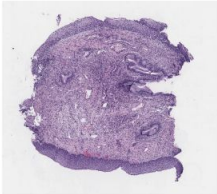
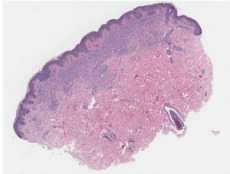
Yes. There is a right pneumothorax. It has not resolved.



# Extending to other modalities (Pathology)



# Extending to other modalities (Pathology)

	Example 1	Example 2	Example 3
WSI thumbnail			
Enlarged view			
Original text	duodenum, biopsy : unremarkable intestinal mucosa.	cervix : biopsy: - low grade squamous intraepithelial lesion (cin 1, mild dysplasia).	skin, biopsy : intradermal nevus.
Top retrieved text	duodenum, third part, biopsy : small bowel mucosa with no pathologic diagnosis.	cervix : biopsy: - high grade squamous intraepithelial lesion (cin-2; hsil).	skin, punch biopsy : intradermal nevus.
Generated text	duodenum, biopsy : duodenal mucosa with no significant pathologic changes.	cervix, biopsy : low grade squamous intraepithelial lesion (cin 1).	skin, punch biopsy : compound nevus.
Pathologist review	Agree with all	Favor HSIL (high grade)	Favor compound nevus

# What's next? Med-Gemini

Medical Modalities  
& Knowledge



Gemini

Med-Gemini



Applications

Classification

Visual Question  
Answering

2D Report  
Generation

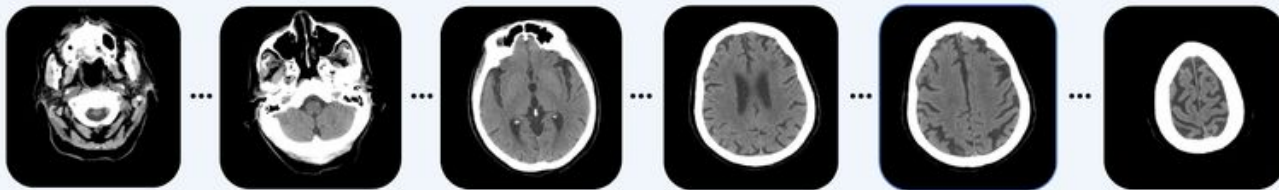
3D Report  
Generation

Polygenic Risk  
Prediction

# What's next? Med-Gemini

## Prompt

Given the image, what are the radiology findings?



## Response

FINDINGS: mild cortical atrophy and periventricular white matter changes are noted compatible with microvasculopathy, there is proportionate ventriculomegaly. no bleed, mass effect, midline shift or abnormal intracranial fluid is seen. no evidence of acute intracranial hemorrhage, territorial infarct, abnormal fluid or mass effect. no abnormal extra-axial fluid collection is identified. no definitive evidence of intracranial hemorrhage or large territorial infarct within the limits of ct imaging. **old left basilar ganglia infarct**. no intracranial mass, swelling, hemorrhage, infarction or extra-axial collection is seen. visualized mastoid air cells: well pneumatized. visualized paranasal sinuses: well pneumatized. atherosclerotic calcification of vasculature at the skull base.

## Radiologist Written Report

FINDINGS: the fourth ventricle is midline. the remaining ventricular system is within normal limits. there is mild chronic white matter microangiopathic disease and moderate generalized cortical atrophy. no bleed, mass effect, midline shift or abnormal intracranial fluid is seen. bone windows demonstrate an intact skull. the remaining visualized paranasal sinuses and mastoids are clear. chronic white matter microangiopathic disease, mild generalized cortical atrophy, not significantly changed since 2015. no evidence of acute intracranial hemorrhage, territorial infarct, abnormal fluid or mass effect.

# A huge thank you

Data Descriptor | [Open Access](#) | [Published: 12 December 2019](#)

## **MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports**

[Alistair E. W. Johnson](#) , [Tom J. Pollard](#), [Seth J. Berkowitz](#), [Nathaniel R. Greenbaum](#), [Matthew P. Lungren](#), [Chih-ying Deng](#), [Roger G. Mark](#) & [Steven Horng](#)

[Scientific Data](#) **6**, Article number: 317 (2019) | [Cite this article](#)

**17k** Accesses | **186** Citations | **11** Altmetric | [Metrics](#)

## **CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison**

[Jeremy Irvin](#), [Pranav Rajpurkar](#), [Michael Ko](#), [Yifan Yu](#), [Silviana Ciurea-Ilcus](#), [Chris Chute](#), [Henrik Marklund](#), [Behzad Haghgoo](#), [Robyn Ball](#), [Katie Shpanskaya](#), [Jayne Seekins](#), [David A. Mong](#), [Safwan S. Halabi](#), [Jesse K. Sandberg](#), [Ricky Jones](#), [David B. Larson](#), [Curtis P. Langlotz](#), [Bhavik N. Patel](#), [Matthew P. Lungren](#), [Andrew Y. Ng](#)