

Modeling Clinical Uncertainty in Radiology Reports: from Explicit Uncertainty Markers to Implicit Reasoning Pathways

Paloma Rabaey, Jong Hak Moon, Jung-Oh Lee, Min Gwan Kim, Hangyul Yoon, Thomas Demeester, Edward Choi

Clinical Uncertainty in Radiology Reports

Structuring radiology reports means turning them into a **machine-readable standardized format** which can be used for evaluation of automated radiology report generation models, as well as training of image-based disease classification models.

The **Language dataset** is a fine-grained, structured benchmark of MIMIC-CXR reports. In this work, we build an extension of Language, called **Language++**, which models two types of **clinical uncertainty** which are inherent to these reports:

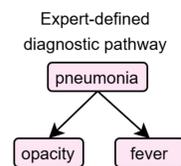
- **Explicit uncertainty**, which arises when radiologists explicitly convey doubt about a finding. We quantify this uncertainty as a probability between 0 and 1, based on the hedging phrases used by the radiologist to describe a finding.
- **Implicit uncertainty**, where radiologists omit portions of their diagnostic reasoning, documenting only key findings. We reconstruct this omitted diagnostic evidence through expert-defined diagnostic pathways, which help to infer sub-findings that support each higher-level diagnosis.

CXR Report

possible early developing right middle lobe pneumonia. the lungs again are hyperinflated suggesting underlying emphysema.

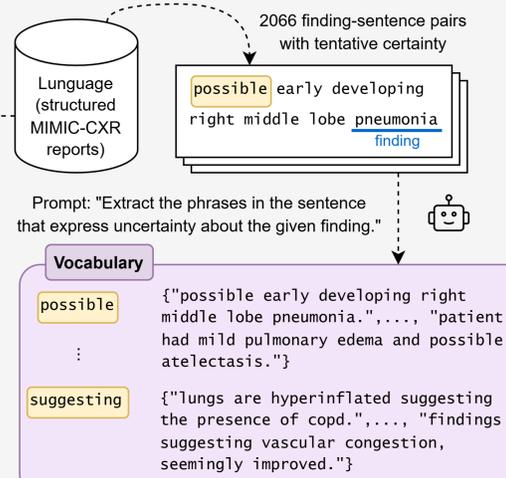
Hedging phrases convey explicit uncertainty about presence or absence of findings.

Implicit uncertainty arises due to the omission of findings which are part of diagnostic reasoning.

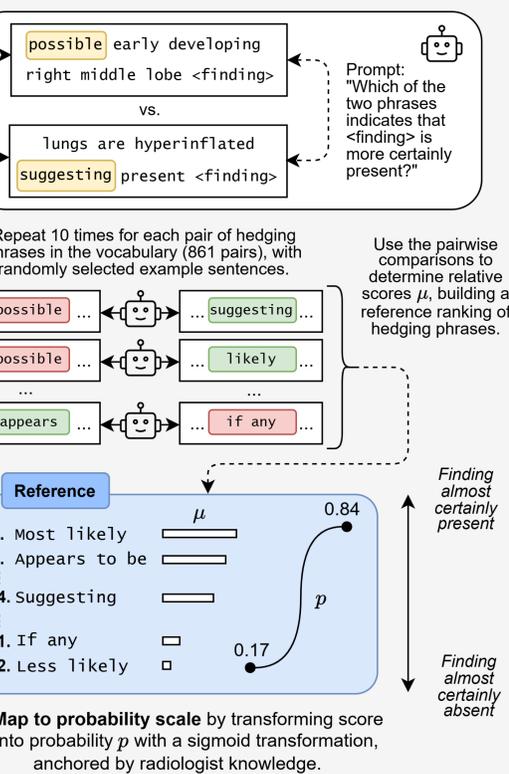


Structured CXR Report				
Finding	Attribute(s)	Certainty	Status	Probability
pneumonia	location: right middle lobe	tentative	-	0.45
lungs	measurement: hyperinflated	definitive	positive	1.00
emphysema	-	tentative	-	0.62
Language				
opacity	-	tentative	-	0.45
fever	-	tentative	-	0.45
Language++				

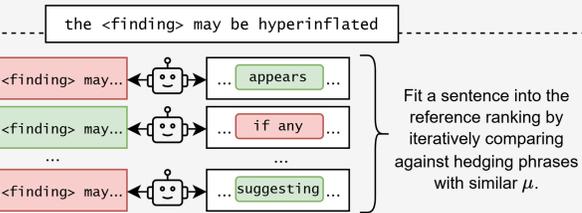
1. Extracting a vocabulary of 42 common hedging phrases



2. Building a reference ranking of hedging phrases



3. Ranking each finding-sentence pair based on reference



Explicit Uncertainty

Language is made up of **finding-sentence pairs**, for which we aim to quantify the explicit uncertainty expressed through hedging phrases. For this, we use a four-part framework, where we **create an expert-validated, LLM*-based reference ranking of common hedging phrases**, mapping each finding to a probability value based on this reference.



*LLMs included in this study are Gemini, GPT-4o, Claude and MedGemma.

Sentence, finding, hedging phrase	Prob.
No radiographic findings <i>suggesting pneumonia</i>	0.102
The lungs remain hyperinflated, <i>suggesting chronic obstructive pulmonary disease</i>	0.508
Given the clinical presentation, <i>pneumonia must be suspected</i>	0.594
Mediastinal silhouettes are stable with <i>possible slight decrease in right paratracheal prominence</i>	0.356

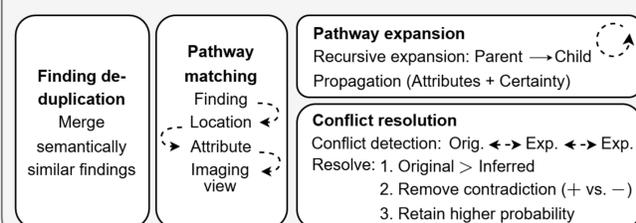
Implicit Uncertainty

Radiologists defined **diagnostic pathways for 14 common CXR diagnoses**. Whenever a finding matches one of these diagnoses, we apply the pathway to infer missing findings. Within Language, 18.7% of findings were eligible for expansion, resulting in an increase of +33.9% with sub-findings which were originally unmentioned.

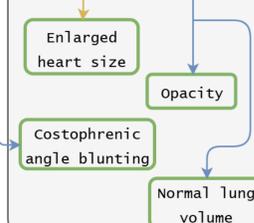
CXR Report View/Position: Erect PA

Acute exacerbation of **chf**. There is moderate **pleural effusion** along with chronic **cardiomegaly**. New right infrahilar **consolidation** could be regional **edema**.

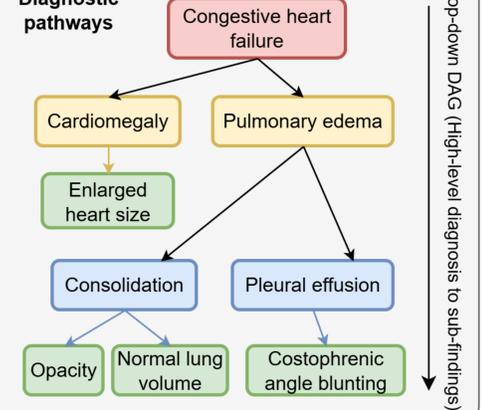
Cascading Expansion Framework



Expanded data



Diagnostic pathways



Utility of Language++

Language++ is a set of structured CXR reports, extended with (1) probabilities capturing explicit uncertainty and (2) previously unmentioned sub-findings addressing implicit uncertainty. This resource enables **uncertainty-aware training and evaluation of image classification and report generation models**, as well as the study of how diagnostic uncertainty may influence downstream clinical outcomes.

✉ paloma.rabaey@gmail.com

🌐 github.com/prabaey/language_uncertainty

📄 /in/paloma-rabaey-3a29091a2

