NTCIR-17
Medical Natural Language Processing for Social media and Clinical texts (MedNLP-SC)

Eiji ARAMAKI, Ph.D. @ NAIST
Yuta NAKAMURA, Ph.D., M.D. @ The University of Tokyo

Organizers

Co-chair (general)

Eiji Aramaki, Ph.D. (NAIST, Japan)

Co-chair (general)

Shoko Wakamiya, Ph.D. (NAIST, Japan)

Co-chair (SM Subtask)

Shuntaro Yada, Ph.D. (NAIST, Japan)

Co-chair (RR Subtask)

Yuta Nakamura, M.D. (The University of Tokyo, Japan)

SM Subtask

Aurélie Névéol, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Patrick Paroubek, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Hui-Syuan Yeh (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Pierre Zweigenbaum, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Gabriel Herman Bernardim Andrade (NAIST, Japan)

SM Subtask

Faith Wavinya Mutinda, Ph.D. (NAIST, Japan)

SM Subtask

Tomohiro Nishiyama (NAIST, Japan)

SM Subtask

Lisa Raithel (DFKI, Germany, TU Berlin, Germany, and Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Akiko Aizawa, Ph.D. (NII, Japan)

RR Subtask

Shouhei Hanaoka, M.D., Ph.D. (The University of Tokyo, Japan)

SM Subtask

Yuji Matsumoto, Ph.D. (RIKEN, Japan)

SM Subtask

Noriki Nishida, Ph.D. (RIKEN, Japan)

SM Subtask

Roland Roller, Ph.D. (DFKI, Germany)

SM Subtask

Philippe Thomas, Ph.D. (DFKI, Germany)

SM Subtask

Cyril Grouin, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Thomas Lavergne, Ph.D. (Université Paris-Saclay, CNRS, LISN, France)

SM Subtask

Hiroki Teranishi, Ph.D. (RIKEN, Japan)

SM Subtask

Narumi Tokunaga (RIKEN, Japan)

SM Subtask

Lis Weiji Kanashiro Pereira Ph.D. (NAIST, Japan)

SM Subtask

Peitao Han (NAIST, Japan)

MedNLP-SC Subtasks

- Social Media Adverse Drug Event detection (SM-ADE)
  - Identify a set of symptoms caused by a drug from short messages written by social media users
  - Social media corpus in Japanese, English, German, and French

Co-chair (general)

Eiji Aramaki, Ph.D. (NAIST, Japan)

- Radiology Report TNM staging (RR-TNM)
  - Determine the clinical stage of lung cancer from radiology reports, which requires clinical knowledge and complex reasoning
  - Radiology report corpus in Japanese

Co-chair (RR Subtask)

Yuta Nakamura, M.D. (The University of Tokyo, Japan)

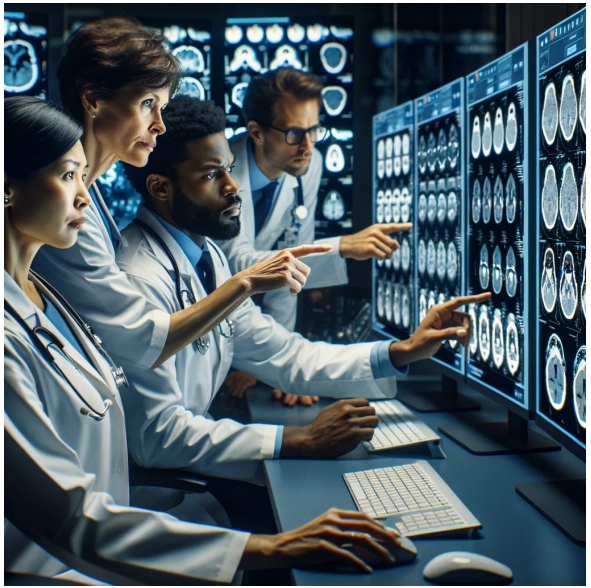
RR-TNM Subtask

Subtask for Radiology Report (RR)

RR-TNM subtask

Radiology Report

Clinical documents from radiologists to physicians

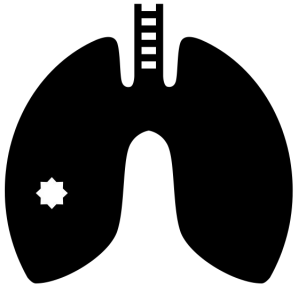


(These illustrations were created by GPT-4V)

RR-TNM subtask

TNM staging of lung cancer

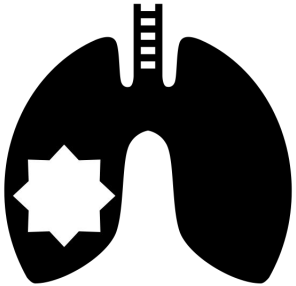
Precise evaluation of cancer progression (staging) is essential



Early cancer

Surgery applicable

Different treatment

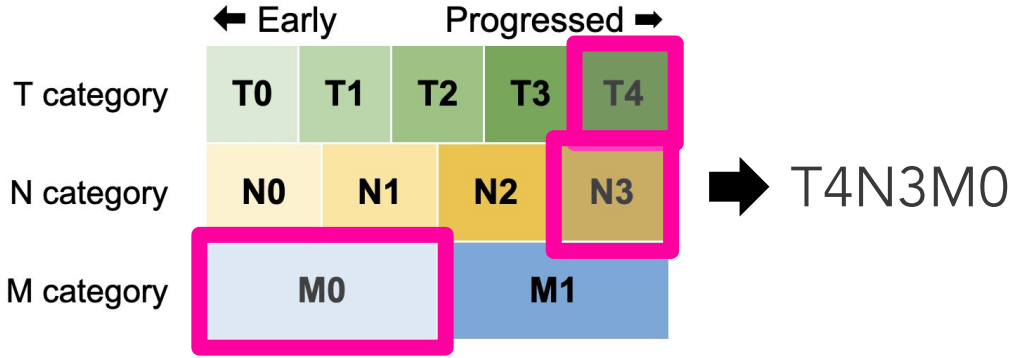


Progressed cancer

Surgery not applicable

Motivation

- Staging is complex
  - Three-label classification
  - Complicated criteria



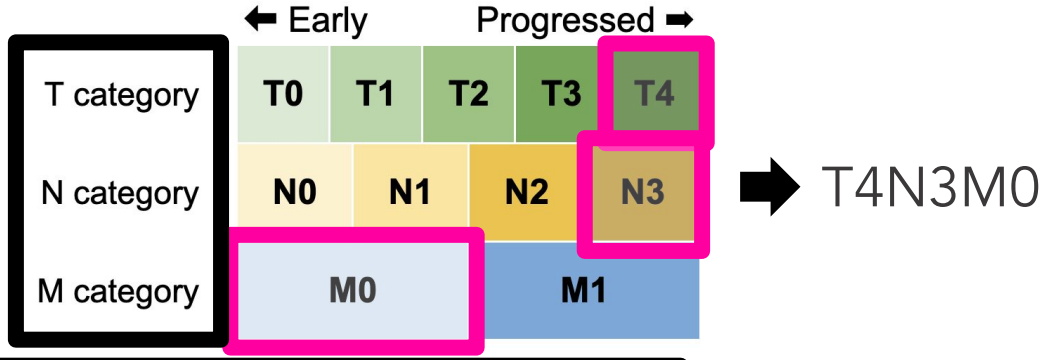
T0: No primary tumor
Tis: Ground-glass nodule without solid component with the total diameter ≤3 cm
T1mi: Ground-glass nodule with solid component ≤0.5 cm and the total diameter ≤3 cm
T1a: Solid component diameter ≤1 cm
T1b: Solid component diameter >1 cm and ≤2 cm
T1c: Solid component diameter >2 cm and ≤3 cm
T2a: Solid component diameter >3 cm and ≤4 cm. Otherwise, extension to main bronchus or visceral pleura, or atelectasis or obstructive pneumonia extending to hilum," with the solid component diameter <3 cm or unknown
T2b: Solid component diameter >4 cm and ≤5 cm
T3: Solid component diameter >5 cm and ≤7 cm. Otherwise, solid component diameter ≤5 cm and either condition holds: direct invasion of parietal pleura, chest wall (including superior sulcus tumor), mediastinal nerve, or pericardium; separate tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm. Otherwise, either condition holds: invasion of diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, spine, or carina; tumor nodule(s) in a different ipsilateral lobe

N0: No regional lymph node metastasis
N1: Metastasis to ipsilateral peribronchial, hilar, or pulmonary lymph nodes, including direct invasion of the primary tumor
N2: Metastasis to ipsilateral mediastinal or subcarinal lymph nodes
N3: Metastasis to contralateral mediastinal, hilar, anterior scalene, or supraclavicular lymph nodes

M0: No distant metastasis
M1a: Contralateral tumor nodule(s), pleural or pericardial nodule(s), malignant pleural effusion, or malignant pericardial effusion
M1b: Single extrathoracic metastasis
M1c: Multiple extrathoracic metastases

Motivation

- Staging is complex
  - Three-label classification
  - Complicated criteria



T: primary tumor
- Tumor size in mm?
- Invading to where?
- How many satellite lesions?

N: metastasis to lymph nodes
- To what lymph nodes?

M: metastasis to distant organs
- No/Single/multiple?

T0: No primary tumor
Tis: Ground-glass nodule without soli
T1mi: Ground-glass nodule with solid
cm
T1a: Solid component diameter ≤1 cr
T1b: Solid component diameter >1 cr
T1c: Solid component diameter >2 cr
T2a: Solid component diameter >3 cr
bronchus or visceral pleura, or atelect
hilum," with the solid component dia
T2b: Solid component diameter >4 cr
T3: Solid component diameter >5 cm
diameter ≤5 cm and either condition
wall (including superior sulcus tumor),
tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm
diaphragm, mediastinum, heart, great
esophagus, spine, or carina; tumor no

e metastasis
peribronchial, hilar, or
cluding direct invasion of

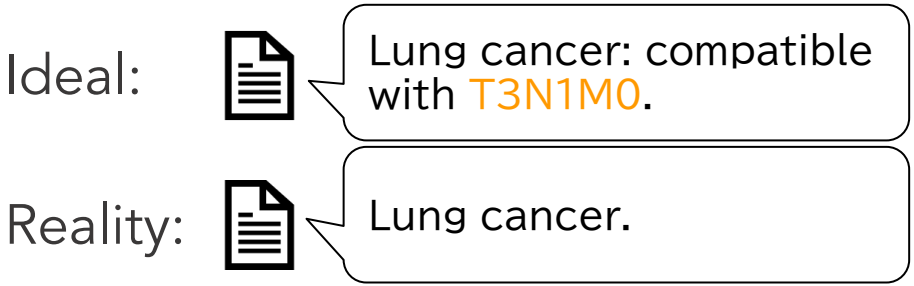
mediastinal or subcarinal

eral mediastinal, hilar,
avicular lymph nodes

odule(s), pleural or
nant pleural effusion, or
on
etastasis
metastases

Motivation

- TNM stage is not often explicitly mentioned in radiology reports (Sexauer et al., 2018)



- Patient care or secondary usage of information can be hampered

- We aim to evaluate the capability of NLP systems to automatically assign TNM stage based on radiology reports

Task scheme

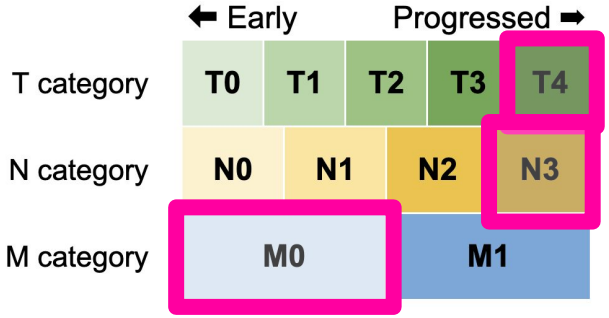
- Three-label document classification

Input
(lung cancer radiology report)

左上下葉に広がる長径 12cm の腫瘤を認めます。既知肺癌が示唆されます。 胸膜に広範囲に接しており左第 3 肋骨の破壊を伴っています。肋骨、壁側胸膜浸潤を疑います。左上葉に小結節あり、副腫瘍結節を疑います。 左縦隔、両側肺門部リンパ節が腫大、転移を疑います。 胸水は認めません。 撮像範囲の上腹部臓器に明らかな異常は認めません。

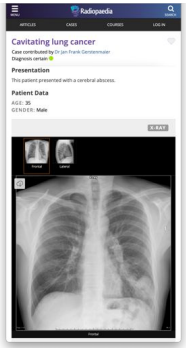
Answer
(clinical stage)

T4N3M0

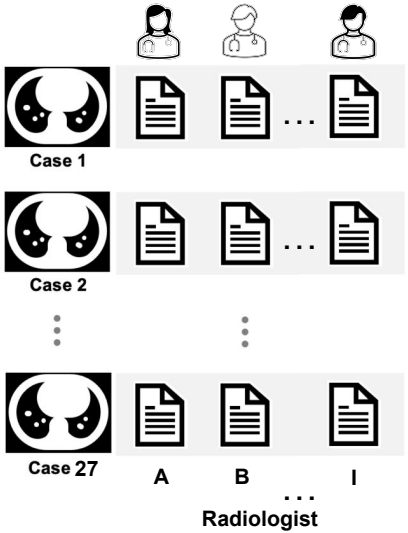


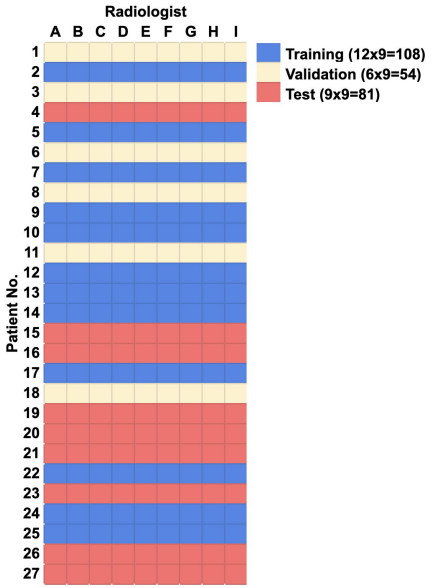
Dataset

- 243 Japanese radiology reports
- Contains no personal information
- 9 radiologists diagnosed 27 cases
- Case-level split (training : validation : test = 12:6:9 cases)
- Gold standards were assigned by the radiologists writing the radiology reports



(https://radiopaedia.org/cases/cavitating-lung-cancer-1)

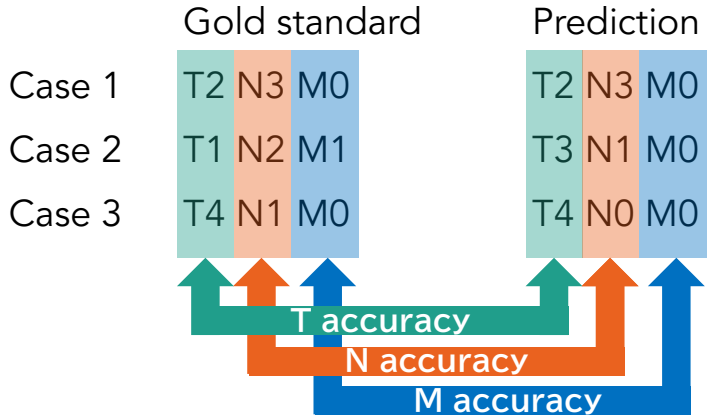




Evaluation metrics

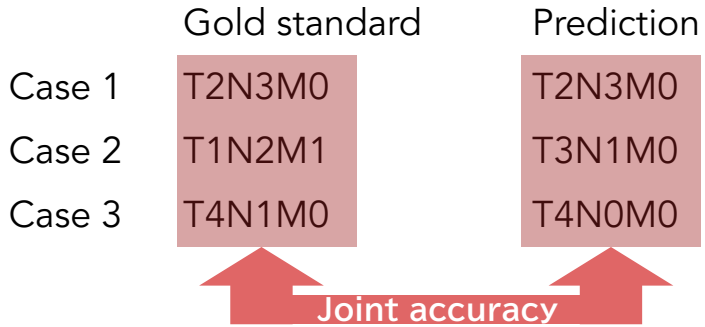
Separate evaluation: T/N/M accuracy

Accuracy scores calculated independently for T, N, and M categories



Joint accuracy

Fraction of cases whose T, N, and M categories were all predicted correctly



Potential challenge

- Complicated criteria requiring detailed domain-specific knowledge

T0: No primary tumor
Tis: Ground-glass nodule without solid component with the total diameter ≤3 cm
T1mi: Ground-glass nodule with solid component ≤0.5 cm and the total diameter ≤3 cm
T1a: Solid component diameter ≤1 cm
T1b: Solid component diameter >1 cm and ≤2 cm
T1c: Solid component diameter >2 cm and ≤3 cm
T2a: Solid component diameter >3 cm and ≤4 cm. Otherwise, extension to main bronchus or visceral pleura, or atelectasis or obstructive pneumonia extending to hilum," with the solid component diameter <3 cm or unknown
T2b: Solid component diameter >4 cm and ≤5 cm
T3: Solid component diameter >5 cm and ≤7 cm. Otherwise, solid component diameter ≤5 cm and either condition holds: direct invasion of parietal pleura, chest wall (including superior sulcus tumor), mediastinal nerve, or pericardium; separate tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm. Otherwise, either condition holds: invasion of diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, spine, or carina; tumor nodule(s) in a different ipsilateral lobe

N0: No regional lymph node metastasis
N1: Metastasis to ipsilateral peribronchial, hilar, or pulmonary lymph nodes, including direct invasion of the primary tumor
N2: Metastasis to ipsilateral mediastinal or subcarinal lymph nodes
N3: Metastasis to contralateral mediastinal, hilar, anterior scalene, or supraclavicular lymph nodes

M0: No distant metastasis
M1a: Contralateral tumor nodule(s), pleural or pericardial nodule(s), malignant pleural effusion, or malignant pericardial effusion
M1b: Single extrathoracic metastasis
M1c: Multiple extrathoracic metastases

Participant solutions

7 solutions from 3 teams (including baseline)

- KRad team (Nishio et al.)
  - LLM gpt-3.5-turbo + in-context learning
- kuhp team (Fujimoto et al.)
  - LLM openCALM-7B + instruction tuning
- NAISTSOCRR team (Fukushima et al.)
  - smaller model BERT/RoBERTa + fine-tuning
  - Majority baseline (always answer T1N0M0)

Participant solutions

7 solutions from 3 teams (including baseline)

- KRad team (Nishio et al.)
  - LLM gpt-3.5-turbo + in-context learning
- kuhp team (Fujimoto et al.)
  - LLM openCALM-7B + instruction tuning
- NAISTSOCRR team (Fukushima et al.)
  - smaller model BERT/RoBERTa + fine-tuning
  - Majority baseline (always answer T1N0M0)

Presenting criteria in the prompt

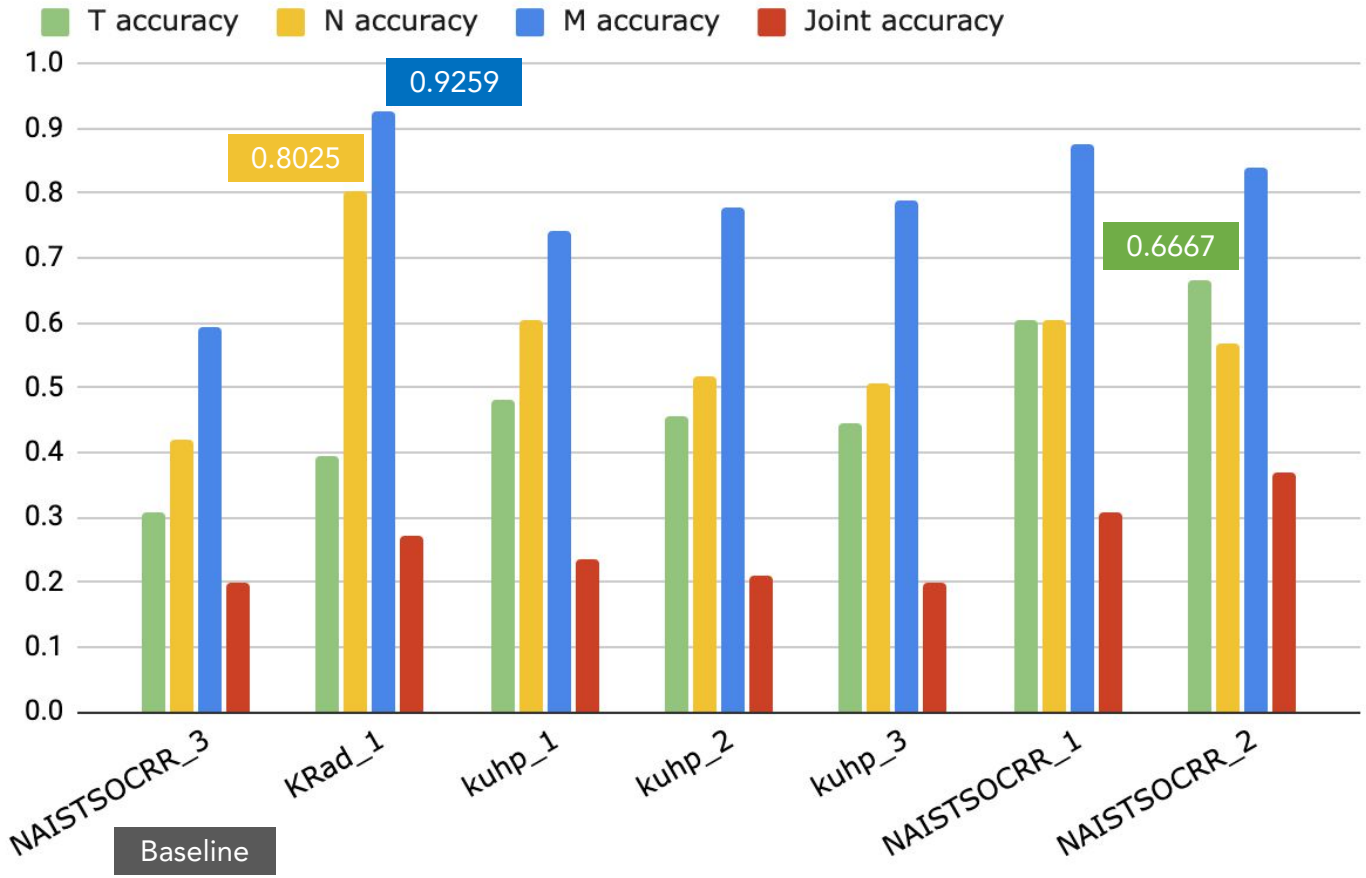
Custom QA dataset for instruction tuning + data augmentation

Pre-trained models in biomedical domain

Subtask for Radiology Report (RR)

Results

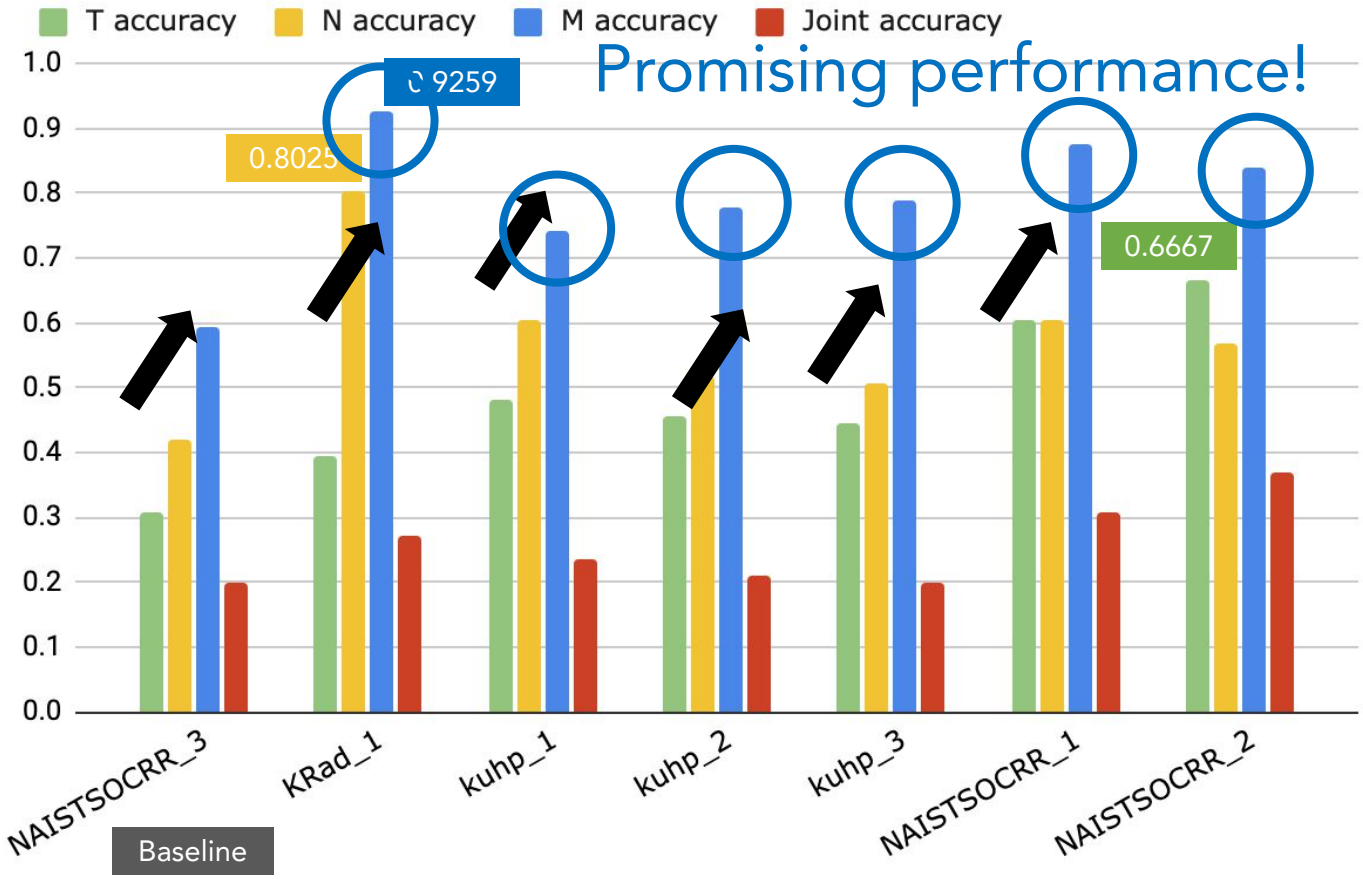
All solutions outperformed baseline



Results

All solutions outperformed baseline

Tendency:
T < N < M

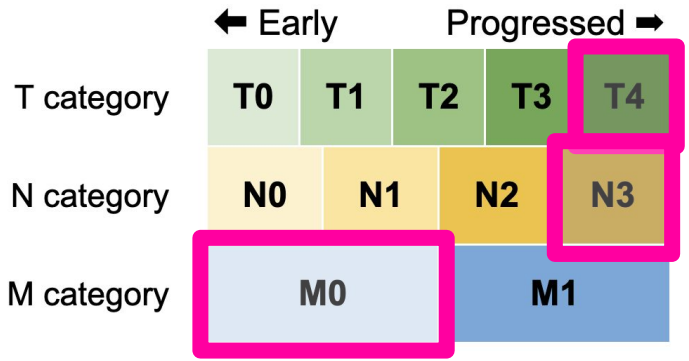


Results

	All teams succeeded 😀	All teams failed 😢
T category	1/81	9/81
N category	20/81	6/81
M category	38/81	5/81

Why T<N<M?

Class:
T(5) > N(4) > M(2)



Complexity of criteria:
T >> N, M

T: primary tumor
- Tumor size in mm?
- Invading to where?
- How many satellite lesions?

N: metastasis to lymph nodes
- To what lymph nodes?

M: metastasis to distant organs
- No/Single/multiple?

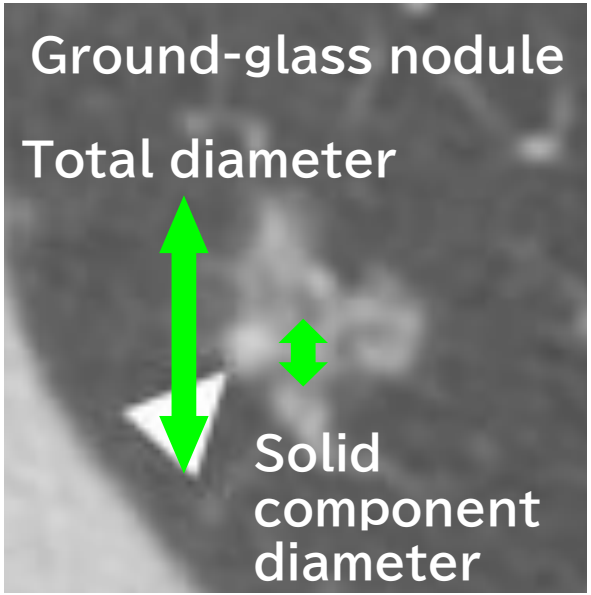
T criteria is complex

T: primary tumor
- Tumor size in mm?
- Invading to where?
- How many satellite lesions?

T0: No primary tumor
Tis: Ground-glass nodule without solid component with the total diameter ≤3 cm
T1mi: Ground-glass nodule with solid component ≤0.5 cm and the total diameter ≤3 cm
T1a: Solid component diameter ≤1 cm
T1b: Solid component diameter >1 cm and ≤2 cm
T1c: Solid component diameter >2 cm and ≤3 cm
T2a: Solid component diameter >3 cm and ≤4 cm. Otherwise, extension to main bronchus or visceral pleura, or atelectasis or obstructive pneumonia extending to hilum," with the solid component diameter <3 cm or unknown
T2b: Solid component diameter >4 cm and ≤5 cm
T3: Solid component diameter >5 cm and ≤7 cm. Otherwise, solid component diameter ≤5 cm and either condition holds: direct invasion of parietal pleura, chest wall (including superior sulcus tumor), mediastinal nerve, or pericardium; separate tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm. Otherwise, either condition holds: invasion of diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, spine, or carina; tumor nodule(s) in a different ipsilateral lobe

T criteria is complex

T: primary tumor
- Tumor size in mm?
- Invading to where?
- How many satellite lesions?



Mark M. Hammer et al., 2020

T0: No primary tumor
Tis: Ground-glass nodule without solid component with the total diameter ≤3 cm
T1mi: Ground-glass nodule with solid component ≤0.5 cm and the total diameter ≤3 cm
T1a: Solid component diameter ≤1 cm
T1b: Solid component diameter >1 cm and ≤2 cm
T1c: Solid component diameter >2 cm and ≤3 cm
T2a: Solid component diameter >3 cm and ≤4 cm. Otherwise, extension to main bronchus or visceral pleura, or atelectasis or obstructive pneumonia extending to hilum," with the solid component diameter <3 cm or unknown
T2b: Solid component diameter >4 cm and ≤5 cm
T3: Solid component diameter >5 cm and ≤7 cm. Otherwise, solid component diameter ≤5 cm and either condition holds: direct invasion of parietal pleura, chest wall (including superior sulcus tumor), mediastinal nerve, or pericardium; separate tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm. Otherwise, either condition holds: invasion of diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, spine, or carina; tumor nodule(s) in a different ipsilateral lobe

T criteria is complex

T: primary tumor
- Tumor size in mm?
- Invading to where?
- How many satellite lesions?

Requires anatomical understanding

Invading ribs ⊂
Invading chest wall

T0: No primary tumor
Tis: Ground-glass nodule without solid component with the total diameter ≤3 cm
T1mi: Ground-glass nodule with solid component ≤0.5 cm and the total diameter ≤3 cm
T1a: Solid component diameter ≤1 cm
T1b: Solid component diameter >1 cm and ≤2 cm
T1c: Solid component diameter >2 cm and ≤3 cm
T2a: Solid component diameter >3 cm and ≤4 cm. Otherwise, extension to main bronchus or visceral pleura, or atelectasis or obstructive pneumonia extending to hilum," with the solid component diameter <3 cm or unknown
T2b: Solid component diameter >4 cm and ≤5 cm
T3: Solid component diameter >5 cm and <7 cm. Otherwise, solid component diameter ≤5 cm and either condition holds: direct invasion of parietal pleura, chest wall (including superior sulcus tumor), mediastinal nerve, or pericardium; separate tumor nodule(s) in the same lobe
T4: Solid component diameter >7 cm. Otherwise, either condition holds: invasion of diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, spine, or carina; tumor nodule(s) in a different ipsilateral lobe

T criteria is complex

Recall scores tended to

- Higher in early (T1–T2) or very progressed (T4) cancers
- Lower in T3 cancers

	KRad_1	NAISTSOCRR_1	NAISTSOCRR_2	baseline_3	kuhp_1	kuhp_2	kuhp_3
T0	0.500000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000
T1	0.040000	0.720000	1.000000	1.0	0.400000	0.520000	0.520000
T2	0.666667	0.000000	0.333333	0.0	0.333333	0.500000	0.500000
T3	0.400000	0.000000	0.000000	0.0	0.200000	0.200000	0.200000
T4	0.560976	0.756098	0.658537	0.0	0.634146	0.487805	0.463415
	LLM	BERT	BERT		LLM	LLM	LLM

Indeterminate mentions

Presence/absence of image findings are often difficult to determine due to indeterminate mentions in radiology reports

😢 A test sample failed by all teams for N category prediction

- 「右肺門リンパ節がやや腫大しており、**転移を否定できません**。」
- Moderate right hilar lymphadenopathy is noted: metastasis cannot be ruled out.
  - ✅ N1 (lymph node metastasis is present)
  - ❌ N0 (lymph node metastasis is absent)

Future direction

- Further exploration for better performance is valuable for real-world applications

Remaining issues

- This subtask was limited to Japanese data
- How to create open radiology report datasets to non-Japanese languages?

Acknowledgement

We greatly appreciate Y's Reading, Inc., a teleradiology company in Kumamoto, Japan, for continuous support in creating an open Japanese radiology report corpus