Supplemental materials for

Ellertsson S, Hlynsson HD, Loftsson H, Sigurdsson EL. triaging patients with artificial intelligence for respiratory symptoms in primary care to improve patient outcomes: a retrospective diagnostic accuracy study. *Ann Fam Med*. 2023;21(3):240-248.

Supplemental Table 1. Presenting complaints used to filter patients into the dataset

Cough	Fever	Sore throat
Common cold symptom	Respiratory symptoms	Headache
Sputum excretion	Nasal congestion	Muscle / bone pain
Blood in sputum	Rhinorrhea	Chest pain
Chest tightness	Dyspnea	Thoracal pain
Pleural pain	Chills / shivers	Flu-like symptoms
Malaise	Pain in ear	

Supplemental Appendix

Data annotation

The annotation method was inspired by a group of researchers who applied similar annotations on medical text¹⁹ and its purpose is to assign binary and numerical values to clinical variables, representing the existence or lack of specific signs or symptoms in the CTNs, as they are textually referenced in the CTNs. Each annotation is formed as a question-answer pair where the question is stated to ask about a specific clinical variable and the answer is annotated depending on the type of variable being annotated. As an example, consider a frequent piece of text description in CTNs: 'The patient has a cough and a sore throat but no fever'. Two questionanswer pairs regarding the patient's cough and sore throat are annotated as positive and the third one, regarding the fever, is annotated as negative. Numerical values variables were assigned a number value, for example, a variable for temperature (e.g. with the value 38.2°C). The annotator was blinded to the diagnoses. Missing binary variables were given the value of 0. Missing value variables were replaced by randomly sampling the normal distribution for that variable to reduce the odds of the model simply learning where value variables are missing which would be more likely for a patient with a less severe disease. Supplemental Table 2. Clinical features included in the RSTM

Binary clinical features (positive or negative)		Value clinical features
Sex (1 for females, 0 for men)	History of asthma	Heart rate
Sputum excretion	History of COPD	Age
Blood in sputum excretion	History of hypertension	Respiratory frequency
Purulent / colored sputum	History of previous smoking	Temperature
Sore throat	Active smoking	Temperature measured by patient during course of symptoms
Cough	Dizziness	Symptom duration (days)
Cough disturbs sleep	Malaise	Symptom duration (weeks)
Dyspnea	Chest tightness	Symptom duration (months)
Back pain	Thoracal pain	Known allergy
Chest pain	Thoracal pain appearing or worsening when coughing	Smoking - years since quitting
Nasal congestion	Thoracal pain appearing or worsening when taking a deep breath	Smoking - number of years
Rhinorrhea	Headache	Smoking - package years
History of allergy	Headache frontal location	
Hoarseness	Flu-like symptoms	
Chills / shivers	Fever at any time during symptom course	
Muscle / bone pain	Common cold symptoms	
Nausea	Pleural pain	
Vomiting	Facial pressure / pain	
Pain in single / both ears	Diarrhea	

Supplemental Figure 1. The mean score from our RSTM compared to the score from the diagnostic rule from Diehr et al, compared for each risk group.



Note: The score from Diehr et al. is normalized to fit the same range as the score from the RSTM, ie. from zero to one.

Supplemental Figure 2. The Anthonisen score for each patient, with J44 ICD-10 code (COPD), stratified by risk group.



Note: Anthonisen et al showed benefit with antibiotic treatment for COPD when 2-3 out of 3 cardinal symptoms are present. The cardinal symptoms are increased sputum production, purulent sputum production and increased dyspnea. No benefit was shown for patients with scores 1 or lower.