BMJ Open

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-079123
Article Type:	Original research
Date Submitted by the Author:	22-Aug-2023
Complete List of Authors:	Cartuliares, Mariana; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Mogensen, Christian Backer; University of Southern Denmark, Institute for Regional Health Research; University Hospital of Southern Denmark, Department of Emergency Medicine Rosenvinge, Flemming; Odense Universitetshospital, Department of Clinical Microbiology; University of Southern Denmark, Research Unit of Clinical Microbiology Skovsted, Thor; University Hospital of Southern Denmark, Department of Biochemistry and Immunology Lorentzen, Morten; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Heltborg, Anne; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Hertz, Mathias ; University of Southern Denmark, Department of Clinical Research; Odense University Hospital, Infectious Diseases Department Kaldan, Frida; University Hospital of Southern Denmark, Department of Emergency Medicine Specht, Jens ; University Hospital of Southern Denmark, Department of Emergency Medicine Skjøt-Arkil, Helene; University Hospital of Southern Denmark, Department of Emergency Medicine; University of Southern Denmark, Department of Regional Health Research
Keywords:	ACCIDENT & EMERGENCY MEDICINE, INFECTIOUS DISEASES, Aged

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

TITLE

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Authors

Mariana Bichuette Cartuliares* [1, 2], Christian Backer Mogensen [1, 2], Flemming Schønning Rosenvinge [3, 4], Thor Aage Skovsted [5], Morten Hjarnø Lorentzen [1,2], Anne Heltborg[1,2], Mathias Amdi Hertz [6,7], Frida Kaldan[1], Jens Juel Specht[1], Helene Skjøt-Arkil [1,2]

*Corresponding author: Emergency Department, University Hospital of Southern Denmark, Kresten Philipsens vej 15, 6200 Aabenraa, Denmark; Email: mbc@rsyd.dk

Author affiliations

[1]Emergency Department, University Hospital of Southern Denmark, Aabenraa, Denmark [2]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark [3]Research Unit of Clinical Microbiology, University of Southern Denmark, Odense, Denmark [4]Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark [5]Department of Biochemistry and Immunology, University Hospital of Southern Denmark, Aabenraa, Denmark [6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark [7] Infectious Diseases Department, Odense University Hospital, Odense, Denmark

Strength and limitations

- This is a multicentre study with prospectively collected data
- Least absolute shrinkage and selection operator regression was used to establish a score for community-acquired pneumonia, and the performance of the diagnostic model was evaluated using the area under the receiver operating characteristic curve and calibration curves.

- This diagnostic prediction model could be improved in the future by adding diagnostic tools such as imaging or serological markers.
- External validation of the model using the clinical score for community-acquired pneumonia is lacking.

ABSTRACT

Objectives: This study aimed to describe the clinical characteristics of adults with acute community-acquired pneumonia (CAP) upon hospitalisation, evaluate their prediction performance for CAP and compare the performance of the model to the initial assessment of the physician.

Design: Cross-sectional, prospective, multicentre study.

Setting: The data originates from the INfectious DisEases in Emergency Departments study and were collected prospectively from patient interviews and medical records. The study included four Danish medical emergency departments (EDs) and was conducted between 1 March 2021 to 28 February 2022.

Participants: A total of 954 patients admitted with suspected infection were included in the study.

Primary and secondary outcome: The primary outcome was CAP diagnosis assessed by an expert panel.

Results: According to expert evaluation, CAP had a 28% prevalence. Thirteen diagnostic predictors were identified using Least absolute shrinkage and selection operator regression to build the prediction model: dyspnea, expectoration, cough, common cold, malaise, chest pain, respiratory rate (>20/min), oxygen saturation (< 96%), abnormal chest auscultation, leucocytes (<3,5 or >8,8 10E9/L) and neutrophilocytes (>7.5 10E9/L). In addition, C-reactive protein (<20 mg/L) and having no previous event of CAP contributed negatively to the final model. The predictors yielded good prediction performance for CAP with an area under the ROC of 85% with a sensitivity of 86% (79%-93%) and specificity of 64% (57%-71%) using a 35% cut-off. However, the initial diagnosis made by the ED physician performed better, with 86% (84%-89%) sensitivity and 75% (72%-78%) specificity.

Conclusion: Typical respiratory symptoms combined with abnormal vital signs and elevated infection biomarkers were predictors for CAP upon admission to an ED. The clinical value of the prediction model is questionable in our setting. Further studies adding novel diagnostic tools and using imaging or serological markers are needed to improve the model, helping diagnose CAP in an ED setting more accurately.

Keywords: community-acquired pneumonia; diagnostic prediction model; emergency department

Word count: 3.771

INTRODUCTION

Community-acquired pneumonia (CAP) is an increasing cause of hospitalisation and mortality, especially among elderly patients [1-5]. Early diagnosis and accurate treatment at the emergency department are essential to avoid serious complications such as bacteremia, sepsis, organ failure, and death [6] and to fight antimicrobial resistance [7].

Traditionally, the diagnosis of CAP generally requires a new infiltrate on chest x-ray with a clinically compatible syndrome [8]. These symptoms aren't sufficient to diagnose or exclude CAP, as they overlap with other diseases [8] and can be subtle in patients with advanced age and/or impaired immune systems [9, 10]. Chest x-ray is imprecise as diagnostic tool for CAP, risking under/over diagnosis [11, 12] and might not the optimal reference standard for CAP. This variability of clinical signs and symptoms combined with non-specific diagnostic tools [12], biomarkers [13, 14], and time-consuming microbiological tests [9] challenges physicians in differentiating CAP from other infections [10, 15].

The CAP population today has also changed with the increasing ageing [16], higher multimorbidities [17], and immunomodulatory treatments. Our knowledge of CAP symptoms and signs therefore need to be adapted to the actual population.

Previously, prediction models for the diagnosis of CAP have been developed based primarily on prognostic factors including severity assessment [18, 19], observations in a primary care setting only [20-22], or an outcome diagnosis based solely on the registered discharge diagnosis in the medical record or positive chest x-ray findings [22, 23]. A valid outcome diagnosis is essential. An expert panel using several available information might be the best reference standard in pragmatic studies [11].

Therefore, there is a need to describe clinical characteristics of the current population of patients admitted with CAP and develop an improved diagnostic model to be used upon arrival at the emergency room that include physical examination, blood tests, vital signs, patient medical history, and healthcare expertise. Given the current diagnostic tool inaccuracies, an expert-panel-based diagnostic model is expected to surpass the ED physicians' initial accuracy.

Hypothesis and objectives

We hypothesised that developing of a diagnostic prediction model using well-defined clinical characteristics could assist an ED physician in an earlier, more precise CAP diagnosis. Therefore, the aim was to identify the clinical characteristics of adults admitted with CAP and evaluate their performance in a prediction model.

The objectives were:

1) To investigate clinical characteristics in patients with a CAP diagnosis from i) all patients admitted with suspected infection and ii) patients suspected of CAP
2) To develop and evaluate a diagnostic model to identify patients with CAP among ED patients suspected of infection and to compare the performance of the model to the initial assessment of the ED physician

METHODS

The study was reported following "The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis" (TRIPOD) statement [24] and conducted in agreement with the Declaration of Helsinki-Ethical principle for medical research involving human subjects. The protocol was approved by the Regional Committee on Health Research Ethics for Southern Denmark (S- 20200188), registered by the Danish Data Protection Agency (no. 20/60508), and by ClinicalTrials.gov (NCT04681963).

Study design, source of data, and setting

This study had a prospective, analytical cross-sectional, multicentre design. The data originates from the INfectious DisEases in Emergency Departments (INDEED) study. The published study protocol provides further detailed information [25]. Four Danish medical EDs participated, covering around 775,000 inhabitants, during March 1, 2021, to February 28, 2022.

In Denmark, patients can be directed to various specialties within the ED, e.g. medical, gastrointestinal surgery, cardiology, orthopedics, gynecology, psychiatry, and intensive care [26]. Suspected infection cases usually are assigned to the medical ED.

Participants

Adult patients (≥ 18 years) admitted to the medical ED were eligible to participate. Patients were included if the ED physician suspected infection and if the patients could provide verbal and written consent. The exclusion criteria included: i) need for urgent, life-saving treatment, ii) transferal to intensive care, iii) admission within the last fortnight, iv) verified SARS-CoV-2 infection at the time of admission or within 14 days before admission, v) severe immunodeficiencies (HIV positive, with a cluster of differentiation 4 cell count <200) or treatment with immunosuppressive medicine (Anatomical Therapeutic Chemical classification L04A), corticosteroids (>20 mg/day prednisone or equivalent for >14 days within the last 30 days) or chemotherapy within 30 days.

Recruitment and data collection

Six project assistants with a healthcare background (three physicians, one physiotherapist, and two final-year medical students) were responsible for inclusion and data collection from Mondays to Fridays, 8 am to 8 pm. A project assistant consecutively identified eligible patients from the patient management system. Immediately following the initial clinical assessment, the project assistant asked the ED physician whether an infection was suspected and the most likely infection focus (CAP, urinary tract infection, or unknown origin). Generally, the clinical assessment took place within 30 minutes upon admission before blood tests or imaging was ordered, and therefore, the ED physician often had only information on the patient's signs, symptoms, and vital parameters. The study assistant collected verbal and written consent from eligible patients. All data collected was registered in the electronic study database REDCap (Research Electronic Data Capture) [27].

Outcome

The outcome was the diagnosis of CAP. An expert panel was established consisting of pairs of experienced infectious diseases and emergency medicine specialists at each site. They conducted a patient file audit and determined the final diagnosis based on all clinical information registered within the first week of ED admission. The information included routine laboratory tests of blood, -urine, and -sputum. In addition, polymerase chain reaction test of sputum, urine flow cytometry, chest x-ray, and chest computed tomography (CT) were available for some patients. The experts were blinded to each other and independently registered their assessments in a standardized electronic template [27] in the study database. Disagreements were discussed until a consensus was reached.

Predictors

All clinical characteristics were collected upon arrival at the ED. Symptoms, demographic data, and lifestyle factors were registered during a standardised bedside interview with the patient. In addition, information

about vital parameters, comorbidities, medical treatment, and blood tests were collected from the patient's medical record. The project assistants collecting data were blinded to the final diagnosis.

Several candidate predictors (70) were selected from the literature and discussed with the specialists and project group [20, 28-37]. The pre-specified potential predictors with their measurement units, groups, cut-offs, and which considerations/assumptions of inclusion were selected and are described in Supplemental material, Supplementary Table S1.

- Demographic information, lifestyle factors, and comorbidities: age, sex, civil status, employment, nursing home residence, smoking, and alcohol consumption, body mass index (BMI), level of physical activity, activities of daily living score, dementia, respiratory, neurological, cardiovascular, endocrinological, nephrological and gastrointestinal comorbidities were collected.

-Patient symptoms the last two weeks before admission: malaise, fatigue, headache, dizziness, altered mental status, e.g. confusion, dyspnea, malnutrition, cough, secretions from the respiratory tract, sore throat, common cold, fever feeling, chest pain, peripheral oedema, nausea, vomiting, decreased appetite, abdominal pain, diarrhoea, and pain in muscles and joints including back pain were collected.

-Severity assessment, clinical parameters with cut-offs based on National Early Warning Score (NEWS) [38] used at the arrival of the ED and the use of medications: CURB-65 ≥3 (confusion, uremia, respiratory rate, blood pressure, age > 65 years), triage [39], Glasgow coma scale (GCS), oxygen saturation <96%, heart rate <51 or >90/min, blood pressure (systolic <111 or >219, diastolic ≤60 mmHg), respiratory rate >20/min, temperature > 38^{o}C, abnormal chest auscultation, abdominal tenderness, polypharmacy (≥ 5 medications), use of analgesics, and vaccination status (SARS-CoV-2, pneumococcus, influenza) were recorded.

-Blood tests with cut-offs routinely applied at our institutions: haematocrit (%), hemoglobin (mmol/L), leukocytes (10E9/L), platelets (10E9/L), neutrophils (10E9/L), lymphocytes (10E9/L), albumin g/L, creatinine (µmol/L), blood urea nitrogen (mmol/L), sodium (mmol/L), prothrombin, bilirubin (µmol), glucose (mmol/L), and CRP (mg/L) were recorded.

Statistical methods

The study sample size was estimated based on the University Hospital of Southern Denmark data. We estimated a need for at least 700 patients admitted with suspected infection. Of those, four hundred patients should be with suspected CAP and two hundred patients should have verified CAP for sufficient multivariable regression analysis. Descriptive statistics for baseline characteristics of the patients were conducted for the 70 potential predictors based on the data from the INDEED study [25]. Data were presented as means and standard deviations (SD), or medians and interquartile ranges (IQRs) for continuous variables, and numbers (n) and percentages (%) for categorical and binary variables. Extensive univariate logistic regression analyses were performed to examine the unadjusted association between each candidate predictor and the outcome CAP. Results of univariate analyses were reported with odds ratio (OR), 95% confidence intervals (CI), and statistical significance levels were two-sided reported with a p-value of <0.05 to present a descriptive overview of the individual's associations in the population. Complete case analyses were performed and the predictors were dichotomised or categorised and presented with percentages (%) for inclusion in the final model. The least absolute shrinkage and selection operator (LASSO) multivariable regression was performed with a random split-sample to develop and validate the model, using 20 % of the data for internal cross-validation. The model calibration was assessed using a likelihood ratio test, and recalibration was done based on the calibration belt and the optimal predicted proportion. In the model, age (≥75 years old) was considered as an effect modifier based on several studies showing differences in symptoms and signs for a CAP diagnosis in older adults [33, 40-42]. An exploratory approach was conducted for the clinical characteristics to achieve a model with the best predictive performance, testing their performance as continuous, dichotomous, or categorical variables. In addition, the receiver-operator characteristic (ROC) curve was created to estimate the model's accuracy, and the area under the ROC curve (AUC) visualized the discrimination between true positives and negatives. The sensitivity, specificity, and positive and negative predictive values with 95% CI were calculated using the best threshold criteria of the predicted probability of the ROC curve.The same

threshold was implemented in developing a CAP score, including the predictor variables. A CAP score> 0 represents the presence of CAP, and < 0 indicates the absence of CAP. Sensitivity, specificity, and positive and negative predictive values with 95% CI were calculated from the initial diagnosis made by the ED physician. Analyses were performed using STATA 17.0 (Texas, USA).

Patient and public involvement

Patients and/or the public were not directly involved in this study.

RESULTS

Participants

We recruited 954 patients admitted to the ED with suspected infection, representing 43% screened for eligibility. Of those, the attending physician suspected 402 (42%) had CAP. Patients with verified CAP diagnosis by the expert panel comprised of 265 (28%) of the recruited patients (Figure 1).

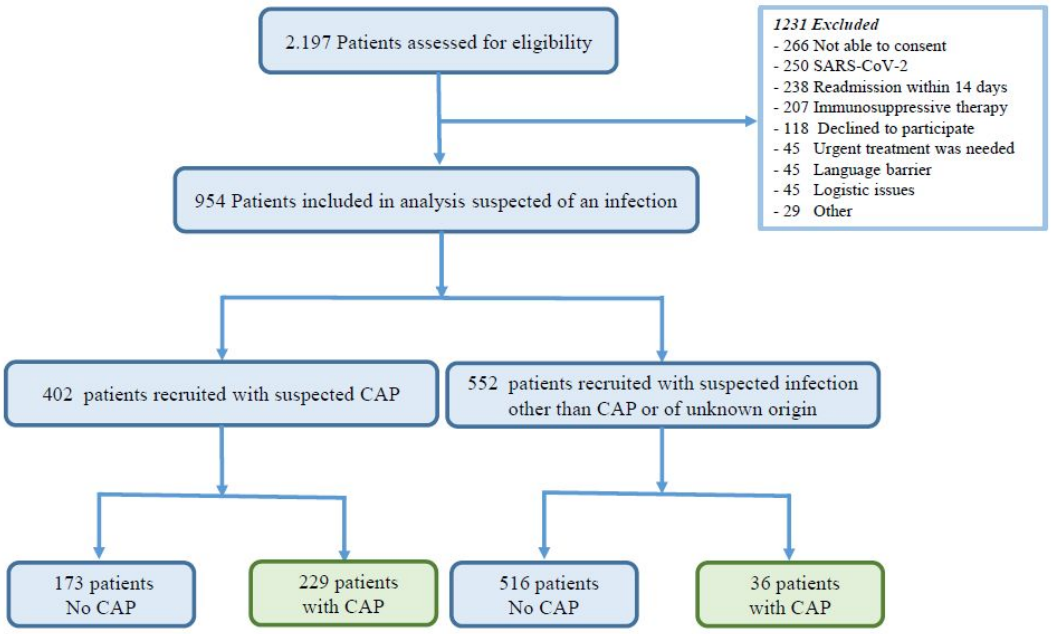


Figure 1: Trial population, green boxes showing the numbers of patients with CAP.

Characteristics of patients with suspected infections

We compared the clinical characteristics of patients with verified CAP to patients with suspected infection (954) without verified CAP. Median age for patients with verified CAP was 75 years (IQR: 63.5; 82.0), and over half admitted with suspected infection were males (53.8%). Univariate analysis revealed that verified CAP patients were more often previous smokers [OR 1.83 (CI: 1.30-2.57) p<0.001] with smoking history compared non-CAP cases. Strongly independent predictors for CAP were symptoms such as dyspnea, cough, expectoration, chest pain, and cold symptoms (all p<0.001). Compared to patients without CAP, the risk of having CAP increased fivefold if the patient had chest auscultation abnormalities [OR 5.67 (CI: 4.15-7.75) p<0.001] and decreased by half in case of abdominal tenderness by palpation [OR 0.52 (CI: 0.35-0.78) p=0.002]. CAP patients often had comorbidities related to other pulmonary diseases (p<0.001) and had more previous CAP infections (p<0.001). These patients were more acutely ill when assessed by triage (p<0.001), with fever > 38°C (p=0.036), higher respiratory rate [median 20.0 (IQR 18.0; 24.0) p<0.001], higher heart rate [mean 93.2 (SD 18.9) (p<0.001], and lower oxygen saturation [median 95.0 (IQR: 93.0; 97.0) p<0.001]. Patients with verified CAP had a median CRP of 125.0 (IQR: 57.0; 203.5) versus 82.0 (IQR: 19.0; 172.0) (p<0.001) compared to the rest of the population and higher levels of neutrophilocytes (p<0.001) and leucocytes (p<0.001). Furthermore, lymphocytes yielded a p-value of 0.018. Patients with verified CAP were more often vaccinated against SARS-CoV-2 (p=0.033) and influenza (p=0.025), but no differences were found regarding pneumococcal vaccination. Table 1 presents the characteristics of the population with statistically significant results of the unadjusted association between each predictor for patients with verified and not verified CAP. See Supplementary Table S2 for the 70 exploratory results from continuous, dichotomous, and categorical variables tested in the diagnostic prediction model.

Table 1: Characteristics of the population with suspected infection (n=954).

Characteristics	Patients suspected of infection at admission		Missings n (%)	OR (95% CI)	p-value
	CAP n (%)	Not CAP n (%)			
Total of patients	265 (27.8)	689 (72.2)	0 (0.0)	-	-
LIFESTYLE FACTORS					

Smoking status			33 (3.5)		
No	66 (26.0)	257 (38.5)		1 (reference)	
Current smoker	54 (21.3)	125 (18.7)		1.68 (1.10-2.55)	0.015
Previous smoker	134 (52.8)	285 (42.7)		1.83 (1.30-2.57)	<0.001
SYMPTOMS					
Malaise	173 (67.8)	386 (58.7)	41 (4.3)	1.48 (1.09-2.01)	0.010
Dyspnea	171 (67.3)	208 (31.5)	39 (4.1)	4.48 (3.29-6.11)	<0.001
Cough	173 (68.1)	185 (28.0)	39 (4.1)	5.49 (4.01-7.52)	<0.001
Expectoration	140 (55.1)	139 (21.0)	39 (4.1)	4.61 (3.38-6.28)	<0.001
Sore throat	39 (15.4)	65 (9.8)	39 (4.1)	1.66 (1.08-2.54)	0.019
Common cold	45 (17.7)	50 (7.6)	39 (4.1)	2.63 (1.70-4.05)	<0.001
Chest pain	71 (28.1)	97 (14.7)	40 (4.2)	2.26 (1.60-3.21)	<0.001
Oedema	10 (4.0)	69 (10.4)	40 (4.2)	0.35 (1,17-0.69)	0.002
Vomiting	40 (15.8)	150 (22.6)	38 (4.0)	0.64 (0.43-0.94)	0.023
Gastrointestinal pain	40 (15.8)	153 (23.1)	38 (4.0)	0.62 (0.42-0.91)	0.016
Muscular pain	79 (31.3)	265 (40.3)	44 (4.6)	0.67 (0.49-0.92)	0.013
COMORBIDITIES					
Pulmonary diseases	105 (39.6)	164 (23.8)	0 (0.0)	2.10 (1.55-2.84)	<0.001
Prior pneumonia			100 (10.5)		
No	79 (33.3)	331 (53.6)		1 (reference)	
Yes, one time	50 (21.1)	130 (21.1)		1.61 (1.07-2.42)	0.022
Yes, more than one time	108 (45.6)	156 (25.3)		2.90 (2.05-4.10)	<0.001
VACCINATIONS					
SARS-CoV-2 †	222 (83.8)	534 (77.5)	0 (0.0)	1.49 (1.03-2.17)	0.033
Influenza	191 (72.1)	444 (64.4)	0 (0.0)	1.42 (1.04-1.94)	0.025
CLINICAL ASSESSMENT					
Abnormal chest auscultation*	168 (65.4)	161 (25.0)	52 (5.4)	5.67 (4.15-7.75)	<0.001
Abdominal tenderness	37 (15.0)	155 (25.0)	86 (9.0)	0.52 (0.35-0.78)	0.002
SEVERITY ASSESSMENT					
Triage**			59 (6.2)		
Green/Blue	37 (14.8)	146 (22.6)		1 (reference)	
Yellow	126 (50.4)	353 (54.7)		1.40 (0.93-2.13)	0.105
Red/Orange	87 (34.8)	146 (22.6)		2.35 (1.50-3.67)	<0.001
VITAL PARAMETERS					
Respiratory rate >20/min	124 (47.0)	161 (23.5)	5 (0.5)	2.88 (2.13-3.88)	<0.001
Oxygen saturation < 96 %	162 (61.1)	231 (33.7)	4 (0.4)	3.09 (2.30-4.14)	<0.001
Heart rate <51 or >90/min	148 (55.8)	312 (45.3)	1 (0.1)	1.52 (1.14-2.02)	0.003
Fever > 38^{o}C	77 (29.3)	156 (22.7)	5 (0.5)	1.40 (1.02-1.93)	0.036
BLOOD TESTS					
Leukocytes <3.5 or > 8.8 10E9/L	214 (80.8)	456 (66.2)	0 (0.0)	2.14 (1.52-3.02)	<0.001
Neutrophilocytes > 7.5 10E9/L	187 (71.1)	362 (53.2)	10 (1.0)	2.16 (1.59-2.94)	<0.001
Lymphocytes† <1.00 or > 4.00 10E9/L	53 (55.2)	92 (40.9)	633 (66.3)	1.78 (1.10-2.88)	0.018
C-Reactive protein mg/L			0 (0.0)		
<20 mg/L	21 (7.9)	175 (25.4)		1 (reference)	

21-99 mg/L	86 (32.5)	205 (29.8)		3.49 (2.08-5.86)	<0.001
≥ 100 mg/L	158 (59.6)	309 (44.8)		4.26 (2.60-6.96)	<0.001

The predictors in the table are those dichotomised or categorised as they were later incorporated into the final diagnostic model. Only statistically significant results of the unadjusted association between each candidate predictor and the outcome CAP are presented. *Abnormal chest auscultation: Any abnormal findings such as crackles and rhonchi. ** Triage: Danish emergency process triage [39]. † Variables not included in the multivariate model.

Characteristics of patients suspected of CAP

Using the 70 candidate predictors, we compared clinical characteristics of patients with verified CAP to patients with suspected (402) but not verified CAP.

Statistically significant differences are shown in Table 2. Of the 402 patients with suspected CAP, half of the patients, 229 (57%) had verified CAP. Patients with suspected CAP had a median age of 74.0 (IQR: 62.0; 82.0), and half were male (52.7%). Patients with verified CAP reported more respiratory symptoms, such as cough (p=0.009) and expectoration (p=0.037), and more gastrointestinal symptoms, such as nausea (p=0.033) and loss of appetite (p=0.030), compared to those without CAP. Fewer patients with verified CAP had a CURB-65 ≥3 (p=0.047), and more patients had oxygen saturation <96% (p<0.001), a heart rate of <51 or >100bpm/min (p=0.045), and fever >38 ^{o}C (p=0.011). Elevated infection biomarkers (leukocytes, neutrophilocytes, CRP, all p<0.001), and plasma natrium (p<0.001) were highly associated with CAP. Fewer patients with CAP had plasma bilirubin values of <5 or >25 mmol/L (p=0.045) (Table 2).

Table 2: Characteristics of the population with suspected CAP (n=402) by the physician at admission.

Characteristics	Patients suspected of CAP at admission		Missings n (%)	OR (95% CI)	p-value
	CAP n (%)	Not CAP n (%)			
Total of patients	229 (57.0)	173 (43.0)	0 (0.0)		
SYMPTOMS					
Cough	168 (75.7)	104 (63.4)	16(4.0)	1.79 (1.15-2.79)	0.009
Expectoration	132 (59.5)	80 (48.8)	16 (4.0)	1.54 (1.02-2.31)	0.037
Nausea	70 (31.8)	36 (22.0)	18 (4.5)	1.65 (1.04-2.64)	0.033
Loss of appetite	137 (62.3)	84 (51.2)	18 (4.5)	1.57 (1.04-2.36)	0.030
SEVERITY ASSESSMENT					
CURB65 ≥3 *	23 (10.4)	30 (17.3)	8 (2.0)	0.55 (0.30-0.99)	0.047
VITAL PARAMETERS					
Oxygen saturation <96%	147 (64.2)	79 (46.0)	1 (0.2)	2.11 (1.40-3.15)	<0.001

Heart rate < 51 or >100 bpm/min	129 (56.3)	80 (46.2)	0 (0.0)	1.49 (1.00-2.23)	0.045
Fever >38°C	64 (28.2)	30 (17.3)	2 (0.5)	1.87 (1.14-3.05)	0.011
BLOOD TESTS					
Leukocytes <3.5 or > 8.8 10E9/L	191 (83.4)	106 (61.3)	0 (0.0)	3.17 (1.99-5.04)	<0.001
Neutrophilocytes > 7.5 10E9/L	166 (73.1)	81 (47.6)	5 (1.2)	2.99 (1.96-4.55)	<0.001
Natrium <137 or > 145 mmol/L	114 (49.8)	55 (31.8)	0 (0.0)	2.12 (1.40-3.21)	<0.001
Bilirubin<5 or >25 mmol/L	32 (14.0)	37 (21.8)	4 (1.0)	0.58 (0.34-0.98)	0.045
C-Reactive Protein mg/L, n (%)			0 (0.0)		
<20 mg/L	15 (6.6)	59 (34.1)		1 (reference)	
21-99 mg/L	74 (32.3)	64 (37.0)		4.54 (2.35-8.78)	<0.001
≥ 100 mg/L	140 (61.1)	50 (28.9)		11.01 (5.73-21.14)	<0.001

Statistically significant results from the unadjusted association between each candidate predictor and the outcome CAP.* CURB65: confusion, uremia, respiratory rate, blood pressure, age > 65 years.

Model development and performance

We developed a prediction model for diagnosing pneumonia in patients admitted with suspected infection (n=954) and compared it with the clinician's presumptive diagnosis. Supplementary table S3 presents the characteristics of the population randomised in the training and validation sets.

The predictors associated with CAP in our final model are presented in Table 3.

Table 3: The complete diagnostic model, including the intercept

Intercept and predictors	ß Coefficient
Intercept	-1.66192
Dyspnea (yes)	0.35172
Expectoration (yes)	0.36250
Cough (yes)	0.39671
Common cold (yes)	0.34374
Malaise (yes)	0.07475
Chest pain (yes)	0.20499
Respiratory rate >20/min	0.14566
Oxygen saturation < 96%	0.24303
Abnormal auscultation findings (yes)	0.56758
Leucocytes*	0.00322
Neutrophilocytes**	0.08338
C-reactive protein <20 mg/L	-0.64269
Previous event of CAP (no)	-0.12006
Age of ≥ 75 and cough (yes)	0.53816
Age of ≥ 75 and oedema (no)	-0.05797
Age of ≥ 75 and glucose >11.0 mmol/L	0.88124
ROC AUC† (95% CI)	0.85 (0.77-0.92)

* Cut-off for leucocyttes: normal values 3.5 -8.8 10E9/L **Neutrophilocytes: > 7.5 10E9/L
† ROC AUC = receiver-operating characteristic area under the curve

The model performance yielded an AUC of 0.85 (CI: 0.77-0.92) and the calibration of the model yielded p=0.227 after recalibration, demonstrating a good prediction of the proportion of CAP patients in the test sample (Supplementary figures S4 and S5).

Based on a lambda result of λ=0.0402856 and a probability threshold of 0.35, the LASSO calculation with characteristics predictive of CAP and the calculation of the final model with a cut-off value greater than 0 indicating the diagnosis CAP are presented in Supplemental material (Supplementary formulas S6 and S7).

At the optimal cut-off of 0.35, the prediction model yielded an 86.1% sensitivity and 64.1% specificity. Based on the trial population (Figure 1), the sensitivity of the prediction model was comparable to the initial diagnosis made by the ED physicians. However, the specificity and positive predictive value were significantly lower (Table 4).

Table 4: Performance of the predictive model compared to the initial diagnosis made by the ED physicians.

Performance	Sensitivity % (CI %)	Specificity % (CI %)	Positive predictive value % (CI %)	Negative predictive value % (CI %)
Predictive model	86.1 (79.1-93.1)	64.1 (57.1-71.1)	41.6 (34.6-48.6)	93.9 (86.9-100)
Physicians	86.4 (84.2-88.6)	74.9 (72.1-77.6)	57.0 (53.8-60.1)	93.5 (92.0-95.0)

The predictive model had a 35% cut-off and a prevalence of 22%. The prevalence of CAP was 28% in the population of 954 patients suspected of infection.

Model specification

The final model did not include the following possible predictors: lymphocytes, SARS-CoV-2, and BMI. The reasons were a high percentage of missings (lymphocytes 66.3%), clinical relevance, and statistical performance (BMI and SARS-CoV-2). These considerations are described in detail in Supplemental material.

DISCUSSION

More than every fourth patient with suspected infection was diagnosed with CAP (28%). The ED physicians suspected CAP in almost half (42%) of patients admitted with suspected infection. Patients with suspected CAP included 57% with a final expert diagnosis of CAP and 43% without CAP. We have identified twenty-seven clinical characteristics for patients diagnosed with CAP among those admitted suspected of infection. Patients with CAP were characterised by having more often a history of smoking, previous CAP, respiratory symptoms, abnormal lung auscultation, worse triage, and abnormal levels of infection biomarkers. Fewer clinic characteristics (thirteen) were identified for patients diagnosed with CAP among patients suspected of CAP by the ED physician and included typical respiratory symptoms but also gastrointestinal symptoms, abnormal vital signs, increased blood markers, and lower CURB-65 scores. The final diagnostic prediction model yielded thirteen diagnostic predictors for CAP recognised by the literature. The model performance was similar to the diagnosis made by the ED physicians regarding sensitivity and negative predictive value but not as good in determining the specificity and positive predictive values.

Our prediction model had a good performance (AUC 85%) and calibration (p=0.227), and with the best cut-off of 35%, the sensitivity reached 86.1% and specificity 64.1%. Therefore, the model could be tested externally and contribute to the initial management of CAP, guiding further clinical investigation. In this study, ED physicians who generally only had the patient's history and the results from a simple clinical examination diagnosed CAP with a comparable negative predictive value (93% vs. 94%) and a better positive predictive value (57% vs. 42%). Even though our model is not entirely comparable to the initial diagnosis made by the ED physicians due to the difference in the prevalence of CAP, our results are similar to a recent systematic review [43]. Other studies reported that ED physicians' accuracy in diagnosing CAP ranged from 76% to 96% [44], and artificial intelligence predicted the presence of pneumonia with a sensitivity of 94% and specificity of 50% [45]. These results show that there is room for improvement in diagnosing CAP. It could be achieved by including additional predictors such as biomarkers, e.g.,

procalcitonin, YKL-40, and surfactant protein-D [46, 47], molecular detection of respiratory pathogens [48], and/or improved imaging modalities [12, 14].

This prospective study highlights the challenges in identifying patients with CAP based on patient history, vital signs, and symptoms upon admission [20, 22, 46]. The initial CAP diagnosis often differs from the discharge diagnosis [10, 49]. A plausible cause for uncertainty in diagnosing CAP was the heterogenic presentation of symptoms overlapping with other diseases. We found that patients with verified CAP often had gastrointestinal symptoms, whereas patients not verified with CAP sometimes presented with typical respiratory symptoms and had more severe conditions measured by CURB-65. Typical respiratory symptoms could explain some CAP misclassification. Misclassification of CAP may lead to unnecessary or ineffective antibiotic treatment, increased healthcare costs, delayed diagnosis, increased mortality, and increased risk of bacterial resistance [44, 50].

The predictors of CAP identified in this study are strongly represented in the literature [9, 20, 36, 37, 42, 46, 49]. Most prediction models for ED patients with CAP aim to predict prognostic outcomes such as disease severity and mortality [51]. Prior studies have investigated only a few diagnostic predictors or studied very selected patients [20, 22, 52]. The main reason for including several potential predictors and having age as a cross-factor in the development of our model was the expectation of finding predictors not represented in the literature and predictors specific for older patients (≥75 years). This is considered very relevant as the population worldwide ages [4, 16]. An age of ≥ 75 interacted with the symptoms of cough, blood glucose levels, and peripheral oedema. Peripheral oedema was associated with an absence of CAP where symptoms may be explained by other infections such as erysipelas or cardiac heart failure patients with respiratory symptoms. In addition, hyperglycemia has been recognized as a predictor associated with poorer patient outcomes for elderly CAP patients, regardless of their history of diabetes [53, 54].

Even though the literature highlights malnutrition as a strong prognostic predictor for CAP [33, 35, 55], we excluded BMI from our final model. Measuring weight and height is not a priority in acute settings where vital parameters, symptoms, and point-of-care biomarkers are the primary observations in the diagnostic

process. Another concern was that BMI was missing in 26.3% of the population, and bias may arise due to systematic differences between subjects with complete datasets and subjects with missing data. Patients with missing BMI data may be more frail, incapable, or difficult to transfer. A model including BMI could be a better choice in a primary care setting, where patients are not necessarily as acutely ill and may be able to weigh themselves.

A major strength of this study is the completeness of data from medical charts and patient interviews combined with CAP diagnoses assigned by a panel of experts. The experts had a range of information from the patient's medical records, including chest x-ray, chest CT for patients suspected of CAP, and microbiology results available for many of the patients. In addition, to identifying possible predictors, we included many relevant and easily accessible clinical parameters. Finally, we excluded patients infected with SARS-CoV-2 from the study to increase the potential generalisability for CAP patients after the pandemic.

This study also has several limitations. Multiple testing and mass significance are potentially a problem in this study. Methods, such as Bonferroni-Holm correction, could have been applied to counteract this problem [56]. However, the univariate analyses were conducted for exploratory and descriptive purposes only. Therefore, these results should be interpreted cautiously, and the findings should be used as hypothesis-generating rather than conclusive. Another concern is that even though the reference standard of CAP was the same for the model performance and the initial diagnosis of the ED physicians, the expert panel might have a better prerequisite to diagnose CAP in suspected CAP patients due to the availability of results from imaging and microbiological tests, and better register of patient's symptoms. It might lead to differential verification bias overestimating the ED physician's accuracy in diagnosing CAP [57]. This assumption may be supported by the higher specificity of CAP diagnoses from ED physicians.

Another limitation is the selected population of the patients allocated to the internal medicine specialty that may have masked atypical predictors from patients assigned to other specialities. Furthermore, some patients with atypical clinical presentation might have an infection that the ED physician had not suspected

upon admission and therefore was not included in our study. Patients with severe condition or acute cognitive impairment who could not consent were excluded. A broader patient inclusion may contribute to a model that identifies other predictors assisting in diagnosing CAP as the clinical presentation might differ from those admitted with suspected CAP and capable of consent. Another limitation of the development of the model, was the choice of cut-offs for blood tests routinely used in our institutions, this pragmatic choice reflects our clinical practice. However, it does raise questions about the applicability in other settings that apply different cut-offs.

This population cohort could be applicable as a test validation cohort for future models as the data collection of these well-known predictors of CAP is reproducible across EDs. The development of automatic extraction for a prediction model from electronic medical records using artificial intelligence could be of great value in a busy ED. In conclusion, typical respiratory symptoms combined with abnormal vital signs and elevated infection biomarkers are predictors for CAP upon admission to an ED. A diagnostic prediction model based on these predictors is of limited value. Future prediction models should include novel diagnostic tools, imaging, PCR analysis, and/or serological markers not routinely used in clinical practice to improve model performance, helping diagnose CAP more accurately at the ED.

Acknowledgements: The authors appreciate text editing from the research consultant Caroline Moos, statistician support from Andreas Kristian Pedersen and Sofie Ronja Petersen at the University Hospital of Southern Denmark, and from OPEN (Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark).

Authors' contributions: MBC, FSR, CBM, TS, HSA, MHL, AH, and MAH were involved in the study's design. MBC performed the literature search and drafted the original work. MBC, MHL, AH, MAH, JJS, and FK recruited patients and collected data. CBM and MAH participated in the expert panel. HSA was the study investigator-, and coordinated and supervised the project. MBC performed the statistical analyses. CBM

was the chief research officer responsible for supervising the overall study. All authors, MBC, FSR, CBM, TS, HSA, MHL, AH, MAH, FK, and JJS critically revised and approved the final manuscript.

Funding: University of Southern Denmark (17/10636), University Hospital of Southern Denmark (20/20505), The funders of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Competing interests: The authors declare that they have no competing interests.

Patient and public involvement: Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication: Not required.

Ethics approval and consent to participate: Ethics approval and consent to participate: Approval was obtained from the Regional Committee for Health Research Ethics in Southern Denmark (S-20200188). In addition, informed verbal and written consent was obtained from each participant before enrolment in the study. This study was conducted in agreement with the Declaration of Helsinki-Ethical principle for medical research involving human subjects.

Availability of data and materials: Due to Danish laws on personal data, data cannot be shared publicly. To request data, please contact the corresponding author for more information. The person responsible for the research was the principal investigator and corresponding author (MBC) in collaboration with the University Hospital of Southern Denmark. This organization owns the data and can provide access to the final data set.

REFERENCES

1. Søgaard M, Nielsen RB, Schønheyder HC, Nørgaard M, Thomsen RW. Nationwide trends in pneumonia hospitalization rates and mortality, Denmark 1997-2011. Respir Med. 2014;108(8):1214-22.doi:10.1016/j.rmed.2014.05.004
2. McLaughlin JM, Khan FL, Thoburn EA, Isturiz RE, Swerdlow DL. Rates of hospitalization for community-acquired pneumonia among US adults: A systematic review. Vaccine. 2020;38(4):741-51.doi:10.1016/j.vaccine.2019.10.101
3. Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. Thorax. 2012;67(1):71-9
4. Laporte L, Hermetet C, Jouan Y, Gaborit C, Rouve E, Shea KM, et al. Ten-year trends in intensive care admissions for respiratory infections in the elderly. Ann Intensive Care. 2018;8(1):84.doi:10.1186/s13613-018-0430-6
5. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis. 2018;18(11):1191-210.doi:10.1016/s1473-3099(18)30310-4
6. Meehan TP, Fine MJ, Krumholz HM, Scinto JD, Galusha DH, Mockalis JT, et al. Quality of care, process, and outcomes in elderly patients with pneumonia. JAMA. 1997;278(23):2080-4.doi:10.1001/jama.1997.03550230056037
7. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infect Dis. 2014;14 (1):13.doi:10.1186/1471-2334-14-13
8. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med. 2019;200(7):e45-e67.doi:10.1164/rccm.201908-1581ST
9. Musher DM, Thorner AR. Community-acquired pneumonia. N Engl J Med. 2014;371(17):1619-28.doi:10.1056/NEJMra1312885
10. Chandra A, Nicks B, Maniago E, Nouh A, Limkakeng A. A multicenter analysis of the ED diagnosis of pneumonia. Am J Emerg Med. 2010;28(8):862-5.doi:10.1016/j.ajem.2009.04.014
11. Claessens YE, Debray MP, Tubach F, Brun AL, Rammaert B, Hausfater P, et al. Early chest computed tomography scan to assist diagnosis and guide treatment decision for suspected community-acquired pneumonia. Am J Respir Crit Care Med. 2015;192(8):974-82.doi:10.1164/rccm.201501-0017OC
12. Ye X, Xiao H, Chen B, Zhang S. Accuracy of Lung Ultrasonography versus Chest Radiography for the Diagnosis of Adult Community-Acquired Pneumonia: Review of the Literature and Meta-Analysis. PLoS One. 2015;10(6):e0130066.doi:10.1371/journal.pone.0130066
13. Kamat IS, Ramachandran V, Eswaran H, Guffey D, Musher DM. Procalcitonin to distinguish viral from bacterial pneumonia: a systematic review and meta-analysis. Clin Infect Dis. 2020;70(3):538-42.doi:10.1093/cid/ciz545
14. Gentilotti E, De Nardo P, Cremonini E, Górska A, Mazzaferri F, Canziani LM, et al. Diagnostic accuracy of point-of-care tests in acute community-acquired lower respiratory tract infections. A systematic review and meta-analysis. Clin Microbiol Infect. 2022;28(1):13-22.doi:10.1016/j.cmi.2021.09.025
15. Welker JA, Huston M, McCue JD. Antibiotic timing and errors in diagnosing pneumonia. Arch Intern Med. 2008;168(4):351-6.doi:10.1001/archinternmed.2007.84
16. World Health Organization. Aging and Health [Internet]. 2022 October 1 [cited 2022 October 28]. Available from: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

17. Weir DL, Majumdar SR, McAlister FA, Marrie TJ, Eurich DT. The impact of multimorbidity on short-term events in patients with community-acquired pneumonia: prospective cohort study. Clin Microbiol Infect. 2015;21(3):264.e7-.e13.doi:10.1016/j.cmi.2014.11.002
18. Sakakibara T, Shindo Y, Kobayashi D, Sano M, Okumura J, Murakami Y, et al. A prediction rule for severe adverse events in all inpatients with community-acquired pneumonia: a multicenter observational study. BMC Pulm Med. 2022;22(1):34.doi:https://dx.doi.org/10.1186/s12890-022-01819-0
19. Gong L, He D, Huang D, Wu Z, Shi Y, Liang Z. Clinical profile analysis and nomogram for predicting in-hospital mortality among elderly severe community-acquired pneumonia patients with comorbid cardiovascular disease: a retrospective cohort study. BMC Pulm Med. 2022;22(1):312.doi:10.1186/s12890-022-02113-9
20. Ding F, Han L, Yin D, Zhou Y, Ji Y, Zhang P, et al. Development and validation of a simple tool composed of items on dyspnea, respiration rates, and C-reactive protein for pneumonia prediction among acute febrile respiratory illness patients in primary care settings. BMC Med. 2022;20(1):360.doi:10.1186/s12916-022-02552-5
21. Hammond A, Halliday A, Thornton HV, Hay AD. Predisposing factors to acquisition of acute respiratory tract infections in the community: a systematic review and meta-analysis. BMC Infect Dis. 2021;21(1):1254.doi:10.1186/s12879-021-06954-3
22. Ebell MH, Chupp H, Cai X, Bentivegna M, Kearney M. Accuracy of signs and symptoms for the diagnosis of community-acquired pneumonia: a meta-analysis. Acad Emerg Med. 2020;27(7):541-53.doi:10.1111/acem.13965
23. Kitazawa T, Yoshihara H, Seo K, Yoshino Y, Ota Y. Characteristics of pneumonia with negative chest radiography in cases confirmed by computed tomography. J Community Hosp Intern Med Perspect. 2020;10(1):19-24.doi:10.1080/20009666.2020.1711639
24. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-W73.doi:10.7326/m14-0698
25. Skjøt-Arkil H, Heltborg A, Lorentzen MH, Cartuliares MB, Hertz MA, Graumann O, et al. Improved diagnostics of infectious diseases in emergency departments: a protocol of a multifaceted multicentre diagnostic study. BMJ Open. 2021;11(9):e049606.doi:10.1136/bmjopen-2021-049606
26. Nørgaard B, Mogensen CB, Teglbjærg LS, Brabrand M, Lassen AT. Diagnostic packages can be assigned accurately in emergency departments. A multi-centre cohort study. Dan Med J. 2016;63(6)
27. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208.doi:10.1016/j.jbi.2019.103208
28. Mutepe ND, Cockeran R, Steel HC, Theron AJ, Mitchell TJ, Feldman C, et al. Effects of cigarette smoke condensate on pneumococcal biofilm formation and pneumolysin. Eur Respir J. 2013;41(2):392-5.doi:10.1183/09031936.00213211
29. Samokhvalov AV, Irving HM, Rehm J. Alcohol consumption as a risk factor for pneumonia: a systematic review and meta-analysis. Epidemiol Infect. 2010;138(12):1789-95.doi:10.1017/s0950268810000774
30. Torres A, Peetermans WE, Viegi G, Blasi F. Risk factors for community-acquired pneumonia in adults in Europe: a literature review. Thorax. 2013;68(11):1057-65.doi:10.1136/thoraxjnl-2013-204282
31. Barbagelata E, Cillóniz C, Dominedò C, Torres A, Nicolini A, Solidoro P. Gender differences in community-acquired pneumonia. Minerva Med. 2020;111(2):153-65.doi:10.23736/s0026-4806.20.06448-4
32. Baik I, Curhan GC, Rimm EB, Bendich A, Willett WC, Fawzi WW. A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women. Arch Intern Med. 2000;160(20):3082-8.doi:10.1001/archinte.160.20.3082
33. Cillóniz C, Dominedò C, Pericàs JM, Rodriguez-Hurtado D, Torres A. Community-acquired pneumonia in critically ill very old patients: a growing problem. Eur Respir Rev. 2020;29(155):190126.doi:10.1183/16000617.0126-2019

34. Reisinger EC, Fritzsche C, Krause R, Krejs GJ. Diarrhea caused by primarily non-gastrointestinal infections. Nat Clin Pract Gastroenterol Hepatol. 2005;2(5):216-22
35. Riquelme OR, Riquelme OM, Rioseco ZML, Gómez MV, Cárdenas G, Torres C. Neumonía adquirida en la comunidad en el anciano hospitalizado: Aspectos clínicos y nutricionales. [Community-acquired pneumonia in the elderly: clinical and nutritional aspects]. Rev Med Chil. 2008;136(5):587-93.doi:10.4067/S0034-98872008000500006
36. Moore M, Stuart B, Little P, Smith S, Thompson MJ, Knox K, et al. Predictors of pneumonia in lower respiratory tract infections: 3C prospective cough complication cohort study. Eur Respir J. 2017;50(5).doi:10.1183/13993003.00434-2017
37. van Vugt SF, Broekhuizen BD, Lammens C, Zuithoff NP, de Jong PA, Coenen S, et al. Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. BMJ. 2013;346:f2450.doi:10.1136/bmj.f2450
38. Alam N, Vegting IL, Houben E, van Berkel B, Vaughan L, Kramer MH, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation. 2015;90:111-5.doi:10.1016/j.resuscitation.2015.02.011
39. Plesner LL, Iversen AKS, Langkjær S, Nielsen TL, Østervig R, Warming PE, et al. The formation and design of the TRIAGE study-baseline data on 6005 consecutive patients admitted to hospital from the emergency department. Scand J Trauma Resusc Emerg Med. 2015;23:106.doi:10.1186/s13049-015-0184-1
40. Ravioli S, Germann C, Gygli R, Exadaktylos AK, Lindner G. Age- and sex-related differences in community-acquired pneumonia at presentation to the emergency department: a retrospective cohort study. Eur J Emerg Med. 2022;29(5):366-72.doi:10.1097/mej.0000000000000933
41. Akhtar A, Hassali MAA, Zainal H, Ali I, Iqbal MS, Khan AH. Respiratory-tract infections among geriatrics: prevalence and factors associated with the treatment outcomes. Ther Adv Respir Dis. 2021;15:1753466620971141.doi:10.1177/1753466620971141
42. Metlay JP, Schulz R, Li YH, Singer DE, Marrie TJ, Coley CM, et al. Influence of age on symptoms at presentation in patients with community-acquired pneumonia. Arch Intern Med. 1997;157(13):1453-9.doi:doi:10.1001/archinte.1997.00440340089009
43. Dale AP, Marchello C, Ebell MH. Clinical gestalt to diagnose pneumonia, sinusitis, and pharyngitis: a meta-analysis. Br J Gen Pract. 2019;69(684):e444-e53.doi:10.3399/bjgp19X704297
44. Ray P, Birolleau S, Lefort Y, Becquemin MH, Beigelman C, Isnard R, et al. Acute respiratory failure in the elderly: etiology, emergency diagnosis and prognosis. Crit Care. 2006;10(3):R82.doi:10.1186/cc4926
45. Heckerling PS, Gerber BS, Tape TG, Wigton RS. Prediction of community-acquired pneumonia using artificial neural networks. Med Decis Making. 2003;23(2):112-21.doi:10.1177/0272989x03251247
46. Htun TP, Sun Y, Chua HL, Pang J. Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review. Sci Rep. 2019;9(1):7600.doi:10.1038/s41598-019-44145-y
47. Spoorenberg SM, Vestjens SM, Rijkers GT, Meek B, van Moorsel CH, Grutters JC, et al. YKL-40, CCL18 and SP-D predict mortality in patients hospitalized with community-acquired pneumonia. Respirology. 2017;22(3):542-50.doi:10.1111/resp.12924
48. Gastli N, Loubinoux J, Daragon M, Lavigne JP, Saint-Sardos P, Pailhoriès H, et al. Multicentric evaluation of BioFire FilmArray Pneumonia Panel for rapid bacteriological documentation of pneumonia. Clin Microbiol Infect. 2021;27(9):1308-14.doi:10.1016/j.cmi.2020.11.014
49. Metlay JP, Fine MJ. Testing strategies in the initial management of patients with community-acquired pneumonia. Ann Intern Med. 2003;138(2):109-18.doi:10.7326/0003-4819-138-2-200301210-00012
50. Kanwar M, Brar N, Khatib R, Fakih MG. Misdiagnosis of community-acquired pneumonia and inappropriate utilization of antibiotics: side effects of the 4-h antibiotic administration rule. Chest. 2007;131(6):1865-9.doi:10.1378/chest.07-0164
51. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. Thorax. 2010;65(10):884-90.doi:10.1136/thx.2009.134072

52. Okimoto N, Yamato K, Kurihara T, Honda Y, Osaki K, Asaoka N, et al. Clinical predictors for the detection of community-acquired pneumonia in adults as a guide to ordering chest radiographs. Respirology. 2006;11(3):322-4.doi:10.1111/j.1440-1843.2006.00846.x

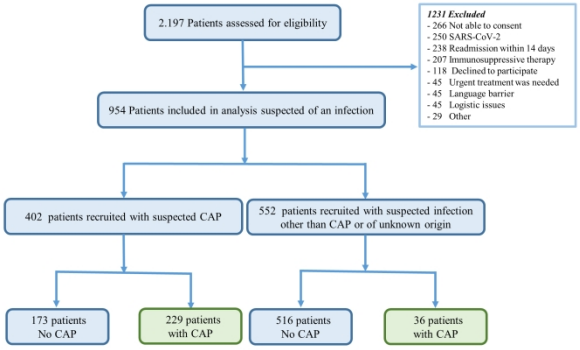
53. Zeng W, Huang X, Luo W, Chen M. Association of admission blood glucose level and clinical outcomes in elderly community-acquired pneumonia patients with or without diabetes. Clin Respir J. 2022;16(8):562-71.doi:10.1111/crj.13526

54. Barmanray RD, Cheuk N, Fourlanos S, Greenberg PB, Colman PG, Worth LJ. In-hospital hyperglycemia but not diabetes mellitus alone is associated with increased in-hospital mortality in community-acquired pneumonia (CAP): a systematic review and meta-analysis of observational studies prior to COVID-19. BMJ Open Diabetes Res Care. 2022;10(4).doi:10.1136/bmjdrc-2022-002880

55. Yeo HJ, Byun KS, Han J, Kim JH, Lee SE, Yoon SH, et al. Prognostic significance of malnutrition for long-term mortality in community-acquired pneumonia: a propensity score matched analysis. Korean J Intern Med. 2019;34(4):841-9.doi:10.3904/kjim.2018.037

56. Sedgwick P. Multiple significance tests: the Bonferroni correction. BMJ. 2012;344

57. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174(4):469-76.doi:10.1503/cmaj.050090



338x190mm (300 x 300 DPI)

Supplemental material

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Mariana Bichuette Cartuliares MSc* [1, 2], Christian Backer Mogensen [1, 2], Flemming Schønning Rosenvinge [3, 4], Thor Aage Skovsted [5], Morten Hjarnø Lorentzen [1,2], Anne Heltborg Kristensen[1,2], Mathias Amdi Hertz [6,7], Frida Kaldan[1], Jens Juel Specht[1], Helene Skjøt-Arkil [1,2]

Affiliations

[1]Emergency Department, University Hospital of Southern Denmark, Aabenraa, Denmark

[2]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark

[3]Research Unit of Clinical Microbiology, University of Southern Denmark, Odense, Denmark

[4]Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark

[5]Department of Biochemistry and Immunology, University Hospital of Southern Denmark, Aabenraa, Denmark

[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark 7 Infectious Diseases Department, Odense University Hospital, Odense, Denmark

* Corresponding author: mbc@rsyd.dk

Table of Contents

Table S1: Description of the 70 pre-specified predictors for CAP

Source: *The patient interview*			
Group	**Variable name**	**Measurement**	**Consideration/assumption** Considerations to collect data from these predictors were based on the described literature and expert consensus together with the project group
Demographic information	Age	Continuous, years	Age is a risk factor for CAP [1]. Several studies stratify age groups when investigating pneumonia due to several atypical symptoms and signs and the absence of respiratory symptoms among the elderly. Stratified age groups differ in cut-offs between the ages of ≥65 to ≥80 years old [2-7].
	Gender	Binary 1=Male 0=Female	The risk of CAP is higher for males [8].CAP is more severe [7] leading to higher mortality in males [9]. Males' lifestyle factors differ from women resulting in a higher risk of CAP [10].
	Civil status (Living alone)	Binary (Yes/no)	Living alone has a two-fold association with having one or more respiratory tract infections [11].
	Nursing home residence	Binary (Yes/no)	Nursing home residents were found to have several comorbidities [12] and lower physical functioning levels, which might result in a higher risk of CAP [13].
	Employment	Categorical: 1=Working 2=Retired 0=Others (e.g. students, flex job)	Low income and unemployment are associated with readmissions after CAP [14].
Symptoms	Feeling unwell/ Malaise	Binary (Yes/No) Symptoms within 14 days prior to ED admission.	Malaise has been identified as one of the most frequent symptoms for patients infected with *Mycoplasma pneumoniae*[15].
	Fatigue		Fatigue is associated with pneumonia especially in elderly patients [4].
	Headache		Headache is one of the clinical findings of symptoms of CAP [7, 15]. However, headaches were less common in the older population [7].
	Dizziness		The rationale of the presence of dizziness as a symptom relied on the assumption that several factors such as polypharmacy[16], combined with comorbidities such as cardiovascular diseases [17], symptoms such as confusion, conditions of frailty and malnutrition [18], and lower oxygen saturation [19]could contribute to dizziness.
	Confusion		Confusion e.g. altered mental status or delirium was significantly more frequent in CAP patients [2, 4].
	Dyspnea		Dyspnea was identified as a strong prediction of CAP among febrile patients [20] and one of the main symptoms of pneumonia [2, 21].
	Cough		Cough is a common symptom and one of the most frequent increasing the likelihood of detecting a viral pathogen among CAP patients [15, 22]. Algorithms included cough as a diagnostic predictor [23], and dry cough was a strong predictor in a prediction model for *Legionella pneumoniae* [24]. Cough was less common in older population [7].
	Secretions		Purulent secretions were a significant symptom and predictor for CAP patients [20, 21].
	Sore throat		Some studies identified sore throat as a symptom of CAP [15], and one included the symptom in the prediction rules of pneumonia [5].

	Cold		Among respiratory diseases, the common cold is one of the most frequent, with symptoms similar to CAP [25].
	Fever feeling		Quantified from reported chills or night sweat or fever measured at home. Included as a rationale of fever.
	Chest pain		Chest pain has been used as a single predictor of CAP [18, 20, 23] or a combined diagnostic predictor [23] and may present as a secondary symptom of coughing or pleuritic involvement [26]. However, chest pain was less common in the older population [7].
	Peripheral edema		The rationale for including peripheral edema as possible predictor is that it is included in the clinical assessment at admission. In case of peripheral edema and respiratory symptoms of dyspnea, chest pain and a history of cardiovascular disease, CAP could be ruled out as a tentative diagnosis replaced with suspicion of cardiovascular disease.
	Nausea		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Vomiting		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Loss of appetite		Loss of appetite could be present in the case of gastrointestinal symptoms [26] and could result from malnutrition [18].
	Abdominal pain		Abdominal pain may be present in the case of gastrointestinal symptoms described above and, therefore, is included in the model [26].
	Diarrhea		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Pain in muscles and joints including back pain		Muscle and joint pain are associated with viral pneumonia as influenza, especially among younger patients and therefore is included in our model [27].
Previous event of CAP	Previous event of CAP	Categorical: 0= Never 1= Once 2= More than once	A previous diagnosis of CAP was reported as having robust evidence as a risk factor for CAP [1]. Furthermore, any hospitalization in the previous five years was reported as a predisposing factor for CAP [8].
Lifestyle factors and aids	Smoke	Categorical: 0=Never been a smoker 1=Current smoker 2=Previous smoker	Smoking has been associated with an increased risk of CAP in several studies [1, 8, 10, 17], and has a strong association with the treatment outcomes of elderly individuals with respiratory tract infections [28].
	Alcohol	Doses per week (a dose=12 grams (1, 5 cl) alcohol). Categories based on the Danish Board of Health recommendations [29]. 0=No alcohol 1=1-7 doses/week maximum doses recommended for women 2=8-14 doses/week maximum dose recommended for men 3= >14 doses	Alcohol has also been associated with increased CAP risk and with treatment outcomes. The risk increases in individuals with higher consumption (>41 g/day) compared to those who consume no alcohol [10, 17, 28].
	Physical activity levels	We categorized physical activity levels based on recommendations from the world health organization for adults with a minimum 150 min/week [30]. 1= Not physically active 2= Less than 2.5hrs/week 3= More than 2.5hrs/week	The risk of CAP decreased in physically active women [10]. In addition, a high level of activity protects against upper respiratory tract infections and reduces the severity and symptoms of the infection [13].
	Activities of daily living	Binary (yes/no) Yes= If the patient had one or more dependencies regarding: bathing, dressing, toileting, transfer, continence and feeding.	Difficulty in maintaining toilet hygiene, preparing meals, and being unable to transfer were associated with an increased risk of respiratory infections [31].

Source: *Variables extracted from the patient's medical report*			
Comorbidities (diseases)	Neurological	Binary (Yes/no) If the patient was diagnosed with one of these diagnoses.	Cerebrovascular disease/stroke and Parkinson's disease approximately doubled the risk of CAP [17].
	Pulmonary		A history of pneumonia increased the risk of a subsequent episode and patients with chronic respiratory diseases, including chronic obstructive pulmonary disease, bronchitis or asthma, had up to a fourfold increase in the risk of CAP [1, 4, 17].
	Endocrinological		Chronic liver conditions were reported as a risk factor of CAP [8]. Recently, diabetes mellitus has been described as an independent risk factor for sepsis secondary to CAP in very old patients [4] and data from several studies showed an association between diabetes mellitus and moderate risk of CAP [17].
	Renal		Chronic renal disease was reported as an independent risk factor for sepsis secondary to CAP in very old patients [4, 8] and chronic renal disease increased the risk of CAP twofold [17].
	Cardiovascular		Chronic cardiovascular disease increased the risk of CAP up to threefold [4, 17].
	Gastrointestinal		The rationale for including gastrointestinal diseases in the model was that CAP patients have gastrointestinal symptoms that could be related to a differential diagnosis besides CAP.
	Dementia		Dementia approximately doubles the risk of CAP [17].
	Cancer		Cancer was associated with a moderate increase in CAP risk, and a single study reported a fivefold increased risk of CAP for patients with lung cancer [17].
	Rheumatological		A moderate risk of CAP was found in patients with rheumatological diseases [17].
Pharmacological treatments	Polypharmacy	Binary (yes/no) Regular consumption of at least five medications	The increased number of comorbidities of older patients increases the risk of polypharmacy [4, 32]. The prevalence of polypharmacy reached almost 40% among individuals with respiratory tract infections above age 65 years and had a twofold association with treatment outcomes of respiratory tract infections [28]. Furthermore, the prevalence of polypharmacy increased from 45% to 74%, irrespective of antibiotic use if patients were hospitalized with CAP [16].
	Analgesics	Binary (Yes/no) Regular consumption of analgesics	A systematic review reported an association between prescribed opioids and CAP [33].
	Vaccination SARS-CoV-2	Binary (Yes/no) Recent vaccination for SARS-CoV-2	SARS-CoV-2 vaccination was reported during the clinical assessment but was taken out of the model, as the model would be used after the pandemic when vaccination for SARS-CoV-2 rates might decrease. However, the inclusion of this variable did not change the final predictive model.
	Vaccination pneumococcus	Binary (Yes/no) Pneumococcus vaccine (not specified) within 5 years	*Streptococcus pneumoniae* is one of the most causative pathogens of CAP and the vaccine could be a possible protective predictor for CAP as the risk of CAP increases among those unvaccinated [1, 34, 35].
	Vaccination influenza	Binary (Yes/no) Season influenza vaccine 2020/2021	Influenza vaccine can reduce hospitalization but is questionable if it could have a protective effect in admitted patients [1, 36], therefore, we included this possible predictor to investigate if it could have a protective role in our population.
Severity assessment	CURB-65	Binary ≥ 3 points (Yes/no) Definition: Confusion, urea >7 mmol/L, respiratory rate ≥ 30 bpm, blood pressure (≤90 for systolic blood pressure or ≤60 for diastolic blood pressure, age > 65 years) Score: one point for each present variable. CURB65≥ 3= severe condition	CURB65 is an assessment tool for the severity of CAP [37] recommended by the guidelines in Europe [38] including in Denmark [39].

	Triage	Based on the 5-level triage system "Danish emergency department triage" (DEPT) [40, 41], we categorized the following: Red/Orange and Green/Blue were pooled due to few patients in the blue and red groups: 1= Red/Orange 2= Yellow 3= Green/Blue	DEPT is a Danish adaption and modification of the "Adaptive Process Triage" (ADAPT) developed in Sweden [42]. DEPT was chosen as it is routinely used in the three included sites. Furthermore, in Denmark, most EDs have implemented formalized triage called "Danish Emergency Process Triage". DEPT shares core similarities with widespread standardized 5-level triage systems [43].
Vital parameters All vital parameters regardless of diastolic blood pressure were based on The National Early Warning Score (NEWS) [44]. This score was chosen as it is routinely used in the three EDs included in this study and cut-offs values in predicting CAP are similar from the literature.	Oxygen saturation	Binary < 96 % (Yes/no) The cut-off was based on The National Early Warning Score (NEWS) [44]. However, we did not differentiate between patients with chronic obstructive pulmonary disease.	A similar cut-off of oxygen saturation has been used in investigating predictors for CAP [19].
	Heart rate	Binary < 51 or >90 bpm (Yes/no)	Some studies have investigated and pointed out that a higher heart rate with similar cut-offs as a predictor for CAP [19, 45, 46].
	Blood pressure systolic	Binary <111 or >219 mmHg (Yes/no)	Other cut-offs based on the CURB65-score or lower level of triage (<90mmHg) have been used to predict a high risk of adverse events among inpatients with CAP [47]. This cut-off was also explored in our model without resulting in any difference.
	Blood pressure diastolic	Binary ≤60 mmHg (Yes/no) Based on severity assessment CURB65-score [37]. The NEWS does not include diastolic blood pressure and therefore the value from CURB-65 was chosen.	CURB-65 is routinely used in Denmark as a severity score and is included in the guidelines for antibiotic treatment [39]. As systolic blood pressure has been investigated in prediction rules, we added diastolic blood pressure to our model to explore this variable as a predictor for CAP.
	Respiratory rate (RR)	Binary >20 breaths/min (Yes/no)	There are different cut-offs of RR in the literature [20, 47]. RR> 20/min was defined as a strong prediction of CAP among febrile patients [20].
	Temperature	Binary >38 ^{o}C (Yes/no) Measured with ear thermometer [48].	Different cut-offs have been investigated, including the cut-off of >38^{o}C used in this study [49]. Independent of cut-offs, several studies have identified fever as a predictor of CAP [19-21, 23, 45]. However, fever is less common and generally absent in the older population [7].
	Glascow coma score	Binary >15 (Yes/no)	Cognitive impairment [32] has been reported as a strong risk factor for delirium and confusion as a predictor of the severity of CAP [47]. Altered mental status is associated with CAP, especially in the elderly [18].
Blood tests The literature does not describe a clear cut-off for the diagnosis of CAP. We chose a pragmatic approach and applied the cut-offs of serum biomarkers used in the EDs from our institution to reflect reality. Most of the serological biomarkers have been studied for prognostic	Hematocrit	Hematocrit (%), median (IQR) Binary (Yes/no) Cut-off: 40-50 for males and 35-46 for females Yes= outside of the cut-off No= within the cut-off	A hematocrit value of less than 35% was an independent predictor for severity and 2 years of mortality (p = 0.035) [50].
	Hemoglobin	Hemoglobin mmol/L, median (IQR) Binary (Yes/no) Cut-off: 8.3-10.5 for males and 7.3-9.5 for females Yes= outside of the cut-off No= within the cut-off	Hemoglobin correlates with frailty in the elderly and indirectly could be a predictor that should be investigated [51].
	Leukocytes	Leukocytes 10E9/L, median (IQR) Binary (Yes/no) Cut-off: 3.5-8.8 Yes= outside of the cut-off No= within the cut-off	Elevated leucocytes have been reported as a predictor for CAP, especially in pneumonia with negative chest x-ray [52].

purposes. We have included these as potential predictors for CAP to investigate their diagnostic prediction performance combined with signs and symptoms. Binary (Yes/no) measures. Yes= abnormal/ outside of the cut-off No= normal/ within the cut-off	Platelets	Platelets 10E9/L, median (IQR) Binary (Yes/no) Cut-off: 145-350 for males and 165-400 for females Yes= outside of the cut-off No= within the cut-off	Platelet count < 171 × 109/L was included in a prediction model for *legionella pneumoniae* showing a high diagnostic accuracy [AUC 0.89 (95% CI 0.86–0.93)] [24].
	Neutrophils	Neutrophilocytes 10E9/L, median (IQR) Binary (Yes/no) Cut-off: > 7.5 Yes= >7.5 No= ≤ 7.5	The neutrophil to lymphocyte ratio had a high diagnostic value for CAP patients [53]. Furthermore, higher mortality risk was found for CAP patients and if measured in the early stage of CAP could contribute to the diagnostic and disease severity [54].
	Lymphocytes	Lymphocytes 10E9/L, median (IQR) Binary (Yes/no) Cut-of: 1.00-4.00 Yes= outside of the cut-off No= within the cut-off	The neutrophil to lymphocyte ratio has been studied in prognostic studies and is associated with higher mortality risk in CAP patients and if measured in the early stage of CAP could contribute to the diagnostic and disease severity [54].
	Albumin	Albumin g/L, median (IQR) Binary (Yes/no) Cut-off: 34-45 Yes= outside of the cut-off No= within the cut-off	The ratio of blood urea and albumin has been investigated as a predictive factor for CAP, but poor model performance advocated for further investigation [55]. Furthermore, albumin correlates with frailty in the elderly and indirectly could be a predictor that should be investigated as frailty has been associated with an increased risk of CAP [51]. In addition, serum albumin (<3.4 g/dl) was associated with higher mortality for elderly patients with CAP [18] and was included in a prediction rule for severe adverse events in patients hospitalized with CAP (< 2 g/dL, 2 points; 2–3 g/dL, 1 point) [47].
	Creatinine	Creatinine µmol/L, median (IQR) Binary (Yes/no) Cut-off: 60-105 for males and 45-90 for females Yes= outside of the cut-off No= within the cut-off	Elevated creatinine levels have been reported with almost a sixfold association of poor CAP outcome (OR=5.67; 95%CI: 1.72-18.65) [56]. This result is supported by another study that showed that serum creatinine levels of ≥ 2.8 were a strong predictor of in-hospital mortality in adults with CAP when compared with five serum biomarkers [57].
	Blood urea	Blood urea nitrogen mmol/L, median (IQR) Binary (Yes/no) Cut-off: 3-5-8.1 for males and 3.1-7.9 for females Yes= outside of the cut-off No= within the cut-off	The ratio of blood urea and albumin has been investigated as a predictive factor for CAP, but poor model performance advocated for further investigation [55].
	Natrium	Natrium mmol/L, median (IQR) Binary (Yes/no) Cut-off: 137-145 Yes= outside of the cut-off No= within the cut-off	Hyponatremia < 133 mmol/L was one of the strong predictors in the prediction of CAP caused by *legionella pneumoniae* [24].
	Prothrombin time-international normalized ratio	Prothrombin (IQR) Binary (Yes/no) Cut-off: <1.2 Yes= ≥ 1.2 No= <1.2	Prothrombin time-international normalized ratio was investigated to distinguish Influenza A (H1N1) from other pneumonia. Prothrombin times were lower in H1N1 compared with non-H1N1 pneumonia patients (p=0.04) [58]. Furthermore, it has been investigated as a factor that could be associated with decreased sensitivity in negative urinary antigen (UAT) tests in CAP caused by pneumococcal. Prothrombin was 50% higher in the UAT-negative patients than in the UAT-positive patients [59]. We chose to include prothrombin in the diagnostic model to explore its significance in or rule out CAP, furthermore, the marker is routinely measured in acutely admitted patients.

	Bilirubin	Bilirubin µmol/L, median (IQR) Binary (Yes/no) Cut-off: <5 or >25 Yes= outside of the cut-off No= within the cut-off	Bilirubin levels were lower in patients with influenza A (H1N1) compared to non-H1N1 pneumonia (p= 0.02) [58]. This marker could add value to a prediction model.
	Glucose	Glucose mmol/L, median (IQR) Binary (Yes/no) Cut-off: > 11.00 Yes= >11.00 No= ≤ 11.00	Patients with CAP frequently present with admission hyperglycemia and have poorer outcomes [60, 61]. Therefore, glucose is included as a potential predictor.
	C- reactive protein (CRP)	C-Reactive Protein, median (IQR) Binary (Yes/no) The cut-off of CRP in our institution is < 5 mg/L at the ED. However, the literature suggests optional cut-offs. Based on the literature and the range of the results from the CRP as continuous variable, we defined the following categories: 1= <20mg/L 2= 20-100 mg/L 3= >100 mg/L	The diagnostic accuracy of CRP in differentiating between bacterial and viral infections of the lower respiratory tract is questionable [62]. However, CRP at different cut-offs increased the performance of prediction models for CAP. It included a cut-off of >20 [20], >30 [63], 50 [23] ≥ 98 [46], and a meta-analysis investigated all three cut-offs of 20, 50, and 100 [64]. CRP levels were found higher when CAP was detected both by a chest x-ray and a chest tomography [52].
Clinical assessment	Stethoscope findings	Binary (Yes/no) Yes for any abnormal stethoscope findings such as crackles and rhonchi.	Several studies investigated associations between abnormal stethoscope findings and the probability of the presence of CAP. They increased the likelihood of CAP [21, 65] and crackles on auscultation had a twofold increase in the prediction of pneumonia [19].
	Abdominal pain on palpation	Binary (Yes/no)	The rationale for including abdominal pain in the clinical assessment was that the literature reported that 20% of symptoms reported by patients with CAP were gastrointestinal symptoms [26].
	Body mass index (BMI).	The BMI was calculated including the high and weight of the patients. The BMI classification was based on "The Centers for diseases control and prevention" [66] and defined with the following categories: 1= Underweight, BMI < 18.5 2= Healthy weight, BMI from 18.5 to <25 3= Overweight, BMI from 25.0 to <30 4= Obesity, BMI from ≥ 30.0	The literature reported the association of several nutritional factors related to CAP and including malnutrition [1, 18], being underweight [8, 17], and BMI was directly associated with an increased risk of CAP among women [10].

Table S2: Characteristics of CAP in the population of patients admitted with an infection (n=954). The values presented of data as continuous, dichotomous or categorical were tested in the model during explorative analysis to identify the best model performance.

Characteristics	Total, n	CAP, n	Not CAP, n	Missings n (%)	OR (95% CI)	p-value
Total of patients	954 (100)	265 (27.8)	689 (72.2)	0 (0.0)		
DEMOGRAPHIC DATA						
Age, median (IQR)	73.0 (59.0; 81.0)	75.0 (63.5; 82.0)	73.0 (57.0; 80.0)	0 (0.0)	1.01 (1.005-1.02)	<0.001
Age ≥75 years	440 (46.1)	133 (50.2)	307 (44.6)	0 (0.0)	1.25 (0.94-1.66)	0.118
Gender male	513 (53.8)	137 (51.7)	376 (54.6)	0 (0.0)	0.89 (0.67-1.18)	0.425
Marital status, Living alone	618 (66.0)	166 (63.8)	452 (66.9)	18 (1.9)	0.87 (0.64-1.18)	0.382
Nursing home resident	66 (7.0)	26 (9.9)	40 (5.9)	13 (1.4)	1.75 (1.05-2.94)	0.317
Occupation				21 (2.2)		
Others	67 (7.2)	17 (6.5)	50 (7.4)		1 (reference)	
Working	202 (21.7)	44 (16.9)	158 (23.5)		0.81 (0.43-1.55)	0.543
Retired	664 (71.2)	200 (76.6)	464 (69.0)		1.26 (0.71-2.25)	0.418
LIFESTYLE FACTORS						
Smoking status				33 (3.5)		
No	323 (35.1)	66 (26.0)	257 (38.5)		1 (reference)	
Current smoker	179 (19.4)	54 (21.3)	125 (18.7)		1.68 (1.10-2.55)	0.015
Previous smoker	419 (45.5)	134 (52.8)	285 (42.7)		1.83 (1.30-2.57)	<0.001
Alcohol status				35 (3.7)		
No alcohol	356 (38.7)	99 (39.1)	257 (38.6)		1 (reference)	
1-7 doses	385 (41.9)	105 (41.5)	280 (42.0)		0.97 (0.70-1.34)	0.870
8-14 doses	105 (11.4)	31 (12.3)	74 (11.1)		1.08 (0.67-1.75)	0.732
> 14 doses	73 (7.9)	18 (7.1)	55 (8.3)		0.84 (0.47-1.51)	0.582
Physically activity				52 (5.4)		
Not physical active	263 (29.2)	74 (29.8)	189 (28.9)		1 (reference)	
Physical activity < 2,5 hr/week	231 (25.6)	64 (25.8)	167 (25.5)		0.97 (0.66-1.45)	0.915
Physical activity ≥ 2,5 hr/week	408 (45.2)	110 (44.4)	298 (45.6)		0.94 (0.66-1.33)	0.735
Body Mass Index, median (IQR)	26.5 (23.2; 30.8)	26.2 (22.9; 29.5)	26.7 (23.3; 31.2)	249 (26.1)	0.97 (0.94-0.99)	0.031
Body Mass Index†				249 (26.1)		
Healthy weight	246 (34.9)	74 (36.1)	172 (34.4)		1 (reference)	
Obese	193 (27.4)	45 (22.0)	148 (29.6)		0.70 (0.45-1.08)	0.114
Overweight	239 (33.9)	74 (36.1)	165 (33.0)		1.04 (0.70-1.53)	0.833
Underweight	27 (3.8)	12 (5.9)	15 (3.0)		1.85 (0.83-4.16)	0.132
ADL dependence*	260 (28.0)	81 (31.2)	179 (26.8)	25 (2.6)	1.23 (0.90-1.69)	0.180
SYMPTOMS						
Feeling unwell	559 (61.2)	173 (67.8)	386 (58.7)	41 (4.3)	1.48 (1.09-2.01)	0.010
Feeling tired	657 (72.6)	190 (75.4)	467 (71.5)	49 (5.1)	1.22 (0.87-1.70)	0.241
Headache	351 (38.3)	99 (38.8)	252 (38.1)	37 (3.9)	1.03 (0.76-1.38)	0.832
Dizziness	346 (37.7)	96 (37.6)	250 (37.8)	37 (3.98)	0.99 (0.73-1.34)	0.973
Confusion	207 (22.6)	58 (22.7)	149 (22.5)	37 (3.89)	1.01 (0.71-1.43)	0.938
Dyspnea	379 (41.4)	171 (67.3)	208 (31.5)	39 (4.1)	4.48 (3.29-6.11)	<0.001
Cough	358 (39.1)	173 (68.1)	185 (28.0)	39 (4.1)	5.49 (4.01-7.52)	<0.001

Expectoration	279 (30.5)	140 (55.1)	139 (21.0)	39 (4.1)	4.61 (3.38-6.28)	<0.001
Sore throat	104 (11.4)	39 (15.4)	65 (9.8)	39 (4.1)	1.66 (1.08-2.54)	0.019
Cold (common cold)	95 (10.4)	45 (17.7)	50 (7.6)	39 (4.1)	2.63 (1.70-4.05)	<0.001
Fever feling at home	612 (64.2)	169 (63.8)	443 (64.3	0 (0.0)	0.97 (0.72-1.31)	0.880
Chest pain	168 (18.4)	71 (28.1)	97 (14.7)	40 (4.2)	2.26 (1.60-3.21)	<0.001
Oedema	79 (8.6)	10 (4.0)	69 (10.4)	39 (4.1)	0.35 (1,17-0.69)	0.002
Nausea	304 (33.2)	76 (30.0)	228 (34.4)	38 (3.9)	0.81 (0.59-1.112)	0.211
Vomiting	190 (20.7)	40 (15.8)	150 (22.6)	38 (3.9)	0.64 (0.43-0.94)	0.023
Loss of appetite	524 (57.2)	149 (58.9)	375 (56.6)	38 (3.9)	1.00 (0.82-1.47)	0.523
Gastrointestinal pain	193 (21.1)	40 (15.8)	153 (23.1)	38 (3.9)	0.62 (0.42-0.91)	0.016
Diarrhoea	134 (14.6)	29 (11.5)	105 (15.8)	38 (3.9)	0.68 (0.44-1.06)	0.095
Muscular pain	344 (37.8)	79 (31.3)	265 (40.3)	44 (4.6)	0.67 (0.49-0.92)	0.013
Back pain	132 (14.5)	33 (13.1)	99 (15.0)	44 (4.6)	0.85 (0.55-1.29)	0.455
CLINICAL ASSESSMENT						
Positive stethoscope findings	329 (36.5)	168 (65.4)	161 (25.0)	52 (5.4)	5.67 (4.15-7.75)	<0.001
Abdominal pain by palpation	192 (22.1)	37 (15.0)	155 (25.0)	86 (9.0)	0.52 (0.35-0.78)	0.002
COMORBIDITIES						
Dementia	32 (3.4)	9 (3.4)	23 (3.3)	0 (0.0)	1.01 (0.46-2.22)	0.964
Neurological diseases	172 (18.0)	53 (20.0)	119 (17.3)	0 (0.0)	1.19 (0.83-1.71)	0.326
Respiratory diseases	269 (28.2)	105 (39.6)	164 (23.8)	0 (0.0)	2.10 (1.55-2.84)	<0.001
Endocrinological diseases	296 (31.0)	80 (30.2)	216 (31.3)	0 (0.0)	0.94 (0.69-1.28)	0.728
Nephrological diseases	252 (26.4)	60 (22.6)	192 (27.9)	0 (0.0)	0.75 (0.54-1.05)	0.101
Cardiovascular diseases	390 (40.9)	116 (43.8)	274 (39.8)	0 (0.0)	1.17 (0.88-1.57)	0.259
Gastrointestinal diseases	100 (10.5)	23 (8.7)	77 (11.2)	0 (0.0)	0.75 (0.46-1.23)	0.260
Rheumatological diseases	118 (12.4)	27 (10.2)	91 (13.2)	0 (0.0)	0.74 (0.47-1.17)	0.205
Cancer diseases	85 (8.9)	26 (9.8)	59 (8.6)	0 (0.0)	1.16 (0.71-1.88)	0.544
Prior pneumonia				100 (10.5)		
No	410 (48.0)	79 (33.3)	331 (53.6)		1 (reference)	
Yes, one time	180 (21.1)	50 (21.1)	130 (21.1)		1.61 (1.07-2.42)	0.022
Yes, more than one time	264 (30.9)	108 (45.6)	156 (25.3)		2.90 (2.05-4.10)	<0.001
SEVERITY ASSESSMENT						
CURB65 ≥3 **	122 (13.0)	29 (11.3)	93 (13.7)	16 (1.7)	0.80 (0.51-1.25)	0.336
Triage***				59 (6.2)		
Green/Blue	183 (20.4)	37 (14.8)	146 (22.6)		1 (reference)	
Yellow	479 (53.5)	126 (50.4)	353 (54.7)		1.40 (0.93-2.13)	0.105
Red/Orange	233 (26.0)	87 (34.8)	146 (22.6)		2.35 (1.50-3.67)	<0.001
VITAL PARAMETERS						
Respiratory rate, median(IQR)	18.0 (16.0; 22.0)	20.0 (18.0; 24.0)	18.0 (16.0; 20.0)	5 (0.5)	1.10 (1.07-1.13)	<0.001
Respiratory rate >20/min	285 (30.0)	124 (47.0)	161 (23.5)	5 (0.5)	2.88 (2.13-3.88)	<0.001
Oxygen saturation % n/min, median (IQR)	96.0 (94.0; 98.0)	95.0 (93.0; 97.0)	97.0 (95.0; 98.0)	4 (0.4)	0.84 (0.80-0.88)	<0.001
Oxygen saturation < 96 %	393 (41.4)	162 (61.1)	231 (33.7)	4 (0.4)	3.09 (2.30-4.14)	<0.001
Heart rate/min, mean (sd)	90.1 (18.3)	93.2 (18.9)	88.9 (18.0)	1 (0.1)	1.01 (1.005-1.02)	0.001
Heart rate <51 or >90/min	460 (48.3)	148 (55.8)	312 (45.3)	1 (0.1)	1.52 (1.14-2.02)	0.003
Systolic blood pressure mmHg, mean (sd)	132.8 (22.5)	134.2 (21.0)	132.2 (23.1)	3 (0.3)	1.003 (0.99-1.01	0.215

Systolic blood pressure <111 or >219 mmHg	156 (16.4)	38 (14.4)	118 (17.2)	3 (0.3)	0.81 (0.54-1.21)	0.314
Diastolic blood pressure mmHg, mean (sd)	74.8 (15.3)	74.2 (13.6)	75.0 (15.8)	3 (0.3)	0.99 (0.98-1.006)	0.483
Diastolic blood pressure ≤60 mmHg	163 (17.1)	40 (15.2)	123 (17.9)	3 (0.3)	0.82 (0.55-1.21)	0.329
Temperature, mean (SD)	37.5 (1.0)	37.6 (1.0)	37.4 (0.9)	5 (0.5)	1.22 (1.05-1.40)	0.006
Fever > 38°C	233 (24.6)	77 (29.3)	156 (22.7)	5 (0.5)	1.40 (1.02-1.93)	0.036
Glascow coma scale <15	31 (3.3)	12 (4.6)	19 (2.8)	5 (0.5)	0.59 (0.28-1.24)	0.168
BLOOD TESTS						
Haematocrit, median (IQR)	38.0 (35.0; 42.0)	38.0 (35.0; 42.0)	39.0 (35.0; 42.0)	260 (27.2)	0.98 (0.95-1.01)	0.465
Haematocrit	268 (38.6)	85 (38.6)	183 (38.6)	260 (27.2)	1.001 (0.72-1.39)	0.994
Haemoglobin mmol/L, median (IQR)	8.0 (7.2; 8.7)	7.9 (7.2; 8.6)	8.0 (7.3; 8.8)	0 (0.0)	0.90 (0.80-1.02)	0.127
Haemoglobin mmol/L	402 (42.1)	118 (44.5)	284 (41.2)	0 (0.0)	1.14 (0.86-1.52)	0.354
Leukocytes 10E9/L, median (IQR)	11.1 (8.3; 14.8)	12.2 (9.5; 15.8)	10.7 (8.0; 14.2)	0 (0.0)	1.05 (1.02-1.07)	<0.001
Leukocytes 10E9/L	670 (70.2)	214 (80.8)	456 (66.2)	0 (0.0)	2.14 (1.52-3.02)	<0.001
Platelets 10E9/L, median (IQR)	240.0 (189.0; 307.8)	260.5 (211.0; 330.8)	232.0 (182.3; 296.0)	10 (1.0)	1.002 (1.001-1.004)	<0.001
Platelets 10E9/L	201 (21.3)	63 (23.9)	138 (20.3)	10 (1.0)	1.23 (0.87-1.72)	0.229
Neutrophilocytes 10E9/L, median (IQR)	8.4 (6.0; 12.2)	9.7 (7.2; 13.0)	8.0 (5.6; 11.6)	10 (1.0)	1.06 (1.03-1.09)	<0.001
Neutrophilocytes 10E9/L	549 (58.2)	187 (71.1)	362 (53.2)	10 (1.0)	2.16 (1.59-2.94)	<0.001
Lymphocytes† 10E9/L, median (IQR)	1.1 (0.7; 1.6)	0.9 (0.6; 1.5)	1.2 (0.8; 1.8)	633 (66.3)	0.98 (0.85-1.12)	0.797
Lymphocytes† 10E9/L	145 (45.2)	53 (55.2)	92 (40.9)	633 (66.3)	1.78 (1.10-2.88)	0.018
Albumin g/L, median (IQR)	39.0 (36.0; 42.0)	39.0 (35.0; 41.0)	39.0 (36.0; 42.0)	7 (0.7)	0.96 (0.93-0.99)	0.029
Albumin g/L	160 (16.9)	39 (14.9)	121 (17.6)	7 (0.7)	0.82 (0.55-1.21)	0.323
Creatinine µmol/L, median (IQR)	84.0 (67.0; 113.0)	81.0 (64.0; 108.0)	86.0 (67.5; 114.0)	0 (0.0)	0.996 (0.993-0.998)	0.003
Creatinine µmol/L	374 (39.2)	106 (40.0)	268 (38.9)	0 (0.0)	1.04 (0.78-1.39)	0.754
Blood urea nitrogen mmol/L, median (IQR)	6.2 (4.4; 8.9)	6.2 (4.5; 8.6)	6.2 (4.4; 9.1)	9 (0.9)	0.99 (0.96-1.02)	0.657
Blood urea nitrogen mmol/L	377 (39.9)	99 (38.1)	278 (40.6)	9 (0.9)	0.90 (0.67-1.20)	0.482
Natrium mmol/L, median (IQR)	137.0 (134.0; 139.0)	137.0 (134.0; 139.0)	137.0 (134.0; 139.0)	0 (0.0)	0.98 (0.95-1.01)	0.394
Natrium mmol/L	432 (45.3)	128 (48.3)	304 (44.1)	0 (0.0)	1.18 (0.89-1.57)	0.245
Prothrombin, median (IQR)	1.1 (1.0; 1.2)	1.1 (1.0; 1.2)	1.1 (1.0; 1.2)	3 (0.3)	1.18 (0.89-1.58)	0.231
Prothrombin	234 (24.6)	65 (24.5)	169 (24.6)	3 (0.3)	0.99 (0.71-1.38)	0.972
Bilirubin µmol/L, median (IQR)	9.0 (6.0; 13.0)	9.0 (6.0; 12.0)	9.0 (6.0; 14.0)	11 (1.1)	0.97 (0.95-0.99)	0.254
Bilirubin µmol/L	152 (16.1)	38 (14.4)	114 (16.8)	11 (1.1)	0.83 (0.55-1.24)	0.369
Glucose mmol/L, median (IQR)	6.7 (5.9; 7.9)	6.9 (6.2; 8.1)	6.6 (5.8; 7.8)	9 (0.9)	1.04 (0.99-1.10)	0.052
Glucose mmol/L	51 (5.4)	19 (7.3)	32 (4.7)	9 (0.9)	1.59 (0.88-2.85)	0.120
C-Reactive Protein mg/L, median (IQR)	95.5 (30.0; 179.3)	125.0 (57.0; 203.5)	82.0 (19.0; 172.0)	0 (0.0)	1.003 (1.001-1.004)	<0.001
C-Reactive Protein mg/L				0 (0.0)		
Low <20mg/L	196 (20.5)	21 (7.9)	175 (25.4)		1 (reference)	
Moderate 21-99 mg/L	291 (30.5)	86 (32.5)	205 (29.8)		3.49 (2.08-5.86)	<0.001
High >=100	467 (49.0)	158 (59.6)	309 (44.8)		4.26 (2.60-6.96)	<0.001
VACCINE AND MEDICAMENTATIONS						
SARS-CoV-2 †	756 (79.2)	222 (83.8)	534 (77.5)	0 (0.0)	1.49 (1.03-2.17)	0.033
Pneumococcal	530 (55.6)	160 (60.4)	370 (53.7)	0 (0.0)	1.31 (0.98-1.75)	0.063
Influenza	635 (66.6)	191 (72.1)	444 (64.4)	0 (0.0)	1.42 (1.04-1.94)	0.025

Analgesics	404 (42.3)	115 (43.4)	289 (41.9)	0 (0.0)	1.06 (0.79-1.41)	0.684
Polypharmacy****	544 (57.0)	163 (61.5)	381 (55.3)	0 (0.0)	1.29 (0.96-1.72)	0.082

Values are numbers (percentages) unless otherwise specified. *ADL dependence: If the patient had one or more dependencies regarding bathing, dressing, toileting, transfer, continence, and feeding. ** CURB65: confusion, uraemia, respiratory rate, blood pressure, age > 65 years. ***Triage: Danish emergency process triage [40] ****Polypharmacy: regular consumption of at least five medications † variables not included in the multivariate model

Table S3: Characteristics of the 954 patients with suspected infection enrolled in the study. It presents the 70 predictors included in the multivariate analysis and randomization of the training set and validation set.

Characteristics	Total, n	Training set, n	Validation set, n	Missings n (%)	p-value
Total of patients	954 (100)	766 (80.3)	188 (19.7)	0 (0.0)	
DEMOGRAPHIC DATA					
Age, median (IQR)	73.0 (59.0; 81.0)	75.0 (63.5; 82.0)	74.0 (60.0; 82.0)	0 (0.0)	0.54
Age ≥75 years	440 (46.1)	348 (45.4)	92 (48.9)	0 (0.0)	0.39
Gender male	513 (53.8)	408 (53.3)	105 (55.9)	0 (0.0)	0.52
Marital status, Living alone	618 (66.0)	488 (65.0)	130 (70.3)	18 (1.9)	0.17
Nursing home resident	66 (7.0)	55 (7.3)	11 (5.9)	13 (1.4)	0.53
Occupation				21 (2.2)	0.62
Others	67 (7.2)	57 (7.6)	10 (5.5)		
Working	202 (21.7)	162 (21.6)	40 (22.0)		
Retired	664 (71.2)	532 (70.8)	132 (72.5)		
LIFESTYLE FACTORS					
Smoking status				33 (3.5)	0.76
No	323 (35.1)	256 (34.5)	67 (37.4)		
Current smoker	179 (19.4)	145 (19.5)	34 (19.0)		
Previous smoker	419 (45.5)	341 (46.0)	78 (43.6)		
Alcohol status				35 (3.7)	0.60
No alcohol	356 (38.7)	283 (38.2)	73 (40.8)		
1-7 doses	385 (41.9)	315 (42.6)	70 (39.1)		
8-14 doses	105 (11.4)	81 (10.9)	24 (13.4)		
> 14 doses	73 (7.9)	61 (8.2)	12 (6.7)		
Physically activity				52 (5.4)	0.76
Not physical active	263 (29.2)	214 (29.4)	49 (28.2)		
Physical activity < 2,5 hr/week	231 (25.6)	189 (26.0)	42 (24.1)		
Physical activity ≥ 2,5 hr/week	408 (45.2)	325 (44.6)	83 (47.7)		
Body Mass Index†				249 (26.1)	0.74

Healthy weight	246 (34.9)	202 (35.8)	44 (31.2)		
Obese	193 (27.4)	154 (27.3)	39 (27.7)		
Overweight	239 (33.9)	187 (33.2)	52 (36.9)		
Underweight	27 (3.8)	21 (3.7)	6 (4.3)		
ADL dependence*	260 (28.0)	203 (27.1)	57 (31.7)	25 (2.6)	0.22
SYMPTOMS					
Malaise	559 (61.2)	458 (62.0)	101 (58.0)	41 (4.3)	0.34
Feeling tired	657 (72.6)	540 (74.0)	117 (66.9)	49 (5.1)	0.06
Headache	351 (38.3)	287 (38.8)	64 (36.0)	37 (3.9)	0.48
Dizziness	346 (37.7)	287 (38.8)	59 (33.1)	37 (3.98)	0.16
Confusion	207 (22.6)	164 (22.2)	43 (24.2)	37 (3.89)	0.57
Dyspnea	379 (41.4)	309 (42.0)	70 (39.1)	39 (4.1)	0.48
Cough	358 (39.1)	294 (39.9)	64 (35.8)	39 (4.1)	0.30
Fever feeling at home	612 (64.2)	464 (64.5)	118 (62.8)	0 (0.0)	0.66
Expectoration	279 (30.5)	224 (30.4)	55 (30.7)	39 (4.1)	0.94
Sore throat	104 (11.4)	86 (11.7)	18 (10.1)	39 (4.1)	0.54
Cold (common cold)	95 (10.4)	81 (11.0)	14 (7.8)	39 (4.1)	0.21
Chest pain	168 (18.4)	134 (18.2)	34 (19.0)	40 (4.2)	0.81
Oedema	79 (8.6)	61 (8.3)	18 (10.1)	39 (4.1)	0.45
Nausea	304 (33.2)	247 (33.4)	57 (32.2)	38 (3.9)	0.76
Vomiting	190 (20.7)	154 (20.8)	36 (20.3)	38 (3.9)	0.88
Loss of appetite	524 (57.2)	424 (57.4)	100 (56.5)	38 (3.9)	0.83
Gastrointestinal pain	193 (21.1)	145 (19.6)	48 (27.1)	38 (3.9)	0.03
Diarrhoea	134 (14.6)	107 (14.5)	27 (15.3)	38 (3.9)	0.79
Muscular pain	344 (37.8)	289 (39.5)	55 (30.9)	44 (4.6)	0.03
Back pain	132 (14.5)	110 (15.0)	22 (12.4)	44 (4.6)	0.36
CLINICAL ASSESSMENT					
Positive stethoscope findings	329 (36.5)	263 (36.5)	66 (36.5)	52 (5.4)	1.00
Abdominal pain by palpation	192 (22.1)	151 (21.7)	41 (23.7)	86 (9.0)	0.58
COMORBIDITIES					
Dementia	23 (3.0)	9 (4.8)	23 (3.3)	0 (0.0)	0.22
Neurological diseases	137 (17.9)	35 (18.6)	119 (17.3)	0 (0.0)	0.82
Pulmonary diseases	212 (27.7)	57 (30.3)	164 (23.8)	0 (0.0)	0.47
Endocrinological diseases	239 (31.2)	57 (30.3)	216 (31.3)	0 (0.0)	0.81
Nephrological diseases	200 (26.1)	52 (27.7)	192 (27.9)	0 (0.0)	0.67

Cardiovascular diseases	303 (39.6)	87 (46.3)	274 (39.8)	0 (0.0)	0.09
Gastrointestinal diseases	81 (10.6)	19 (10.1)	77 (11.2)	0 (0.0)	0.85
Rheumatological diseases	93 (12.1)	25 (13.3)	91 (13.2)	0 (0.0)	0.67
Cancer diseases	66 (8.6)	19 (10.1)	59 (8.6)	0 (0.0)	0.52
Prior pneumonia				100 (10.5)	0.05
No	343 (50.1)	67 (39.6)	331 (53.6)		
Yes, one time	139 (20.3)	41 (24.3)	130 (21.1)		
Yes, more than one time	203 (29.6)	61 (36.1)	156 (25.3)		
SEVERITY ASSESSMENT					
CURB65 ≥3 **	103 (13.6)	19 (10.4)	93 (13.7)	16 (1.7)	0.25
Triage***				59 (6.2)	0.53
Green/Blue	185 (25.6)	48 (27.9)	146 (22.6)		
Yellow	385 (53.3)	94 (54.7)	353 (54.7)		
Red/Orange	153 (21.2)	30 (17.4)	146 (22.6)		
VITAL PARAMETERS					
Respiratory rate >20/min	285 (30.0)	235 (30.8)	50 (26.7)	5 (0.5)	0.27
Oxygen saturation < 96 %	393 (41.4)	324 (42.5)	69 (36.7)	4 (0.4)	0.15
Heart rate <51 or >90/min	460 (48.3)	377 (49.3)	83 (44.1)	1 (0.1)	0.21
Systolic blood pressure <111 or >219 mmHg	156 (16.4)	125 (16.4)	31 (16.6)	3 (0.3)	0.94
Diastolic blood pressure ≤60 mmHg	163 (17.1)	131 (17.1)	32 (17.1)	3 (0.3)	0.99
Fever > 38°C	233 (24.6)	190 (24.9)	43 (23.1)	5 (0.5)	0.61
Glascow coma scale <15	31 (3.3)	23 (3.0)	8 (4.3)	5 (0.5)	0.39
BLOOD TESTS					
Haematocrit	268 (38.6)	218 (39.2)	50 (36.2)	260 (27.2)	0.52
Haemoglobin mmol/L	402 (42.1)	329 (43.0)	73 (38.8)	0 (0.0)	0.31
Leukocytes 10E9/L	670 (70.2)	548 (71.5)	122 (64.9)	0 (0.0)	0.07
Platelets 10E9/L	201 (21.3)	168 (22.2)	33 (17.6)	10 (1.0)	0.17
Neutrophilocytes 10E9/L	549 (58.2)	454 (59.9)	95 (51.1)	10 (1.0)	0.03
Albumin g/L	160 (16.9)	130 (17.1)	30 (16.1)	7 (0.7)	0.76
Creatinine µmol/L	374 (39.2)	303 (39.6)	71 (37.8)	0 (0.0)	0.65
Blood urea nitrogen mmol/L	377 (39.9)	308 (40.5)	69 (37.5)	9 (0.9)	0.46
Natrium mmol/L	432 (45.3)	362 (47.3)	70 (37.2)	0 (0.0)	0.01
Prothrombin	234 (24.6)	186 (24.3)	48 (25.7)	3 (0.3)	0.71
Bilirubin µmol/L	152 (16.1)	119 (15.7)	33 (17.8)	11 (1.1)	0.48

Glucose mmol/L	51 (5.4)	42 (5.5)	9 (4.8)	9 (0.9)	0.71
C-Reactive Protein mg/L				0 (0.0)	0.07
<20 mg/L	196 (20.5)	151 (19.7)	45 (23.9)		
21-99 mg/L	291 (30.5)	226 (29.5)	65 (34.6)		
≥ 100 mg/L	467 (49.0)	389 (50.8)	78 (41.5)		
VACCINE AND MEDICAMENTATIONS					
Pneumococcal	530 (55.6)	414 (54.0)	116 (61.7)	0 (0.0)	0.06
Influenza	635 (66.6)	512 (66.8)	123 (65.4)	0 (0.0)	0.71
Analgesics	404 (42.3)	336 (43.9)	68 (36.2)	0 (0.0)	0.06
Polypharmacy****	544 (57.0)	443 (57.8)	101 (53.7)	0 (0.0)	0.31

Values are numbers (percentages) unless otherwise specified. *ADL dependence: If the patient had one or more dependencies regarding bathing, dressing, toileting, transfer, continence, and feeding. ** CURB65: confusion, uraemia, respiratory rate, blood pressure, age > 65 years. ***Triage: Danish emergency process triage [40] ****Polypharmacy: regular consumption of at least five medications

Figure S4: Performance of the prediction model presented with the area receiver operating characteristic curve

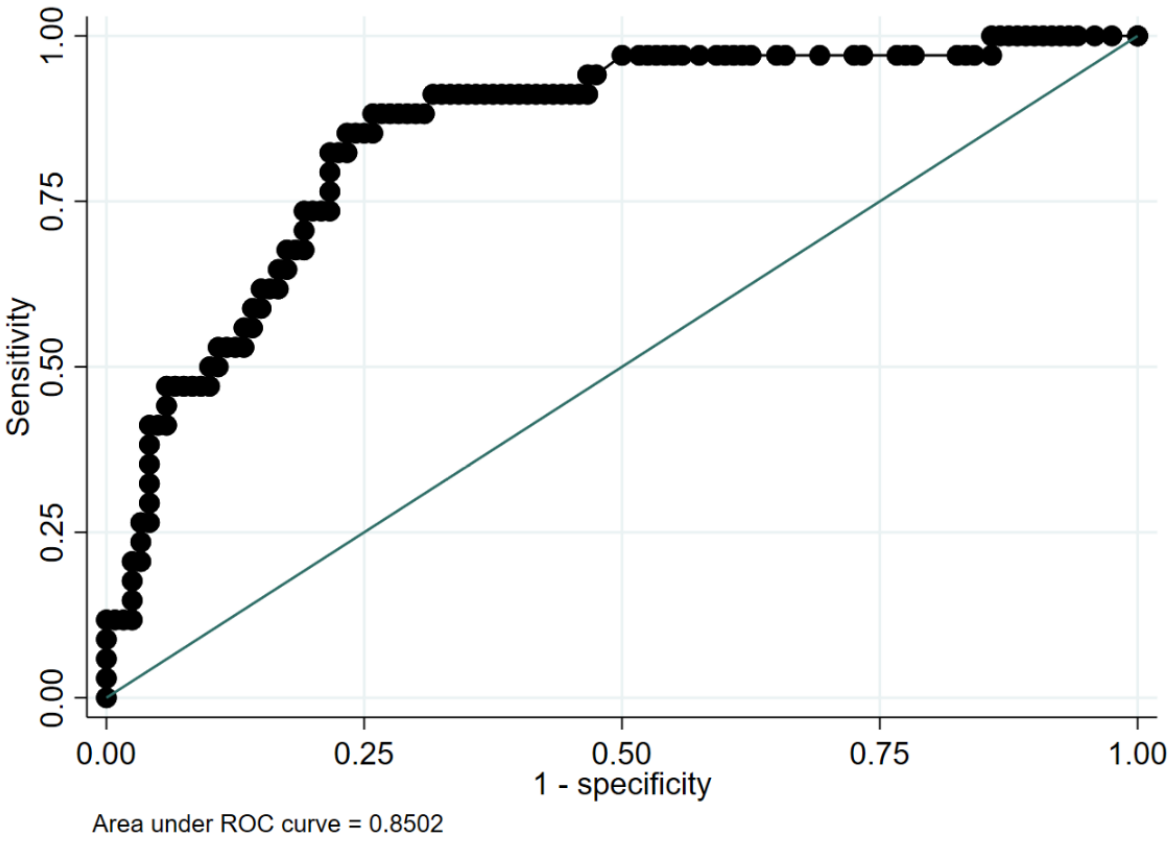
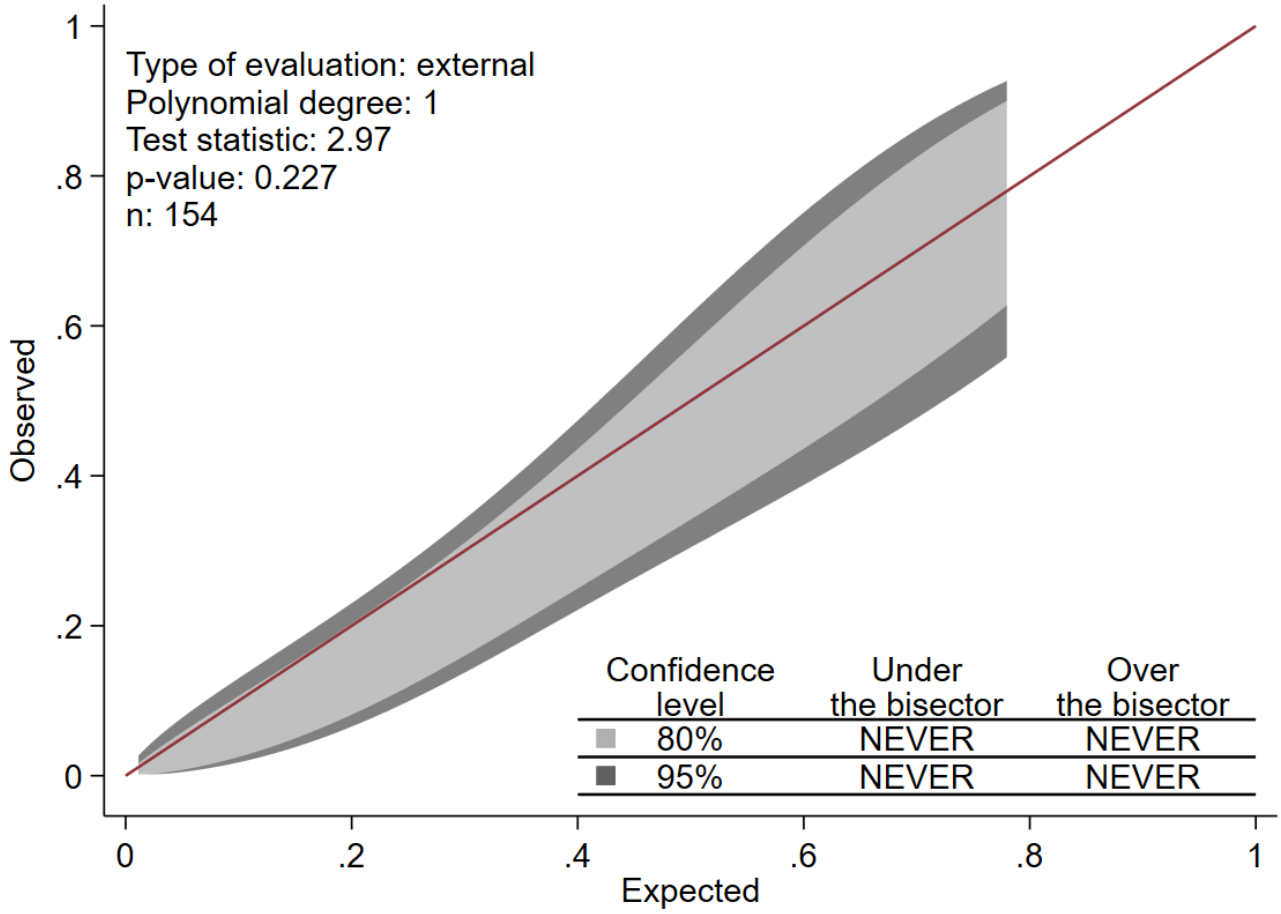


Figure S5: The calibration of the model after recalibration



Formula S6: Based on a lambda result of λ=0.0402856 and a probability threshold of 0.35, the LASSO calculation with characteristics predictive of CAP as follows:

$$
\begin{aligned}
CAP-score = {} & 0.07 \cdot 1_{Unwell=yes} + 0.35 \cdot 1_{Dyspnea=yes} + 0.36 \cdot 1_{Expectoration=yes} + 0.39 \cdot 1_{Cough=yes} \\
& + 0.34 \cdot 1_{Cold=yes} + 0.14 \cdot 1_{Respiratory\ rate\ >20/min=yes} + 0.24 \\
& \cdot 1_{Oxygen\ saturation<96\%=yes} + 0.20 \cdot 1_{Chest\ pain=yes} + 0.56 \cdot 1_{Stethoscope=yes} - 0.12 \\
& \cdot 1_{Previous\ CAP=no} + 0.003 \cdot 1_{Leucocytes<3.5\ or\ >8.8\ 10E9\ /L=yes} + 0.08 \\
& \cdot 1_{Neutrophilocytes>7.5\ 10E9\ /L=yes} - 0.64 \cdot 1_{CRP<20mg\ /L=yes} + 0.53 \cdot 1_{Cough=yes} \cdot 1_{age\geq 75} \\
& - 0.05.\ 1_{Edema=yes} \cdot 1_{age\geq 75} + 0.88 \cdot 1_{Glucose>11\ mmol\ /L=yes} \cdot 1_{age\geq 75} + 0.0402856 \\
& \cdot (0.07 + 0.35 + 0.36 + 0.39 + 0.015 + 0.34 + 0.14 + 0.24 + 0.20 + 0.56 + 0.12 \\
& + 0.003 + 0.08 + 0.64 + 0.53 + 0.05 + 0.88) - 1.66192 - \log\left(\frac{0.35}{0.65}\right)
\end{aligned}
$$

For best calibration, 0.07 must be subtracted from the score if the score is between 0.08 and 0.47.

Formula S7: A cutoff value greater than 0 indicates the diagnosis CAP according to our model and can be calculated using the following formula:

$$
\begin{aligned}
CAP - score = {} & 0.07 \cdot 1_{Unwell=yes} + 0.35 \cdot 1_{Dyspnea=yes} + 0.36 \cdot 1_{Expectoration=yes} + 0.39 \cdot 1_{Cough=yes} \\
& + 0.34 \cdot 1_{Cold=yes} + 0.14 \cdot 1_{Respiratory\ rate\ >20/min=yes} + 0.24 \\
& \cdot 1_{Oxygen\ saturation<96\%=yes} + 0.20 \cdot 1_{Chest\ pain=yes} + 0.56 \cdot 1_{Stethoscope=yes} - 0.12 \\
& \cdot 1_{Previous\ CAP=no} + 0.003 \cdot 1_{Leucocytes<3.5\ or\ >8.8\ 10E9\ /L=yes} + 0.08 \\
& \cdot 1_{Neutrophilocytes>7.5\ 10E9\ /L=yes} - 0.64 \cdot 1_{CRP<20mg\ /L=yes} + 0.53 \cdot 1_{Cough=yes} \cdot 1_{age\geq 75} \\
& - 0.05 \cdot 1_{Edema=yes} \cdot 1_{age\geq 75} + 0.88 \cdot 1_{Glucose>11\ mmol\ /L=yes} \cdot 1_{age\geq 75} - 0.842742
\end{aligned}
$$

For best calibration, 0.07 must be subtracted from the score if the score is between 0.08 and 0.47.

Model specification

Besides the high percentage of missings from lymphocytes (66.3%), lymphocytes contributed to a significantly decreased model performance below 80% and a narrower calibration belt (p<0.001), furthermore lymphocytes were missing for 66.3% of the patients. SARS-CoV-2 vaccine was not included in the final model as the vaccine was related to a specific pandemic and did not change any final predictors or values. The inclusion of the BMI had better prediction performance AUC: 0.86 (CI: 0.79-0.93) and yielded more predictors especially related to lifestyle. The predictors that differed from the final model were: Alcohol (8-14 doses/week) 0.01792, level of physical activity under 2,5 hours/week yielded 0.01067, and obesity appeared with a coefficient of -0.93861. In addition, a symptom of diarrhea (-0.17572), muscular pain (-0.00225), gastrointestinal symptoms (-0.807885), sore throat (0.074709 for patients ≥ 75 years old) and the presence of nephrological diseases (-0.18776 for patients ≥ 75 years old) were predictors of CAP in the model constructed including BMI. From a clinical perspective, we chose to exclude the BMI as the final model would be more useful in an acute setting where reliable information about BMI is not always available. From a statistical perspective, BMI had almost 27% of missings, which would be classified as MAR and possibly selected from the population.

References

1. Almirall, J., et al., *Risk Factors for Community-Acquired Pneumonia in Adults: A Systematic Review of Observational Studies.* Respiration, 2017. **94**(3): p. 299-311.
2. Riquelme, R., et al., *Community-acquired pneumonia in the elderly. Clinical and nutritional aspects.* Am J Respir Crit Care Med, 1997. **156**(6): p. 1908-14.
3. Janssens, J.P., *Pneumonia in the elderly (geriatric) population.* Curr Opin Pulm Med, 2005. **11**(3): p. 226-30.
4. Cillóniz, C., et al., *Community-acquired pneumonia in critically ill very old patients: a growing problem.* Eur Respir Rev, 2020. **29**(155).
5. Metlay, J.P., et al., *Influence of age on symptoms at presentation in patients with community-acquired pneumonia.* Arch Intern Med, 1997. **157**(13): p. 1453-9.
6. Laporte, L., et al., *Ten-year trends in intensive care admissions for respiratory infections in the elderly.* Ann Intensive Care, 2018. **8**(1): p. 84.
7. Ravioli, S., et al., *Age- and sex-related differences in community-acquired pneumonia at presentation to the emergency department: a retrospective cohort study.* Eur J Emerg Med, 2022. **29**(5): p. 366-372.
8. Hammond, A., et al., *Predisposing factors to acquisition of acute respiratory tract infections in the community: a systematic review and meta-analysis.* BMC Infect Dis, 2021. **21**(1): p. 1254.
9. Barbagelata, E., et al., *Gender differences in community-acquired pneumonia.* Minerva Med, 2020. **111**(2): p. 153-165.
10. Baik, I., et al., *A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women.* Arch Intern Med, 2000. **160**(20): p. 3082-8.
11. Heath, G.W., et al., *Exercise and the incidence of upper respiratory tract infections.* Medicine and science in sports and exercise, 1991. **23**(2): p. 152-157.
12. Kim, N.E., et al., *Clinical characteristics and outcomes among older nursing home residents hospitalized with pneumonia.* Arch Gerontol Geriatr, 2021. **95**: p. 104394.
13. Nieman, D.C., et al., *Upper respiratory tract infection is reduced in physically fit and active adults.* British journal of sports medicine, 2011. **45**(12): p. 987-992.
14. Calvillo–King, L., et al., *Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review.* Journal of general internal medicine, 2013. **28**(2): p. 269-282.
15. Clyde, W.A., Jr., *Clinical overview of typical Mycoplasma pneumoniae infections.* Clin Infect Dis, 1993. **17 Suppl 1**: p. S32-6.
16. Gamble, J.M., et al., *Medication transitions and polypharmacy in older adults following acute care.* Ther Clin Risk Manag, 2014. **10**: p. 189-96.
17. Torres, A., et al., *Risk factors for community-acquired pneumonia in adults in Europe: a literature review.* Thorax, 2013. **68**(11): p. 1057-65.
18. Riquelme, R., et al., *Community-acquired pneumonia in the elderly: clinical and nutritional aspects.* Revista médica de Chile, 2008. **136**(5): p. 587-593.
19. Moore, M., et al., *Predictors of pneumonia in lower respiratory tract infections: 3C prospective cough complication cohort study.* Eur Respir J, 2017. **50**(5).
20. Ding, F., et al., *Development and validation of a simple tool composed of items on dyspnea, respiration rates, and C-reactive protein for pneumonia prediction among acute febrile respiratory illness patients in primary care settings.* BMC Med, 2022. **20**(1): p. 360.
21. Nakanishi, M., et al., *Significance of the progression of respiratory symptoms for predicting community-acquired pneumonia in general practice.* Respirology, 2010. **15**(6): p. 969-74.

22. Huijskens, E.G.W., et al., *The value of signs and symptoms in differentiating between bacterial, viral and mixed aetiology in patients with community-acquired pneumonia.* J Med Microbiol, 2014. **63**(Pt 3): p. 441-452.
23. Loubet, P., et al., *Community-acquired pneumonia in the emergency department: an algorithm to facilitate diagnosis and guide chest CT scan indication.* Clin Microbiol Infect, 2020. **26**(3): p. 382.e1-382.e7.
24. Beekman, R., et al., *Validating a clinical prediction score for Legionella-related community acquired pneumonia.* BMC Infect Dis, 2022. **22**(1): p. 442.
25. den Engelsen, C., et al., *Infectious diseases and the use of antibiotics in outpatients at the emergency department of the University Hospital of León, Nicaragua.* Int J Infect Dis, 2009. **13**(3): p. 349-54.
26. Mandell, L.A., *Community-acquired pneumonia: An overview.* Postgrad Med, 2015. **127**(6): p. 607-15.
27. Takase, R., et al., *Clinical Manifestations of Patients with Influenza Differ by Age : A Prospective, Multi-centered Study in the Setouchi Marine Area.* Acta Med Okayama, 2021. **75**(5): p. 567-574.
28. Akhtar, A., et al., *Respiratory-tract infections among geriatrics: prevalence and factors associated with the treatment outcomes.* Therapeutic advances in respiratory disease, 2021. **15**: p. 1753466620971141.
29. Sundhedsstyrelsen. *Sundhedsstyrelsens udmeldinger om alkohol*. 2022 [cited 2022 December 06]; Available from: https://www.sst.dk/da/Viden/Forebyggelse/Alkohol/Alkoholforebyggelse/Sundhedsstyrelsens-udmeldinger-om-alkohol
30. (WHO), W.H.O. [cited 2022 December 6]; Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://apps.who.int/iris/bitstream/handle/10665/337001/9789240014886-eng.pdf
31. Shang, J., et al., *Risk factors for infection in home health care: Analysis of national Outcome and Assessment Information Set data.* Res Nurs Health, 2020. **43**(4): p. 373-386.
32. Guidet, B., et al., *Caring for the critically ill patients over 80: a narrative review.* Ann Intensive Care, 2018. **8**(1): p. 114.
33. Steffens, C., et al., *The Association Between Prescribed Opioid Receipt and Community-Acquired Pneumonia in Adults: a Systematic Review and Meta-analysis.* J Gen Intern Med, 2020. **35**(11): p. 3315-3322.
34. Walters, J.A., et al., *Pneumococcal vaccines for preventing pneumonia in chronic obstructive pulmonary disease.* Cochrane Database Syst Rev, 2017. **1**(1): p. Cd001390.
35. Kraicer-Melamed, H., S. O'Donnell, and C. Quach, *The effectiveness of pneumococcal polysaccharide vaccine 23 (PPV23) in the general population of 50 years of age and older: A systematic review and meta-analysis.* Vaccine, 2016. **34**(13): p. 1540-1550.
36. Liang, C.Y., et al., *Effectiveness of influenza vaccination in the elderly: a population-based case-crossover study.* BMJ Open, 2022. **12**(2): p. e050594.
37. Chalmers, J.D., et al., *Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis.* Thorax, 2010. **65**(10): p. 878-83.
38. Lim, W.S., et al., *BTS guidelines for the management of community acquired pneumonia in adults: update 2009.* Thorax, 2009. **64 Suppl 3**: p. iii1-55.
39. Rosenvinge, F.S. *Antibiotikavejledning for Region Syddanmark*. 06.10.2021 [cited 2022 22 september]; Available from: https://ekstern.infonet.regionsyddanmark.dk/Files/Dokument547684.htm.
40. Plesner, L.L., et al., *The formation and design of the TRIAGE study-baseline data on 6005 consecutive patients admitted to hospital from the emergency department.* Scandinavian journal of trauma, resuscitation and emergency medicine, 2015. **23**(1): p. 1-9.

41. *User Manuel Danish Emergency Process Triage*.
42. Nordberg, M., S. Lethvall, and M. Castrén, *The validity of the triage system ADAPT.* Scandinavian journal of trauma, resuscitation and emergency medicine, 2010. **18**: p. 1-1.
43. Farrohknia, N., et al., *Emergency department triage scales and their components: a systematic review of the scientific evidence.* Scand J Trauma Resusc Emerg Med, 2011. **19**: p. 42.
44. Jones, M., *NEWSDIG: The National Early Warning Score Development and Implementation Group.* Clin Med (Lond), 2012. **12**(6): p. 501-3.
45. Htun, T.P., et al., *Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review.* Sci Rep, 2019. **9**(1): p. 7600.
46. Gong, L., et al., *Clinical profile analysis and nomogram for predicting in-hospital mortality among elderly severe community-acquired pneumonia patients with comorbid cardiovascular disease: a retrospective cohort study.* BMC Pulm Med, 2022. **22**(1): p. 312.
47. Sakakibara, T., et al., *A prediction rule for severe adverse events in all inpatients with community-acquired pneumonia: a multicenter observational study.* BMC pulmonary medicine, 2022. **22**(1): p. 34.
48. Mogensen, C.B., et al., *Ear measurement of temperature is only useful for screening for fever in an adult emergency department.* BMC Emerg Med, 2018. **18**(1): p. 51.
49. Mackowiak, P.A., F.A. Chervenak, and A. Grünebaum, *Defining Fever.* Open Forum Infect Dis, 2021. **8**(6): p. ofab161.
50. Waterer, G.W., L.A. Kessler, and R.G. Wunderink, *Medium-term survival after hospitalization with community-acquired pneumonia.* Am J Respir Crit Care Med, 2004. **169**(8): p. 910-4.
51. Zhao, L.H., J. Chen, and R.X. Zhu, *The relationship between frailty and community-acquired pneumonia in older patients.* Aging Clin Exp Res, 2023. **35**(2): p. 349-355.
52. Kitazawa, T., et al., *Characteristics of pneumonia with negative chest radiography in cases confirmed by computed tomography.* J Community Hosp Intern Med Perspect, 2020. **10**(1): p. 19-24.
53. Huang, Y., et al., *Diagnostic value of blood parameters for community-acquired pneumonia.* Int Immunopharmacol, 2018. **64**: p. 10-15.
54. Alzoubi, O. and A. Khanfar, *Association between neutrophil to lymphocyte ratio and mortality among community acquired pneumonia patients: a meta-analysis.* Monaldi Arch Chest Dis, 2021. **92**(3).
55. Milas, G.P., V. Issaris, and V. Papavasileiou, *Blood urea nitrogen to albumin ratio as a predictive factor for pneumonia: A meta-analysis.* Respir Med Res, 2022. **81**: p. 100886.
56. Kassaw, G., et al., *Outcomes and Predictors of Severe Community-acquired Pneumonia Among Adults Admitted to the University of Gondar Comprehensive Specialized Hospital: A Prospective Follow-up Study.* Infect Drug Resist, 2023. **16**: p. 619-635.
57. Adnan, M., et al., *Prognostic value of five serum markers predicting in-hospital mortality among adults with community acquired pneumonia.* J Infect Dev Ctries, 2022. **16**(1): p. 166-172.
58. Rendón-Ramirez, E.J., et al., *TGF-β Blood Levels Distinguish Between Influenza A (H1N1)pdm09 Virus Sepsis and Sepsis due to Other Forms of Community-Acquired Pneumonia.* Viral Immunol, 2015. **28**(5): p. 248-54.
59. Watanabe, H., et al., *Clinical factors associated with negative urinary antigen tests implemented for the diagnosis of community-acquired pneumococcal pneumonia in adult patients.* Med Princ Pract, 2015. **24**(2): p. 189-94.
60. Zeng, W., et al., *Association of admission blood glucose level and clinical outcomes in elderly community-acquired pneumonia patients with or without diabetes.* Clin Respir J, 2022. **16**(8): p. 562-571.
61. Barmanray, R.D., et al., *In-hospital hyperglycemia but not diabetes mellitus alone is associated with increased in-hospital mortality in community-acquired pneumonia (CAP): a systematic review and meta-analysis of observational studies prior to COVID-19.* BMJ Open Diabetes Res Care, 2022. **10**(4).

62. van der Meer, V., et al., *Diagnostic value of C reactive protein in infections of the lower respiratory tract: systematic review.* Bmj, 2005. **331**(7507): p. 26.
63. van Vugt, S.F., et al., *Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study.* Bmj, 2013. **346**: p. f2450.
64. Ebell, M.H., et al., *Accuracy of Biomarkers for the Diagnosis of Adult Community-acquired Pneumonia: A Meta-analysis.* Acad Emerg Med, 2020. **27**(3): p. 195-206.
65. Ebell, M.H., et al., *Accuracy of Signs and Symptoms for the Diagnosis of Community-acquired Pneumonia: A Meta-analysis.* Acad Emerg Med, 2020. **27**(7): p. 541-553.
66. Division of Nutrition, P.A., and Obesity, National Center for Chronic Disease Prevention and Health Promotion. 3 June 2022 [cited 2023 2 March]; Available from: https://www.cdc.gov/obesity/basics/adult-defining.html.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	4
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	n/a
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6+77 + additional file (table S1 and S2)
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	7
Sample size	8	D;V	Explain how the study size was arrived at.	8
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8
	10c	V	For validation, describe how the predictions were calculated.	9
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	8
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n/a
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n/a
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	9
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10-12 (Table 1, Table 2) + additional file (table S2)
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	additional file (table S3)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	10-14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	10-14(Table 1) + table 2
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	14
	15b	D	Explain how to the use the prediction model.	14 + additional file (formula S6 +S7)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	13
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	14
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	17
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n/a
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	15+16+ 17
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	15+16+17+18
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	5
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	19

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2023-079123.R1
Article Type:	Original research
Date Submitted by the Author:	25-Apr-2024
Complete List of Authors:	Cartuliares, Mariana; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Mogensen, Christian Backer; University of Southern Denmark, Institute for Regional Health Research; University Hospital of Southern Denmark, Department of Emergency Medicine Rosenvinge, Flemming; Odense Universitetshospital, Department of Clinical Microbiology; University of Southern Denmark, Research Unit of Clinical Microbiology Skovsted, Thor; University Hospital of Southern Denmark, Department of Biochemistry and Immunology Lorentzen, Morten; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Heltborg, Anne; University Hospital of Southern Denmark, Department of Emergency Medicine ; University of Southern Denmark, Department of Regional Health Research Hertz, Mathias ; University of Southern Denmark, Department of Clinical Research; Odense University Hospital, Infectious Diseases Department Kaldan, Frida; University Hospital of Southern Denmark, Department of Emergency Medicine Specht, Jens ; University Hospital of Southern Denmark, Department of Emergency Medicine Skjøt-Arkil, Helene; University Hospital of Southern Denmark, Department of Emergency Medicine; University of Southern Denmark, Department of Regional Health Research
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Infectious diseases
Keywords:	ACCIDENT & EMERGENCY MEDICINE, INFECTIOUS DISEASES, Aged

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

TITLE

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Authors

Mariana Bichuette Cartuliares* [1, 2], Christian Backer Mogensen [1, 2], Flemming Schønning Rosenvinge [3, 4], Thor Aage Skovsted [5], Morten Hjarnø Lorentzen [1,2], Anne Heltborg[1,2], Mathias Amdi Hertz [6,7], Frida Kaldan[1], Jens Juel Specht[1], Helene Skjøt-Arkil [1,2]

*Corresponding author: Emergency Department, University Hospital of Southern Denmark, Kresten Philipsens vej 15, 6200 Aabenraa, Denmark; Email: mbc@rsyd.dk

Author affiliations

[1]Emergency Department, University Hospital of Southern Denmark, Aabenraa, Denmark

[2]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark

[3]Research Unit of Clinical Microbiology, University of Southern Denmark, Odense, Denmark

[4]Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark

[5]Department of Biochemistry and Immunology, University Hospital of Southern Denmark, Aabenraa, Denmark

[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[7] Infectious Diseases Department, Odense University Hospital, Odense, Denmark

ABSTRACT

Objectives: This study aimed to describe the clinical characteristics of adults with suspected acute community-acquired pneumonia (CAP) upon hospitalisation, evaluate their prediction performance for CAP and compare the performance of the model to the initial assessment of the physician.

Design: Cross-sectional, multicentre study.

Setting: The data originated from the INfectious DisEases in Emergency Departments study and were collected prospectively from patient interviews and medical records. The study included four Danish medical emergency departments (EDs) and was conducted between 1 March 2021 and 28 February 2022.

Participants: A total of 954 patients admitted with suspected infection were included in the study.

Primary and secondary outcome: The primary outcome was CAP diagnosis assessed by an expert panel.

Results: According to expert evaluation, CAP had a 28% prevalence. Thirteen diagnostic predictors were identified using Least absolute shrinkage and selection operator regression to build the prediction model: dyspnea, expectoration, cough, common cold, malaise, chest pain, respiratory rate (>20/min), oxygen saturation (< 96%), abnormal chest auscultation, leukocytes (<3,5 or >8,8 10E9/L) and neutrophils (>7.5 10E9/L). C-reactive protein (<20 mg/L) and having no previous event of CAP contributed negatively to the final model. The predictors yielded good prediction performance for CAP with an area under the ROC of 0.85 [CI: 0.77-0.92]. However, the initial diagnosis made by the ED physician performed better, with an AUROC of 0.86 [CI:84%-89%].

Conclusion: Typical respiratory symptoms combined with abnormal vital signs and elevated infection biomarkers were predictors for CAP upon admission to an ED. The clinical value of the prediction model is questionable in our setting as it does not outperform the clinician's assessment. Further studies that add novel diagnostic tools and use imaging or serological markers are needed to improve a model that would help diagnose CAP in an ED setting more accurately.

Strength and limitations

-This was a multicentre study with prospectively collected data

-Least absolute shrinkage and selection operator regression was used to establish a score for CAP, and the performance of the diagnostic model was evaluated using the area under the receiver operating characteristic curve and calibration curves.

-This diagnostic prediction model could have been improved by adding other diagnostic tools, such as imaging or serological markers.

-Lack of external validation of the model using the clinical score for community-acquired pneumonia was a limitation

Keywords: community-acquired pneumonia; diagnostic prediction model; emergency department

Word count: 3.966

INTRODUCTION

Community-acquired pneumonia (CAP) is an increasing cause of hospitalisation and mortality, especially among elderly patients [1-5]. Early diagnosis and accurate treatment at the emergency department are essential to avoid serious complications such as bacteremia, sepsis, organ failure, and death [6] and to fight antimicrobial resistance [7].

The diagnosis of CAP generally requires a new infiltrate on a chest x-ray with a clinically compatible syndrome (e.g. fever, dyspnea, cough and sputum production) [8]. These symptoms are not sufficient to diagnose or exclude CAP, as they overlap with other diseases [8] and can be subtle in patients with advanced age and/or impaired immune systems [9, 10]. The chest X-ray is an imprecise diagnostic tool for CAP, risking under/overdiagnosis [11, 12] and might not be the optimal reference standard for CAP. This variability in clinical signs and symptoms combined with non-specific diagnostic tools [12], biomarkers [13, 14], and time-consuming microbiological tests [9] challenges physicians in differentiating CAP from other infections [10, 15].

The CAP population today has also changed with increased ageing [16], multimorbidities [17], and immunomodulatory treatments. Therefore, our knowledge of CAP symptoms and signs needs to be adapted to the actual population.

Previously, prediction models for the diagnosis of CAP were developed on prognostic factors, including severity assessment [18, 19], observations in a primary care setting only [20-22], or a reference diagnosis based solely on the registered discharge diagnosis in the medical record or positive chest X-ray findings [22, 23]. A valid outcome diagnosis was essential. However, in pragmatic studies, an expert panel using available information has been deemed a better reference standard [11].

Therefore, there is a need to describe the clinical characteristics of the current population of patients admitted with suspected CAP and develop a diagnostic model that includes physical examination, blood tests, vital signs, patient medical history, and healthcare expertise. Given the current diagnostic tool inaccuracies, an expert-panel-based diagnostic model was expected to surpass the ED physicians' initial accuracy.

Hypothesis and objectives

We hypothesised that a diagnostic prediction model based on well-defined clinical characteristics could assist an ED physician to make an earlier, more precise CAP diagnosis. Therefore, the aim was to identify the clinical characteristics of adults admitted with CAP and evaluate the performance of these characteristics in a prediction model.

The objectives were:

1) To compare clinical characteristics of patients with a CAP diagnosis from i) all patients admitted with suspected infection and ii) patients suspected of CAP
2) To develop and evaluate a diagnostic model to identify patients with CAP among ED patients suspected of infection and to compare the performance of the model to the initial assessment of the ED physician

METHODS

The study was reported following "The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis" (TRIPOD) statement [24] and conducted in agreement with the Declaration of Helsinki-Ethical principle for medical research involving human subjects. The protocol was approved by the Regional Committee on Health Research Ethics for Southern Denmark (S- 20200188), registered by the Danish Data Protection Agency (no. 20/60508), and in ClinicalTrials.gov (NCT04681963).

Study design, source of data, and setting

This study had an analytical, cross-sectional, multicentre design. The data was collected prospectively and originated from the INfectious DisEases in Emergency Departments (INDEED) study. The published study protocol provides further detailed information [25]. Four Danish medical EDs participated, with a catchment area of around 775,000 inhabitants, between March 1, 2021 and February 28, 2022.

In Denmark, patients can be directed to various specialties within the ED, e.g. medical, gastrointestinal surgery, cardiology, orthopaedics, gynaecology, psychiatry, and intensive care [26]. Suspected infection cases are usually assigned to a medical ED.

Participants

Adult patients (≥ 18 years) admitted to a medical ED were eligible to participate. Patients were included if the ED physician suspected infection and the patient could provide verbal and written consent. The exclusion criteria were: i) need for urgent, life-saving treatment, ii) transferal to intensive care, iii) admission within the last fortnight, iv) verified SARS-CoV-2 infection at the time of admission or within 14 days before admission, v) severe immunodeficiencies (HIV positive, with a cluster of differentiation 4 cell count <200) or treatment with immunosuppressive medicine (Anatomical Therapeutic Chemical classification L04A), corticosteroids (>20 mg/day prednisone or equivalent for >14 days within the last 30 days) or chemotherapy within 30 days.

Recruitment and data collection

Six project assistants with healthcare backgrounds (three physicians, one physiotherapist, and two final-year medical students) were responsible for inclusion and data collection from Mondays to Fridays, 8 am to 8 pm. The population was a convenient sample of eligible patients consecutively identified from the patient management system by a project assistant. Immediately following the initial clinical assessment, the project assistant asked the ED physician whether an infection was suspected and the most likely infection focus (CAP, urinary tract infection, or unknown origin). Generally, the clinical assessment took place within the first 30 minutes of admission before blood tests or imaging were ordered, and therefore, the ED physician often relied only on information from the patient's signs, symptoms, and vital parameters. The study assistant collected verbal and written consent from eligible patients. All data collected was registered in the electronic study database REDCap (Research Electronic Data Capture) [27].

Reference diagnosis

The reference diagnosis was the diagnosis of CAP assessed by an expert panel. The expert panel consisted of eight clinical experts at consultant level in the fields of infectious diseases and emergency medicine working in pairs. They conducted a patient file audit and determined the final diagnosis based on all clinical information registered within the first week of ED admission. The information included routine laboratory tests of blood, -urine, and -sputum. In addition, polymerase chain reaction tests of sputum, urine flow cytometry, chest X-ray, and chest computed tomography (CT) were available for some patients. The experts had access to all images, including the radiologist’s interpretation and documentation. The experts were blinded to each other and independently registered their assessments in a standardized electronic template [27] in the study database. In case of disagreement, the two specialists re-evaluated the medical record and collectively reached a consensus.

Predictors

All clinical characteristics were collected upon arrival at the ED. Symptoms, demographic data, and lifestyle factors were registered during a standardised bedside interview with the patient. In addition, information about vital parameters, comorbidities, medical treatment, and blood tests were collected from the patient's medical record. The project assistants collecting data were blinded to the final diagnosis.

Seventy candidate predictors were selected from the literature and discussed with the specialists and project group [20, 28-37]. The pre-specified potential predictors with measurement units, groups, cut-offs, and considerations/assumptions of inclusion were selected (see Supplementary Table S1).

- Demographic information, lifestyle factors, and comorbidities: age, sex, civil status, employment, nursing home residence, smoking, and alcohol consumption, body mass index (BMI), level of physical activity, activities of daily living score, dementia, respiratory, neurological, cardiovascular, endocrinological, nephrological and gastrointestinal comorbidities were collected.

-Patient symptoms two weeks before admission: malaise, fatigue, headache, dizziness, altered mental status, e.g. confusion, dyspnea, malnutrition, cough, secretions from the respiratory tract, sore throat, common cold, fever feeling, chest pain, peripheral oedema, nausea, vomiting, decreased appetite, abdominal pain, diarrhoea, and pain in muscles and joints including back pain were collected.

-Severity assessment, clinical parameters with cut-offs based on National Early Warning Score (NEWS) [38] used at the arrival of the ED and the use of medications: CURB-65 ≥3 (confusion, uremia, respiratory rate, blood pressure, age > 65 years), triage [39], Glasgow coma scale (GCS), oxygen saturation <96%, heart rate <51 or >90/min, blood pressure (systolic <111 or >219, diastolic ≤60 mmHg), respiratory rate >20/min, temperature > 38^{o}C, abnormal chest auscultation, abdominal tenderness, polypharmacy (≥ 5 medications), use of analgesics, and vaccination status (SARS-CoV-2, pneumococcus, influenza) were recorded.

-Blood tests with cut-offs routinely applied at our institutions: haematocrit (%), haemoglobin (mmol/L), leukocytes (10E9/L), platelets (10E9/L), neutrophils (10E9/L), lymphocytes (10E9/L), albumin g/L, creatinine (µmol/L), blood urea nitrogen (mmol/L), sodium (mmol/L), prothrombin, bilirubin (µmol), glucose (mmol/L), and CRP (mg/L) were recorded.

Statistical methods

The study sample size was estimated using data from the University Hospital of Southern Denmark. We estimated a need for at least 700 patients admitted with suspected infection. Of those, four hundred patients should have suspected CAP and two hundred patients should have verified CAP to complete a reasonable multivariable regression analysis. Descriptive statistics for baseline characteristics of the patients were conducted for the 70 potential predictors based on the data from the INDEED study [25]. Data were presented as means and standard deviations (SD), or medians and interquartile ranges (IQRs) for continuous variables, and numbers (n) and percentages (%) for categorical and binary variables. Extensive univariate logistic regression analyses were performed to examine the unadjusted association between each candidate predictor and the outcome CAP. Results of univariate analyses were reported with odds ratio (OR), 95% confidence intervals (CI), and statistical significance levels were two-sided reported with a p-value of <0.05 to present a descriptive overview of the individual's associations in the population. Complete case analyses were performed, and the predictors were dichotomised or categorised and presented with percentages (%) for inclusion in the final model. The least absolute shrinkage and selection operator (LASSO) multivariable regression was performed with a random split-sample to develop and validate the model, using 20 % of the data for internal cross-validation. The model calibration was assessed using a likelihood ratio test, and recalibration was done based on the calibration belt and the optimal predicted proportion. In the model, age (≥75 years old) was considered an effect modifier based on several studies showing differences in symptoms and signs of a CAP diagnosis in older adults [33, 40-42]. An exploratory approach was conducted for the clinical characteristics to achieve the model with the best predictive performance, testing performances with continuous, dichotomous, or categorical variables. In

addition, the receiver-operator characteristic (ROC) curve was created to estimate the model's accuracy, and the area under the ROC curve (AUC) visualized any discrimination between true positives and negatives. The sensitivity, specificity, and positive and negative predictive values with 95% CI were calculated using the best threshold criteria of the predicted probability of the ROC curve. The same threshold was implemented in developing a CAP score, including the predictor variables. A CAP score> 0 represents the presence of CAP, and < 0 indicates the absence of CAP. Sensitivity, specificity, and positive and negative predictive values with 95% CI were calculated from the initial diagnosis made by the ED physician. Analyses were performed using STATA 17.0 (Texas, USA).

Patient and public involvement

Patients and/or the public were not directly involved in this study.

RESULTS

Participants

We recruited 954 patients admitted to the ED with suspected infection, representing 43% of the population screened for eligibility. Of those, the attending physician suspected that 402 (42%) had a CAP diagnosis. The expert panel verified a CAP diagnosis in 265 (28%) of the recruited patients (Figure 1). The evaluation of 332 chest CT scans showed that 188 (57%) patients had verified pneumonia, and from those, 148 (76%) had CAP assessed by the expert panel and confirmed by a chest CT scan. Most patients (65%) with CAP were discharged to an internal medicine ward, whilst 29% of the patients diagnosed with CAP by the expert panel were discharged directly home. There were 2.5% , 2.5% and 1.0% of the population with CAP that were discharged to the ICU, surgical, other wards respectively.

Characteristics of patients with suspected infections

We compared the clinical characteristics of patients with verified CAP to patients with suspected infection (954). Median age for patients with verified CAP was 75 years (IQR: 63.5; 82.0), and over half admitted with suspected infection were males (53.8%). Univariate analysis revealed that verified CAP patients were more often previous smokers [OR 1.83 (CI: 1.30-2.57) p<0.001] with smoking history compared to suspected infection cases. Strongly independent predictors for CAP were symptoms such as dyspnea, cough, expectoration, chest pain, and cold symptoms (all p<0.001). Compared to patients without CAP, the risk of having CAP increased fivefold if the patient had chest auscultation abnormalities [OR 5.67 (CI: 4.15-7.75) p<0.001] and decreased by half in case of abdominal tenderness by palpation [OR 0.52 (CI: 0.35-0.78) p=0.002]. CAP patients often had comorbidities related to other pulmonary diseases (p<0.001) and often had had previous CAP infections (p<0.001). These patients were more acutely ill when assessed by triage (p<0.001), with fever > 38^{o}C (p=0.036), higher respiratory rate [median 20.0 (IQR 18.0; 24.0) p<0.001], higher heart rate [mean 93.2 (SD 18.9) (p<0.001], and lower oxygen saturation [median 95.0 (IQR: 93.0; 97.0) p<0.001]. Patients with verified CAP had a median CRP of 125.0 (IQR: 57.0; 203.5) versus 82.0 (IQR: 19.0; 172.0) (p<0.001) compared to the rest of the population and higher levels of neutrophils (p<0.001) and leukocytes (p<0.001). Furthermore, lymphocytes yielded a p-value of 0.018. Patients with verified CAP were more often vaccinated against SARS-CoV-2 (p=0.033) and influenza (p=0.025), but no differences were found regarding pneumococcal vaccination. Table 1 presents the characteristics of the population with statistically significant results of the unadjusted association between each predictor for patients with verified and not verified CAP. See Supplementary Table S2 for the 70 exploratory results from continuous, dichotomous, and categorical variables tested in the diagnostic prediction model.

Table 1: Characteristics of the population with suspected infection (n=954).

Characteristics	Patients suspected of infection at admission		Missings n (%)	OR (95% CI)	p-value
	CAP n (%)	Not CAP n (%)			
Total of patients	265 (27.8)	689 (72.2)	0 (0.0)	-	-
LIFESTYLE FACTORS					

Smoking status			33 (3.5)		
No	66 (26.0)	257 (38.5)		1 (reference)	
Current smoker	54 (21.3)	125 (18.7)		1.68 (1.10-2.55)	0.015
Previous smoker	134 (52.8)	285 (42.7)		1.83 (1.30-2.57)	<0.001
SYMPTOMS					
Malaise	173 (67.8)	386 (58.7)	41 (4.3)	1.48 (1.09-2.01)	0.010
Dyspnea	171 (67.3)	208 (31.5)	39 (4.1)	4.48 (3.29-6.11)	<0.001
Cough	173 (68.1)	185 (28.0)	39 (4.1)	5.49 (4.01-7.52)	<0.001
Expectoration	140 (55.1)	139 (21.0)	39 (4.1)	4.61 (3.38-6.28)	<0.001
Sore throat	39 (15.4)	65 (9.8)	39 (4.1)	1.66 (1.08-2.54)	0.019
Common cold	45 (17.7)	50 (7.6)	39 (4.1)	2.63 (1.70-4.05)	<0.001
Chest pain	71 (28.1)	97 (14.7)	40 (4.2)	2.26 (1.60-3.21)	<0.001
Oedema	10 (4.0)	69 (10.4)	40 (4.2)	0.35 (1,17-0.69)	0.002
Vomiting	40 (15.8)	150 (22.6)	38 (4.0)	0.64 (0.43-0.94)	0.023
Gastrointestinal pain	40 (15.8)	153 (23.1)	38 (4.0)	0.62 (0.42-0.91)	0.016
Muscular pain	79 (31.3)	265 (40.3)	44 (4.6)	0.67 (0.49-0.92)	0.013
COMORBIDITIES					
Pulmonary diseases	105 (39.6)	164 (23.8)	0 (0.0)	2.10 (1.55-2.84)	<0.001
Prior pneumonia			100 (10.5)		
No	79 (33.3)	331 (53.6)		1 (reference)	
Yes, one time	50 (21.1)	130 (21.1)		1.61 (1.07-2.42)	0.022
Yes, more than one time	108 (45.6)	156 (25.3)		2.90 (2.05-4.10)	<0.001
VACCINATIONS					
SARS-CoV-2 †	222 (83.8)	534 (77.5)	0 (0.0)	1.49 (1.03-2.17)	0.033
Influenza	191 (72.1)	444 (64.4)	0 (0.0)	1.42 (1.04-1.94)	0.025
CLINICAL ASSESSMENT					
Abnormal chest auscultation*	168 (65.4)	161 (25.0)	52 (5.4)	5.67 (4.15-7.75)	<0.001
Abdominal tenderness	37 (15.0)	155 (25.0)	86 (9.0)	0.52 (0.35-0.78)	0.002
SEVERITY ASSESSMENT					
Triage**			59 (6.2)		
Green/Blue	37 (14.8)	146 (22.6)		1 (reference)	
Yellow	126 (50.4)	353 (54.7)		1.40 (0.93-2.13)	0.105
Red/Orange	87 (34.8)	146 (22.6)		2.35 (1.50-3.67)	<0.001
VITAL PARAMETERS					
Respiratory rate >20/min	124 (47.0)	161 (23.5)	5 (0.5)	2.88 (2.13-3.88)	<0.001
Oxygen saturation < 96 %	162 (61.1)	231 (33.7)	4 (0.4)	3.09 (2.30-4.14)	<0.001
Heart rate <51 or >90/min	148 (55.8)	312 (45.3)	1 (0.1)	1.52 (1.14-2.02)	0.003
Fever > 38°C	77 (29.3)	156 (22.7)	5 (0.5)	1.40 (1.02-1.93)	0.036
BLOOD TESTS					
Leukocytes <3.5 or > 8.8 10E9/L	214 (80.8)	456 (66.2)	0 (0.0)	2.14 (1.52-3.02)	<0.001
Neutrophils > 7.5 10E9/L	187 (71.1)	362 (53.2)	10 (1.0)	2.16 (1.59-2.94)	<0.001
Lymphocytes† <1.00 or > 4.00 10E9/L	53 (55.2)	92 (40.9)	633 (66.3)	1.78 (1.10-2.88)	0.018
C-Reactive protein mg/L			0 (0.0)		
<20 mg/L	21 (7.9)	175 (25.4)		1 (reference)	

21-99 mg/L	86 (32.5)	205 (29.8)		3.49 (2.08-5.86)	<0.001
≥ 100 mg/L	158 (59.6)	309 (44.8)		4.26 (2.60-6.96)	<0.001

The predictors in the table are those dichotomised or categorised as they were later incorporated into the final diagnostic model. Only statistically significant results of the unadjusted association between each candidate predictor and the outcome CAP are presented. *Abnormal chest auscultation: Any abnormal findings such as crackles and rhonchi. ** Triage: Danish emergency process triage [39]. † Variables not included in the multivariate model.

Characteristics of patients suspected of CAP

Using the 70 candidate predictors, we compared the clinical characteristics of patients with verified CAP to patients with suspected but not verified CAP (402).

Statistically significant differences are shown in Table 2. Of the 402 patients with suspected CAP, half of the patients, 229 (57%) had verified CAP. Patients with suspected CAP had a median age of 74.0 (IQR: 62.0; 82.0), and half were male (52.7%). Patients with verified CAP reported more respiratory symptoms, such as cough (p=0.009) and expectoration (p=0.037), and more gastrointestinal symptoms, such as nausea (p=0.033) and loss of appetite (p=0.030), compared to those without CAP. Fewer patients with verified CAP had a CURB-65 ≥3 (p=0.047), and more patients had oxygen saturation <96% (p<0.001), a heart rate of <51 or >100bpm/min (p=0.045), and fever >38 ^{o}C (p=0.011). Elevated infection biomarkers (leukocytes, neutrophils, CRP, all p<0.001), and plasma natrium (p<0.001) were highly associated with CAP. Fewer patients with CAP had plasma bilirubin values of <5 or >25 mmol/L (p=0.045) (Table 2).

Table 2: Characteristics of the population with suspected CAP (n=402) by the physician at admission.

Characteristics	Patients suspected of CAP at admission		Missings n (%)	OR (95% CI)	p-value
	CAP n (%)	Not CAP n (%)			
Total of patients	229 (57.0)	173 (43.0)	0 (0.0)		
SYMPTOMS					
Cough	168 (75.7)	104 (63.4)	16(4.0)	1.79 (1.15-2.79)	0.009
Expectoration	132 (59.5)	80 (48.8)	16 (4.0)	1.54 (1.02-2.31)	0.037
Nausea	70 (31.8)	36 (22.0)	18 (4.5)	1.65 (1.04-2.64)	0.033
Loss of appetite	137 (62.3)	84 (51.2)	18 (4.5)	1.57 (1.04-2.36)	0.030
SEVERITY ASSESSMENT					
CURB65 ≥3 *	23 (10.4)	30 (17.3)	8 (2.0)	0.55 (0.30-0.99)	0.047
VITAL PARAMETERS					
Oxygen saturation <96%	147 (64.2)	79 (46.0)	1 (0.2)	2.11 (1.40-3.15)	<0.001

Heart rate < 51 or >100 bpm/min	129 (56.3)	80 (46.2)	0 (0.0)	1.49 (1.00-2.23)	0.045
Fever >38°C	64 (28.2)	30 (17.3)	2 (0.5)	1.87 (1.14-3.05)	0.011
BLOOD TESTS					
Leukocytes <3.5 or > 8.8 10E9/L	191 (83.4)	106 (61.3)	0 (0.0)	3.17 (1.99-5.04)	<0.001
Neutrophils > 7.5 10E9/L	166 (73.1)	81 (47.6)	5 (1.2)	2.99 (1.96-4.55)	<0.001
Natrium <137 or > 145 mmol/L	114 (49.8)	55 (31.8)	0 (0.0)	2.12 (1.40-3.21)	<0.001
Bilirubin<5 or >25 mmol/L	32 (14.0)	37 (21.8)	4 (1.0)	0.58 (0.34-0.98)	0.045
C-Reactive Protein mg/L, n (%)			0 (0.0)		
<20 mg/L	15 (6.6)	59 (34.1)		1 (reference)	
21-99 mg/L	74 (32.3)	64 (37.0)		4.54 (2.35-8.78)	<0.001
≥ 100 mg/L	140 (61.1)	50 (28.9)		11.01 (5.73-21.14)	<0.001

Statistically significant results from the unadjusted association between each candidate predictor and the outcome CAP.* CURB65: confusion, uremia, respiratory rate, blood pressure, age > 65 years.

Model development and performance

We developed a prediction model for diagnosing pneumonia in patients admitted with suspected infection (n=954) and compared it with the clinician's presumptive diagnosis. Supplementary Table S3 presents the characteristics of the population randomised in the training and validation sets.

The predictors associated with CAP in our final model are presented in Table 3.

Table 3: The complete diagnostic model, including the intercept

Intercept and predictors	ß Coefficient
Intercept	-1.66192
Dyspnea (yes)	0.35172
Expectoration (yes)	0.36250
Cough (yes)	0.39671
Common cold (yes)	0.34374
Malaise (yes)	0.07475
Chest pain (yes)	0.20499
Respiratory rate >20/min	0.14566
Oxygen saturation < 96%	0.24303
Abnormal auscultation findings (yes)	0.56758
Leukocytes*	0.00322
Neutrophils**	0.08338
C-reactive protein <20 mg/L	-0.64269
Previous event of CAP (no)	-0.12006
Age of ≥ 75 and cough (yes)	0.53816
Age of ≥ 75 and oedema (no)	-0.05797
Age of ≥ 75 and glucose >11.0 mmol/L	0.88124
ROC AUC† (95% CI)	0.85 [0.77-0.92]

* Cut-off for leucocyttes: normal values 3.5 -8.8 10E9/L **Neutrophils: > 7.5 10E9/L
† ROC AUC = receiver-operating characteristic area under the curve

The model performance yielded an AUC of 0.85 [CI: 0.77-0.92], and the calibration of the model yielded p=0.227 after recalibration, demonstrating a good prediction of the proportion of CAP patients in the test sample (Supplementary figure S1 and Supplementary figure S2).

Based on a lambda result of λ=0.0402856 and a probability threshold of 0.35, the LASSO calculation with characteristics predictive of CAP and the calculation of the final model with a cut-off value greater than 0 indicating the diagnosis CAP are presented in supplemental material (Supplementary formula S1 and Supplementary formula S2).

At the optimal cut-off of 0.35, the prediction model yielded an 86.1% sensitivity and 64.1% specificity. Based on the trial population (Figure 1), the sensitivity of the prediction model was comparable to the initial diagnosis made by the ED physicians. However, the specificity and positive predictive value were significantly lower (Table 4).

Table 4: Performance of the predictive model compared to the initial diagnosis made by the ED physicians.

Performance	Sensitivity % [CI %]	Specificity % [CI %]	Positive predictive value % [CI %]	Negative predictive value % [CI %]
Predictive model	86.1 [79.1-93.1]	64.1 [57.1-71.1]	41.6 [34.6-48.6]	93.9 [86.9-100]
Physicians	86.4 [84.2-88.6]	74.9 [72.1-77.6]	57.0 [53.8-60.1]	93.5 [92.0-95.0]

The predictive model had a 35% cut-off and a prevalence of 22%. The prevalence of CAP was 28% in the population of 954 patients suspected of infection.

Model specification

The final model did not include the following possible predictors: lymphocytes, SARS-CoV-2, and BMI. The reasons were a high percentage of missings (lymphocytes 66.3%), clinical relevance, and statistical performance (BMI and SARS-CoV-2). These considerations are described in detail in Supplemental material.

DISCUSSION

More than every fourth patient with suspected infection was diagnosed with CAP (28%). The ED physicians suspected CAP in almost half (42%) of patients admitted with suspected infection. Patients with suspected CAP included 57% with a final expert diagnosis of CAP and 43% without CAP. We have identified twenty-seven clinical characteristics for patients diagnosed with CAP among those admitted suspected of infection. Patients with CAP were characterised more often with a history of smoking, previous CAP, respiratory symptoms, abnormal lung auscultation, worse triage, and abnormal levels of infection biomarkers. Fewer clinic characteristics (thirteen) were identified for patients diagnosed with CAP among patients suspected of CAP by the ED physician and included typical respiratory symptoms but also gastrointestinal symptoms, abnormal vital signs, increased blood markers, and lower CURB-65 scores. The final diagnostic prediction model yielded thirteen diagnostic predictors for CAP recognised by the literature. The model performance was similar to the diagnosis made by the ED physicians regarding sensitivity and negative predictive value but not as good in determining the specificity and positive predictive values.

Our prediction model had a good performance (AUC 0.85) and calibration (p=0.227), and with the best cut-off at 35%, the sensitivity reached 86.1% and specificity 64.1%. Therefore, the model could be tested externally at other sites, especially where clinicians are not always available due to the lack of resources, and contribute to the initial management of CAP, guiding further clinical investigation. In this study, ED physicians relied upon the patient's history and the results from a simple clinical examination to diagnose CAP with a comparable negative predictive value (93% vs. 94%) and a better positive predictive value (57% vs. 42%). Even though our model is not entirely comparable to the initial diagnosis made by the ED physicians due to the difference in the prevalence of CAP, our results are similar to a recent systematic review [43]. Other studies reported that ED physicians' accuracy in diagnosing CAP ranged from 76% to 96% [44], and artificial intelligence predicted the presence of pneumonia with a sensitivity of 94% and specificity of 50% [45]. These results show that there is room for improvement in diagnosing CAP. It could be achieved

by including additional predictors such as biomarkers, e.g., procalcitonin, YKL-40, and surfactant protein-D [46, 47], molecular detection of respiratory pathogens [48], and/or improved imaging modalities [12, 14].

This prospective study highlights the challenges in identifying patients with CAP based on patient history, vital signs, and symptoms upon admission [20, 22, 46]. An initial CAP diagnosis may often differ from the discharge diagnosis [10, 49]. A plausible cause for uncertainty in diagnosing CAP was the heterogenic presentation of symptoms overlapping with other diseases. We found that patients with verified CAP often had gastrointestinal symptoms, whereas patients not verified with CAP sometimes presented with typical respiratory symptoms and had more severe conditions measured by CURB-65. Typical respiratory symptoms could explain some CAP misclassification. Misclassification of CAP may lead to unnecessary or ineffective antibiotic treatment, increased healthcare costs, delayed diagnosis, increased mortality, and increased risk of bacterial resistance [44, 50].

The predictors of CAP identified in this study are strongly discussed in the literature [9, 20, 36, 37, 42, 46, 49]. Most prediction models for ED patients with CAP aim to predict prognostic outcomes such as disease severity and mortality [51]. Prior studies have either included few diagnostic predictors or very selected patients [20, 22, 52]. The main reason for including several potential predictors and having age as a cross-factor in the development of our model was the expectation of finding predictors not represented in the literature and predictors specific for older patients (≥75 years). This is considered very relevant as the population worldwide ages [4, 16]. An age of ≥ 75 interacted with the symptoms of cough, blood glucose levels, and peripheral oedema. Peripheral oedema was associated with an absence of CAP, and symptoms may be explained by other infections, such as erysipelas or heart failure. In addition, hyperglycemia has been recognized as a predictor associated with poorer patient outcomes for elderly CAP patients, regardless of their history of diabetes [53, 54].

Even though the literature highlights malnutrition as a strong prognostic predictor for CAP [33, 35, 55], we excluded BMI from our final model. Measuring weight and height is not a priority in acute settings where vital parameters, symptoms, and point-of-care biomarkers are the primary observations in the diagnostic

process. Another concern was that BMI was missing in 26.3% of the population, and bias may arise due to systematic differences between subjects with complete datasets and subjects with missing data. Patients with missing BMI data may be more frail, incapable, or difficult to transfer. A model including BMI could be a better choice in a primary care setting, where patients are not necessarily as acutely ill and may be able to weigh themselves.

A major strength of this study is the completeness of data from medical charts and patient interviews combined with CAP diagnoses assigned by a panel of experts. The experts had a range of information from the patient's medical records, including chest X-ray, chest CT for patients suspected of CAP, and microbiology results. In addition to identifying possible predictors, we included many relevant and easily accessible clinical parameters. Finally, we excluded patients infected with SARS-CoV-2 from the study to increase the potential generalisability for CAP patients after the pandemic.

This study has limitations. Multiple testing and mass significance are potentially a problem in this study. Methods, such as Bonferroni-Holm correction, could have been applied to counteract this problem [56]. However, the univariate analyses were conducted for exploratory and descriptive purposes only. Therefore, these results should be interpreted cautiously, and the findings should be used as hypothesis-generating rather than conclusive. Another concern is that even though the reference standard of CAP was the same for the model performance and the initial diagnosis of the ED physicians, the expert panel may have had better opportunities to diagnose CAP in suspected CAP patients due to the availability of results from imaging and microbiological tests, and better register of patient’s symptoms. This could lead to differential verification bias overestimating the ED physician's accuracy in diagnosing CAP [57]. This assumption was supported by the higher specificity of CAP diagnoses from ED physicians.

Another limitation is the selected population of patients allocated to the internal medicine speciality that may have masked atypical predictors from patients assigned to other specialities. Furthermore, some patients with atypical clinical presentation may have an infection that the ED physician had not suspected upon admission and, therefore, was not included in our study. Patients with severe conditions or acute

cognitive impairment who could not consent were excluded. Furthermore, the inclusion of patients took place during work hours and weekdays, which may have reduced the number of severe cases as admission during out-of-hours and weekends are associated with increased mortality and ICU admissions [58]. Therefore, our results can only be generalised to patients suspected of CAP and admitted on weekdays during the daytime.

A broader patient inclusion may contribute to a model that identifies other predictors to diagnose CAP as the clinical presentation may differ from those admitted with suspected CAP and capable of consent. Another limitation was the pragmatic choice of cut-offs for blood tests routinely used in our institutions, which reflected our clinical practice. However, it does raise questions about the applicability in other settings that apply different cut-offs.

This population cohort could be applicable as a test validation cohort for future models as the data collection of these well-known predictors of CAP is reproducible across EDs. The development of automatic extraction for a prediction model from electronic medical records using artificial intelligence could be of great value in a busy ED. In conclusion, typical respiratory symptoms combined with abnormal vital signs and elevated infection biomarkers are predictors for CAP upon admission to an ED. A diagnostic prediction model based on these predictors is of limited value. Future prediction models should include novel diagnostic tools, imaging, PCR analysis, and/or serological markers not routinely used in clinical practice to improve model performance and diagnose CAP more accurately in the ED.

Acknowledgements: The authors appreciate text editing from the research consultant Caroline Moos, statistician support from Andreas Kristian Pedersen and Sofie Ronja Petersen at the University Hospital of Southern Denmark, and from OPEN (Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark).

Authors' contributions: MBC, FSR, CBM, TS, HSA, MHL, AH, and MAH were involved in the study's design. MBC performed the literature search and drafted the original work. MBC, MHL, AH, MAH, JJS, and FK

recruited patients and collected data. CBM and MAH participated in the expert panel. HSA was the study investigator-, and coordinated and supervised the project. MBC performed the statistical analyses. CBM was the chief research officer responsible for supervising the overall study. All authors, MBC, FSR, CBM, TS, HSA, MHL, AH, MAH, FK, and JJS critically revised and approved the final manuscript.

Funding: University of Southern Denmark (17/10636), University Hospital of Southern Denmark (20/20505), The funders of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Competing interests: The authors declare that they have no competing interests.

Patient and public involvement: Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication: Not required.

Ethics approval and consent to participate: Ethics approval and consent to participate: Approval was obtained from the Regional Committee for Health Research Ethics in Southern Denmark (S-20200188). In addition, informed verbal and written consent was obtained from each participant before enrolment in the study. This study was conducted in agreement with the Declaration of Helsinki-Ethical principle for medical research involving human subjects.

Availability of data and materials: Due to Danish laws on personal data, data cannot be shared publicly. To request data, please contact the corresponding author for more information. The person responsible for the research was the principal investigator and corresponding author (MBC) in collaboration with the University Hospital of Southern Denmark. This organization owns the data and can provide access to the final data set.

REFERENCES

1. Søgaard M, Nielsen RB, Schønheyder HC, Nørgaard M, Thomsen RW. Nationwide trends in pneumonia hospitalization rates and mortality, Denmark 1997-2011. Respir Med. 2014;108(8):1214-22.doi:10.1016/j.rmed.2014.05.004
2. McLaughlin JM, Khan FL, Thoburn EA, Isturiz RE, Swerdlow DL. Rates of hospitalization for community-acquired pneumonia among US adults: A systematic review. Vaccine. 2020;38(4):741-51.doi:10.1016/j.vaccine.2019.10.101
3. Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. Thorax. 2012;67(1):71-9
4. Laporte L, Hermetet C, Jouan Y, Gaborit C, Rouve E, Shea KM, et al. Ten-year trends in intensive care admissions for respiratory infections in the elderly. Ann Intensive Care. 2018;8(1):84.doi:10.1186/s13613-018-0430-6
5. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis. 2018;18(11):1191-210.doi:10.1016/s1473-3099(18)30310-4
6. Meehan TP, Fine MJ, Krumholz HM, Scinto JD, Galusha DH, Mockalis JT, et al. Quality of care, process, and outcomes in elderly patients with pneumonia. JAMA. 1997;278(23):2080-4.doi:10.1001/jama.1997.03550230056037
7. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infect Dis. 2014;14 (1):13.doi:10.1186/1471-2334-14-13
8. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med. 2019;200(7):e45-e67.doi:10.1164/rccm.201908-1581ST
9. Musher DM, Thorner AR. Community-acquired pneumonia. N Engl J Med. 2014;371(17):1619-28.doi:10.1056/NEJMra1312885
10. Chandra A, Nicks B, Maniago E, Nouh A, Limkakeng A. A multicenter analysis of the ED diagnosis of pneumonia. Am J Emerg Med. 2010;28(8):862-5.doi:10.1016/j.ajem.2009.04.014
11. Claessens YE, Debray MP, Tubach F, Brun AL, Rammaert B, Hausfater P, et al. Early chest computed tomography scan to assist diagnosis and guide treatment decision for suspected community-acquired pneumonia. Am J Respir Crit Care Med. 2015;192(8):974-82.doi:10.1164/rccm.201501-0017OC
12. Ye X, Xiao H, Chen B, Zhang S. Accuracy of Lung Ultrasonography versus Chest Radiography for the Diagnosis of Adult Community-Acquired Pneumonia: Review of the Literature and Meta-Analysis. PLoS One. 2015;10(6):e0130066.doi:10.1371/journal.pone.0130066
13. Kamat IS, Ramachandran V, Eswaran H, Guffey D, Musher DM. Procalcitonin to distinguish viral from bacterial pneumonia: a systematic review and meta-analysis. Clin Infect Dis. 2020;70(3):538-42.doi:10.1093/cid/ciz545
14. Gentilotti E, De Nardo P, Cremonini E, Górska A, Mazzaferri F, Canziani LM, et al. Diagnostic accuracy of point-of-care tests in acute community-acquired lower respiratory tract infections. A systematic review and meta-analysis. Clin Microbiol Infect. 2022;28(1):13-22.doi:10.1016/j.cmi.2021.09.025
15. Welker JA, Huston M, McCue JD. Antibiotic timing and errors in diagnosing pneumonia. Arch Intern Med. 2008;168(4):351-6.doi:10.1001/archinternmed.2007.84

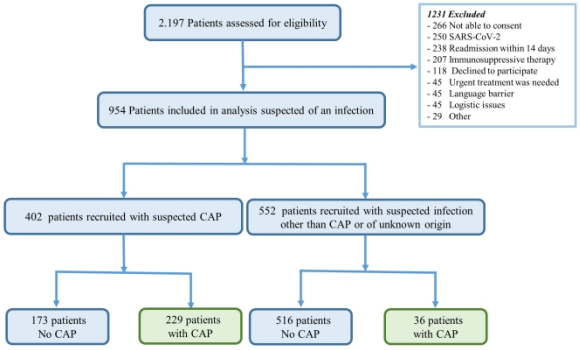
16. World Health Organization. Aging and Health [Internet]. 2022 October 1 [cited 2022 October 28]. Available from: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.
17. Weir DL, Majumdar SR, McAlister FA, Marrie TJ, Eurich DT. The impact of multimorbidity on short-term events in patients with community-acquired pneumonia: prospective cohort study. Clin Microbiol Infect. 2015;21(3):264.e7-.e13.doi:10.1016/j.cmi.2014.11.002
18. Sakakibara T, Shindo Y, Kobayashi D, Sano M, Okumura J, Murakami Y, et al. A prediction rule for severe adverse events in all inpatients with community-acquired pneumonia: a multicenter observational study. BMC Pulm Med. 2022;22(1):34.doi:https://dx.doi.org/10.1186/s12890-022-01819-0
19. Gong L, He D, Huang D, Wu Z, Shi Y, Liang Z. Clinical profile analysis and nomogram for predicting in-hospital mortality among elderly severe community-acquired pneumonia patients with comorbid cardiovascular disease: a retrospective cohort study. BMC Pulm Med. 2022;22(1):312.doi:10.1186/s12890-022-02113-9
20. Ding F, Han L, Yin D, Zhou Y, Ji Y, Zhang P, et al. Development and validation of a simple tool composed of items on dyspnea, respiration rates, and C-reactive protein for pneumonia prediction among acute febrile respiratory illness patients in primary care settings. BMC Med. 2022;20(1):360.doi:10.1186/s12916-022-02552-5
21. Hammond A, Halliday A, Thornton HV, Hay AD. Predisposing factors to acquisition of acute respiratory tract infections in the community: a systematic review and meta-analysis. BMC Infect Dis. 2021;21(1):1254.doi:10.1186/s12879-021-06954-3
22. Ebell MH, Chupp H, Cai X, Bentivegna M, Kearney M. Accuracy of signs and symptoms for the diagnosis of community-acquired pneumonia: a meta-analysis. Acad Emerg Med. 2020;27(7):541-53.doi:10.1111/acem.13965
23. Kitazawa T, Yoshihara H, Seo K, Yoshino Y, Ota Y. Characteristics of pneumonia with negative chest radiography in cases confirmed by computed tomography. J Community Hosp Intern Med Perspect. 2020;10(1):19-24.doi:10.1080/20009666.2020.1711639
24. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-W73.doi:10.7326/m14-0698
25. Skjøt-Arkil H, Heltborg A, Lorentzen MH, Cartuliares MB, Hertz MA, Graumann O, et al. Improved diagnostics of infectious diseases in emergency departments: a protocol of a multifaceted multicentre diagnostic study. BMJ Open. 2021;11(9):e049606.doi:10.1136/bmjopen-2021-049606
26. Nørgaard B, Mogensen CB, Teglbjærg LS, Brabrand M, Lassen AT. Diagnostic packages can be assigned accurately in emergency departments. A multi-centre cohort study. Dan Med J. 2016;63(6)
27. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208.doi:10.1016/j.jbi.2019.103208
28. Mutepe ND, Cockeran R, Steel HC, Theron AJ, Mitchell TJ, Feldman C, et al. Effects of cigarette smoke condensate on pneumococcal biofilm formation and pneumolysin. Eur Respir J. 2013;41(2):392-5.doi:10.1183/09031936.00213211
29. Samokhvalov AV, Irving HM, Rehm J. Alcohol consumption as a risk factor for pneumonia: a systematic review and meta-analysis. Epidemiol Infect. 2010;138(12):1789-95.doi:10.1017/s0950268810000774
30. Torres A, Peetermans WE, Viegi G, Blasi F. Risk factors for community-acquired pneumonia in adults in Europe: a literature review. Thorax. 2013;68(11):1057-65.doi:10.1136/thoraxjnl-2013-204282
31. Barbagelata E, Cillóniz C, Dominedò C, Torres A, Nicolini A, Solidoro P. Gender differences in community-acquired pneumonia. Minerva Med. 2020;111(2):153-65.doi:10.23736/s0026-4806.20.06448-4
32. Baik I, Curhan GC, Rimm EB, Bendich A, Willett WC, Fawzi WW. A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women. Arch Intern Med. 2000;160(20):3082-8.doi:10.1001/archinte.160.20.3082

33. Cillóniz C, Dominedò C, Pericàs JM, Rodriguez-Hurtado D, Torres A. Community-acquired pneumonia in critically ill very old patients: a growing problem. Eur Respir Rev. 2020;29(155):190126.doi:10.1183/16000617.0126-2019
34. Reisinger EC, Fritzsche C, Krause R, Krejs GJ. Diarrhea caused by primarily non-gastrointestinal infections. Nat Clin Pract Gastroenterol Hepatol. 2005;2(5):216-22
35. Riquelme OR, Riquelme OM, Rioseco ZML, Gómez MV, Cárdenas G, Torres C. Neumonía adquirida en la comunidad en el anciano hospitalizado: Aspectos clínicos y nutricionales. [Community-acquired pneumonia in the elderly: clinical and nutritional aspects]. Rev Med Chil. 2008;136(5):587-93.doi:10.4067/S0034-98872008000500006
36. Moore M, Stuart B, Little P, Smith S, Thompson MJ, Knox K, et al. Predictors of pneumonia in lower respiratory tract infections: 3C prospective cough complication cohort study. Eur Respir J. 2017;50(5).doi:10.1183/13993003.00434-2017
37. van Vugt SF, Broekhuizen BD, Lammens C, Zuithoff NP, de Jong PA, Coenen S, et al. Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. BMJ. 2013;346:f2450.doi:10.1136/bmj.f2450
38. Alam N, Vegting IL, Houben E, van Berkel B, Vaughan L, Kramer MH, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation. 2015;90:111-5.doi:10.1016/j.resuscitation.2015.02.011
39. Plesner LL, Iversen AKS, Langkjær S, Nielsen TL, Østervig R, Warming PE, et al. The formation and design of the TRIAGE study-baseline data on 6005 consecutive patients admitted to hospital from the emergency department. Scand J Trauma Resusc Emerg Med. 2015;23:106.doi:10.1186/s13049-015-0184-1
40. Ravioli S, Germann C, Gygli R, Exadaktylos AK, Lindner G. Age- and sex-related differences in community-acquired pneumonia at presentation to the emergency department: a retrospective cohort study. Eur J Emerg Med. 2022;29(5):366-72.doi:10.1097/mej.0000000000000933
41. Akhtar A, Hassali MAA, Zainal H, Ali I, Iqbal MS, Khan AH. Respiratory-tract infections among geriatrics: prevalence and factors associated with the treatment outcomes. Ther Adv Respir Dis. 2021;15:1753466620971141.doi:10.1177/1753466620971141
42. Metlay JP, Schulz R, Li YH, Singer DE, Marrie TJ, Coley CM, et al. Influence of age on symptoms at presentation in patients with community-acquired pneumonia. Arch Intern Med. 1997;157(13):1453-9.doi:doi:10.1001/archinte.1997.00440340089009
43. Dale AP, Marchello C, Ebell MH. Clinical gestalt to diagnose pneumonia, sinusitis, and pharyngitis: a meta-analysis. Br J Gen Pract. 2019;69(684):e444-e53.doi:10.3399/bjgp19X704297
44. Ray P, Birolleau S, Lefort Y, Becquemin MH, Beigelman C, Isnard R, et al. Acute respiratory failure in the elderly: etiology, emergency diagnosis and prognosis. Crit Care. 2006;10(3):R82.doi:10.1186/cc4926
45. Heckerling PS, Gerber BS, Tape TG, Wigton RS. Prediction of community-acquired pneumonia using artificial neural networks. Med Decis Making. 2003;23(2):112-21.doi:10.1177/0272989x03251247
46. Htun TP, Sun Y, Chua HL, Pang J. Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review. Sci Rep. 2019;9(1):7600.doi:10.1038/s41598-019-44145-y
47. Spoorenberg SM, Vestjens SM, Rijkers GT, Meek B, van Moorsel CH, Grutters JC, et al. YKL-40, CCL18 and SP-D predict mortality in patients hospitalized with community-acquired pneumonia. Respirology. 2017;22(3):542-50.doi:10.1111/resp.12924
48. Gastli N, Loubinoux J, Daragon M, Lavigne JP, Saint-Sardos P, Pailhoriès H, et al. Multicentric evaluation of BioFire FilmArray Pneumonia Panel for rapid bacteriological documentation of pneumonia. Clin Microbiol Infect. 2021;27(9):1308-14.doi:10.1016/j.cmi.2020.11.014
49. Metlay JP, Fine MJ. Testing strategies in the initial management of patients with community-acquired pneumonia. Ann Intern Med. 2003;138(2):109-18.doi:10.7326/0003-4819-138-2-200301210-00012
50. Kanwar M, Brar N, Khatib R, Fakih MG. Misdiagnosis of community-acquired pneumonia and inappropriate utilization of antibiotics: side effects of the 4-h antibiotic administration rule. Chest. 2007;131(6):1865-9.doi:10.1378/chest.07-0164

51. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. Thorax. 2010;65(10):884-90.doi:10.1136/thx.2009.134072
52. Okimoto N, Yamato K, Kurihara T, Honda Y, Osaki K, Asaoka N, et al. Clinical predictors for the detection of community-acquired pneumonia in adults as a guide to ordering chest radiographs. Respirology. 2006;11(3):322-4.doi:10.1111/j.1440-1843.2006.00846.x
53. Zeng W, Huang X, Luo W, Chen M. Association of admission blood glucose level and clinical outcomes in elderly community-acquired pneumonia patients with or without diabetes. Clin Respir J. 2022;16(8):562-71.doi:10.1111/crj.13526
54. Barmanray RD, Cheuk N, Fourlanos S, Greenberg PB, Colman PG, Worth LJ. In-hospital hyperglycemia but not diabetes mellitus alone is associated with increased in-hospital mortality in community-acquired pneumonia (CAP): a systematic review and meta-analysis of observational studies prior to COVID-19. BMJ Open Diabetes Res Care. 2022;10(4).doi:10.1136/bmjdrc-2022-002880
55. Yeo HJ, Byun KS, Han J, Kim JH, Lee SE, Yoon SH, et al. Prognostic significance of malnutrition for long-term mortality in community-acquired pneumonia: a propensity score matched analysis. Korean J Intern Med. 2019;34(4):841-9.doi:10.3904/kjim.2018.037
56. Sedgwick P. Multiple significance tests: the Bonferroni correction. BMJ. 2012;344
57. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174(4):469-76.doi:10.1503/cmaj.050090
58. Vest-Hansen B, Riis AH, Sørensen HT, Christiansen CF. Out-of-hours and weekend admissions to Danish medical departments: admission rates and 30-day mortality for 20 common medical conditions. BMJ Open. 2015;5(3):e006731.doi:10.1136/bmjopen-2014-006731

Figure Legend:

Figure 1 - Trial population, green boxes showing the numbers of patients with CAP.



338x190mm (300 x 300 DPI)

Supplemental material

Community-acquired pneumonia – Use of clinical characteristics of acutely admitted patients for the development of a diagnostic model: A cross-sectional multicentre study

Mariana Bichuette Cartuliares MSc* [1, 2], Christian Backer Mogensen [1, 2], Flemming Schønning Rosenvinge [3, 4], Thor Aage Skovsted [5], Morten Hjarnø Lorentzen [1,2], Anne Heltborg Kristensen[1,2], Mathias Amdi Hertz [6,7], Frida Kaldan[1], Jens Juel Specht[1], Helene Skjøt-Arkil [1,2]

Affiliations

[1]Emergency Department, University Hospital of Southern Denmark, Aabenraa, Denmark

[2]Department of Regional Health Research, University of Southern Denmark, Odense, Denmark

[3]Research Unit of Clinical Microbiology, University of Southern Denmark, Odense, Denmark

[4]Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark

[5]Department of Biochemistry and Immunology, University Hospital of Southern Denmark, Aabenraa, Denmark

[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark 7 Infectious Diseases Department, Odense University Hospital, Odense, Denmark

* Corresponding author: mbc@rsyd.dk

Table of Contents

Table S1: Description of the 70 pre-specified predictors for CAP

Source: *The patient interview*			
Group	**Variable name**	**Measurement**	**Consideration/assumption** Considerations to collect data from these predictors were based on the described literature and expert consensus together with the project group
Demographic information	Age	Continuous, years	Age is a risk factor for CAP [1]. Several studies stratify age groups when investigating pneumonia due to several atypical symptoms and signs and the absence of respiratory symptoms among the elderly. Stratified age groups differ in cut-offs between the ages of ≥65 to ≥80 years old [2-7].
	Gender	Binary 1=Male 0=Female	The risk of CAP is higher for males [8].CAP is more severe [7] leading to higher mortality in males [9]. Males' lifestyle factors differ from women resulting in a higher risk of CAP [10].
	Civil status (Living alone)	Binary (Yes/no)	Living alone has a two-fold association with having one or more respiratory tract infections [11].
	Nursing home residence	Binary (Yes/no)	Nursing home residents were found to have several comorbidities [12] and lower physical functioning levels, which might result in a higher risk of CAP [13].
	Employment	Categorical: 1=Working 2=Retired 0=Others (e.g. students, flex job)	Low income and unemployment are associated with readmissions after CAP [14].
Symptoms	Feeling unwell/ Malaise	Binary (Yes/No) Symptoms within 14 days prior to ED admission.	Malaise has been identified as one of the most frequent symptoms for patients infected with *Mycoplasma pneumoniae*[15].
	Fatigue		Fatigue is associated with pneumonia especially in elderly patients [4].
	Headache		Headache is one of the clinical findings of symptoms of CAP [7, 15]. However, headaches were less common in the older population [7].
	Dizziness		The rationale of the presence of dizziness as a symptom relied on the assumption that several factors such as polypharmacy[16], combined with comorbidities such as cardiovascular diseases [17], symptoms such as confusion, conditions of frailty and malnutrition [18], and lower oxygen saturation [19]could contribute to dizziness.
	Confusion		Confusion e.g. altered mental status or delirium was significantly more frequent in CAP patients [2, 4].
	Dyspnea		Dyspnea was identified as a strong prediction of CAP among febrile patients [20] and one of the main symptoms of pneumonia [2, 21].
	Cough		Cough is a common symptom and one of the most frequent increasing the likelihood of detecting a viral pathogen among CAP patients [15, 22]. Algorithms included cough as a diagnostic predictor [23], and dry cough was a strong predictor in a prediction model for *Legionella pneumoniae* [24]. Cough was less common in older population [7].
	Secretions		Purulent secretions were a significant symptom and predictor for CAP patients [20, 21].
	Sore throat		Some studies identified sore throat as a symptom of CAP [15], and one included the symptom in the prediction rules of pneumonia [5].

	Cold		Among respiratory diseases, the common cold is one of the most frequent, with symptoms similar to CAP [25].
	Fever feeling		Quantified from reported chills or night sweat or fever measured at home. Included as a rationale of fever.
	Chest pain		Chest pain has been used as a single predictor of CAP [18, 20, 23] or a combined diagnostic predictor [23] and may present as a secondary symptom of coughing or pleuritic involvement [26]. However, chest pain was less common in the older population [7].
	Peripheral edema		The rationale for including peripheral edema as possible predictor is that it is included in the clinical assessment at admission. In case of peripheral edema and respiratory symptoms of dyspnea, chest pain and a history of cardiovascular disease, CAP could be ruled out as a tentative diagnosis replaced with suspicion of cardiovascular disease.
	Nausea		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Vomiting		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Loss of appetite		Loss of appetite could be present in the case of gastrointestinal symptoms [26] and could result from malnutrition [18].
	Abdominal pain		Abdominal pain may be present in the case of gastrointestinal symptoms described above and, therefore, is included in the model [26].
	Diarrhea		Gastrointestinal symptoms such as nausea, vomiting and diarrhea manifests in 20% of the CAP population [26].
	Pain in muscles and joints including back pain		Muscle and joint pain are associated with viral pneumonia as influenza, especially among younger patients and therefore is included in our model [27].
Previous event of CAP	Previous event of CAP	Categorical: 0= Never 1= Once 2= More than once	A previous diagnosis of CAP was reported as having robust evidence as a risk factor for CAP [1]. Furthermore, any hospitalization in the previous five years was reported as a predisposing factor for CAP [8].
Lifestyle factors and aids	Smoke	Categorical: 0=Never been a smoker 1=Current smoker 2=Previous smoker	Smoking has been associated with an increased risk of CAP in several studies [1, 8, 10, 17], and has a strong association with the treatment outcomes of elderly individuals with respiratory tract infections [28].
	Alcohol	Doses per week (a dose=12 grams (1, 5 cl) alcohol). Categories based on the Danish Board of Health recommendations [29]. 0=No alcohol 1=1-7 doses/week maximum doses recommended for women 2=8-14 doses/week maximum dose recommended for men 3= >14 doses	Alcohol has also been associated with increased CAP risk and with treatment outcomes. The risk increases in individuals with higher consumption (>41 g/day) compared to those who consume no alcohol [10, 17, 28].
	Physical activity levels	We categorized physical activity levels based on recommendations from the world health organization for adults with a minimum 150 min/week [30]. 1= Not physically active 2= Less than 2.5hrs/week 3= More than 2.5hrs/week	The risk of CAP decreased in physically active women [10]. In addition, a high level of activity protects against upper respiratory tract infections and reduces the severity and symptoms of the infection [13].
	Activities of daily living	Binary (yes/no) Yes= If the patient had one or more dependencies regarding: bathing, dressing, toileting, transfer, continence and feeding.	Difficulty in maintaining toilet hygiene, preparing meals, and being unable to transfer were associated with an increased risk of respiratory infections [31].

Source: *Variables extracted from the patient's medical report*			
Comorbidities (diseases)	Neurological	Binary (Yes/no) If the patient was diagnosed with one of these diagnoses.	Cerebrovascular disease/stroke and Parkinson's disease approximately doubled the risk of CAP [17].
	Pulmonary		A history of pneumonia increased the risk of a subsequent episode and patients with chronic respiratory diseases, including chronic obstructive pulmonary disease, bronchitis or asthma, had up to a fourfold increase in the risk of CAP [1, 4, 17].
	Endocrinological		Chronic liver conditions were reported as a risk factor of CAP [8]. Recently, diabetes mellitus has been described as an independent risk factor for sepsis secondary to CAP in very old patients [4] and data from several studies showed an association between diabetes mellitus and moderate risk of CAP [17].
	Renal		Chronic renal disease was reported as an independent risk factor for sepsis secondary to CAP in very old patients [4, 8] and chronic renal disease increased the risk of CAP twofold [17].
	Cardiovascular		Chronic cardiovascular disease increased the risk of CAP up to threefold [4, 17].
	Gastrointestinal		The rationale for including gastrointestinal diseases in the model was that CAP patients have gastrointestinal symptoms that could be related to a differential diagnosis besides CAP.
	Dementia		Dementia approximately doubles the risk of CAP [17].
	Cancer		Cancer was associated with a moderate increase in CAP risk, and a single study reported a fivefold increased risk of CAP for patients with lung cancer [17].
	Rheumatological		A moderate risk of CAP was found in patients with rheumatological diseases [17].
Pharmacological treatments	Polypharmacy	Binary (yes/no) Regular consumption of at least five medications	The increased number of comorbidities of older patients increases the risk of polypharmacy [4, 32]. The prevalence of polypharmacy reached almost 40% among individuals with respiratory tract infections above age 65 years and had a twofold association with treatment outcomes of respiratory tract infections [28]. Furthermore, the prevalence of polypharmacy increased from 45% to 74%, irrespective of antibiotic use if patients were hospitalized with CAP [16].
	Analgesics	Binary (Yes/no) Regular consumption of analgesics	A systematic review reported an association between prescribed opioids and CAP [33].
	Vaccination SARS-CoV-2	Binary (Yes/no) Recent vaccination for SARS-CoV-2	SARS-CoV-2 vaccination was reported during the clinical assessment but was taken out of the model, as the model would be used after the pandemic when vaccination for SARS-CoV-2 rates might decrease. However, the inclusion of this variable did not change the final predictive model.
	Vaccination pneumococcus	Binary (Yes/no) Pneumococcus vaccine (not specified) within 5 years	*Streptococcus pneumoniae* is one of the most causative pathogens of CAP and the vaccine could be a possible protective predictor for CAP as the risk of CAP increases among those unvaccinated [1, 34, 35].
	Vaccination influenza	Binary (Yes/no) Season influenza vaccine 2020/2021	Influenza vaccine can reduce hospitalization but is questionable if it could have a protective effect in admitted patients [1, 36], therefore, we included this possible predictor to investigate if it could have a protective role in our population.
Severity assessment	CURB-65	Binary ≥ 3 points (Yes/no) Definition: Confusion, urea >7 mmol/L, respiratory rate ≥ 30 bpm, blood pressure (≤90 for systolic blood pressure or ≤60 for diastolic blood pressure, age > 65 years) Score: one point for each present variable. CURB65≥ 3= severe condition	CURB65 is an assessment tool for the severity of CAP [37] recommended by the guidelines in Europe [38] including in Denmark [39].

	Triage	Based on the 5-level triage system "Danish emergency department triage" (DEPT) [40, 41], we categorized the following: Red/Orange and Green/Blue were pooled due to few patients in the blue and red groups: 1= Red/Orange 2= Yellow 3= Green/Blue	DEPT is a Danish adaption and modification of the "Adaptive Process Triage" (ADAPT) developed in Sweden [42]. DEPT was chosen as it is routinely used in the three included sites. Furthermore, in Denmark, most EDs have implemented formalized triage called "Danish Emergency Process Triage". DEPT shares core similarities with widespread standardized 5-level triage systems [43].
Vital parameters All vital parameters regardless of diastolic blood pressure were based on The National Early Warning Score (NEWS) [44]. This score was chosen as it is routinely used in the three EDs included in this study and cut-offs values in predicting CAP are similar from the literature.	Oxygen saturation	Binary < 96 % (Yes/no) The cut-off was based on The National Early Warning Score (NEWS) [44]. However, we did not differentiate between patients with chronic obstructive pulmonary disease.	A similar cut-off of oxygen saturation has been used in investigating predictors for CAP [19].
	Heart rate	Binary < 51 or >90 bpm (Yes/no)	Some studies have investigated and pointed out that a higher heart rate with similar cut-offs as a predictor for CAP [19, 45, 46].
	Blood pressure systolic	Binary <111 or >219 mmHg (Yes/no)	Other cut-offs based on the CURB65-score or lower level of triage (<90mmHg) have been used to predict a high risk of adverse events among inpatients with CAP [47]. This cut-off was also explored in our model without resulting in any difference.
	Blood pressure diastolic	Binary ≤60 mmHg (Yes/no) Based on severity assessment CURB65-score [37]. The NEWS does not include diastolic blood pressure and therefore the value from CURB-65 was chosen.	CURB-65 is routinely used in Denmark as a severity score and is included in the guidelines for antibiotic treatment [39]. As systolic blood pressure has been investigated in prediction rules, we added diastolic blood pressure to our model to explore this variable as a predictor for CAP.
	Respiratory rate (RR)	Binary >20 breaths/min (Yes/no)	There are different cut-offs of RR in the literature [20, 47]. RR> 20/min was defined as a strong prediction of CAP among febrile patients [20].
	Temperature	Binary >38 ^{o}C (Yes/no) Measured with ear thermometer [48].	Different cut-offs have been investigated, including the cut-off of >38^{o}C used in this study [49]. Independent of cut-offs, several studies have identified fever as a predictor of CAP [19-21, 23, 45]. However, fever is less common and generally absent in the older population [7].
	Glascow coma score	Binary >15 (Yes/no)	Cognitive impairment [32] has been reported as a strong risk factor for delirium and confusion as a predictor of the severity of CAP [47]. Altered mental status is associated with CAP, especially in the elderly [18].
Blood tests The literature does not describe a clear cut-off for the diagnosis of CAP. We chose a pragmatic approach and applied the cut-offs of serum biomarkers used in the EDs from our institution to reflect reality. Most of the serological biomarkers have been studied for prognostic	Hematocrit	Hematocrit (%), median (IQR) Binary (Yes/no) Cut-off: 40-50 for males and 35-46 for females Yes= outside of the cut-off No= within the cut-off	A hematocrit value of less than 35% was an independent predictor for severity and 2 years of mortality (p = 0.035) [50].
	Hemoglobin	Hemoglobin mmol/L, median (IQR) Binary (Yes/no) Cut-off: 8.3-10.5 for males and 7.3-9.5 for females Yes= outside of the cut-off No= within the cut-off	Hemoglobin correlates with frailty in the elderly and indirectly could be a predictor that should be investigated [51].
	Leukocytes	Leukocytes 10E9/L, median (IQR) Binary (Yes/no) Cut-off: 3.5-8.8 Yes= outside of the cut-off No= within the cut-off	Elevated leukocytes have been reported as a predictor for CAP, especially in pneumonia with negative chest x-ray [52].

purposes. We have included these as potential predictors for CAP to investigate their diagnostic prediction performance combined with signs and symptoms. Binary (Yes/no) measures. Yes= abnormal/ outside of the cut-off No= normal/ within the cut-off	Platelets	Platelets 10E9/L, median (IQR) Binary (Yes/no) Cut-off: 145-350 for males and 165-400 for females Yes= outside of the cut-off No= within the cut-off	Platelet count < 171 × 109/L was included in a prediction model for *legionella pneumoniae* showing a high diagnostic accuracy [AUC 0.89 (95% CI 0.86–0.93)] [24].
	Neutrophils	Neutrophils 10E9/L, median (IQR) Binary (Yes/no) Cut-off: > 7.5 Yes= >7.5 No= ≤ 7.5	The neutrophil to lymphocyte ratio had a high diagnostic value for CAP patients [53]. Furthermore, higher mortality risk was found for CAP patients and if measured in the early stage of CAP could contribute to the diagnostic and disease severity [54].
	Lymphocytes	Lymphocytes 10E9/L, median (IQR) Binary (Yes/no) Cut-of: 1.00-4.00 Yes= outside of the cut-off No= within the cut-off	The neutrophil to lymphocyte ratio has been studied in prognostic studies and is associated with higher mortality risk in CAP patients and if measured in the early stage of CAP could contribute to the diagnostic and disease severity [54].
	Albumin	Albumin g/L, median (IQR) Binary (Yes/no) Cut-off: 34-45 Yes= outside of the cut-off No= within the cut-off	The ratio of blood urea and albumin has been investigated as a predictive factor for CAP, but poor model performance advocated for further investigation [55]. Furthermore, albumin correlates with frailty in the elderly and indirectly could be a predictor that should be investigated as frailty has been associated with an increased risk of CAP [51]. In addition, serum albumin (<3.4 g/dl) was associated with higher mortality for elderly patients with CAP [18] and was included in a prediction rule for severe adverse events in patients hospitalized with CAP (< 2 g/dL, 2 points; 2–3 g/dL, 1 point) [47].
	Creatinine	Creatinine µmol/L, median (IQR) Binary (Yes/no) Cut-off: 60-105 for males and 45-90 for females Yes= outside of the cut-off No= within the cut-off	Elevated creatinine levels have been reported with almost a sixfold association of poor CAP outcome (OR=5.67; 95%CI: 1.72-18.65) [56]. This result is supported by another study that showed that serum creatinine levels of ≥ 2.8 were a strong predictor of in-hospital mortality in adults with CAP when compared with five serum biomarkers [57].
	Blood urea	Blood urea nitrogen mmol/L, median (IQR) Binary (Yes/no) Cut-off: 3-5-8.1 for males and 3.1-7.9 for females Yes= outside of the cut-off No= within the cut-off	The ratio of blood urea and albumin has been investigated as a predictive factor for CAP, but poor model performance advocated for further investigation [55].
	Natrium	Natrium mmol/L, median (IQR) Binary (Yes/no) Cut-off: 137-145 Yes= outside of the cut-off No= within the cut-off	Hyponatremia < 133 mmol/L was one of the strong predictors in the prediction of CAP caused by *legionella pneumoniae* [24].
	Prothrombin time-international normalized ratio	Prothrombin (IQR) Binary (Yes/no) Cut-off: <1.2 Yes= ≥ 1.2 No= <1.2	Prothrombin time-international normalized ratio was investigated to distinguish Influenza A (H1N1) from other pneumonia. Prothrombin times were lower in H1N1 compared with non-H1N1 pneumonia patients (p=0.04) [58]. Furthermore, it has been investigated as a factor that could be associated with decreased sensitivity in negative urinary antigen (UAT) tests in CAP caused by pneumococcal. Prothrombin was 50% higher in the UAT-negative patients than in the UAT-positive patients [59]. We chose to include prothrombin in the diagnostic model to explore its significance in or rule out CAP, furthermore, the marker is routinely measured in acutely admitted patients.
	Bilirubin	Bilirubin µmol/L, median (IQR) Binary (Yes/no)	Bilirubin levels were lower in patients with influenza A (H1N1) compared to non-H1N1 pneumonia (p= 0.02)

		Cut-off: <5 or >25 Yes= outside of the cut-off No= within the cut-off	[58]. This marker could add value to a prediction model.
	Glucose	Glucose mmol/L, median (IQR) Binary (Yes/no) Cut-off: > 11.00 Yes= >11.00 No= ≤ 11.00	Patients with CAP frequently present with admission hyperglycemia and have poorer outcomes [60, 61]. Therefore, glucose is included as a potential predictor.
	C- reactive protein (CRP)	C-Reactive Protein, median (IQR) Binary (Yes/no) The cut-off of CRP in our institution is < 5 mg/L at the ED. However, the literature suggests optional cut-offs. Based on the literature and the range of the results from the CRP as continuous variable, we defined the following categories: 1= <20mg/L 2= 20-100 mg/L 3= >100 mg/L	The diagnostic accuracy of CRP in differentiating between bacterial and viral infections of the lower respiratory tract is questionable [62]. However, CRP at different cut-offs increased the performance of prediction models for CAP. It included a cut-off of >20 [20], >30 [63], 50 [23] ≥ 98 [46], and a meta-analysis investigated all three cut-offs of 20, 50, and 100 [64]. CRP levels were found higher when CAP was detected both by a chest x-ray and a chest tomography [52].
Clinical assessment	Stethoscope findings	Binary (Yes/no) Yes for any abnormal stethoscope findings such as crackles and rhonchi.	Several studies investigated associations between abnormal stethoscope findings and the probability of the presence of CAP. They increased the likelihood of CAP [21, 65] and crackles on auscultation had a twofold increase in the prediction of pneumonia [19].
	Abdominal pain on palpation	Binary (Yes/no)	The rationale for including abdominal pain in the clinical assessment was that the literature reported that 20% of symptoms reported by patients with CAP were gastrointestinal symptoms [26].
	Body mass index (BMI).	The BMI was calculated including the high and weight of the patients. The BMI classification was based on "The Centers for diseases control and prevention" [66] and defined with the following categories: 1= Underweight, BMI < 18.5 2= Healthy weight, BMI from 18.5 to <25 3= Overweight, BMI from 25.0 to <30 4= Obesity, BMI from ≥ 30.0	The literature reported the association of several nutritional factors related to CAP and including malnutrition [1, 18], being underweight [8, 17], and BMI was directly associated with an increased risk of CAP among women [10].

Table S2: Characteristics of CAP in the population of patients admitted with an infection (n=954). The values presented of data as continuous, dichotomous or categorical were tested in the model during explorative analysis to identify the best model performance.

Characteristics	Total, n	CAP, n	Not CAP, n	Missings n (%)	OR (95% CI)	p-value
Total of patients	954 (100)	265 (27.8)	689 (72.2)	0 (0.0)		
DEMOGRAPHIC DATA						
Age, median (IQR)	73.0 (59.0; 81.0)	75.0 (63.5; 82.0)	73.0 (57.0; 80.0)	0 (0.0)	1.01 (1.005-1.02)	<0.001
Age ≥75 years	440 (46.1)	133 (50.2)	307 (44.6)	0 (0.0)	1.25 (0.94-1.66)	0.118
Gender male	513 (53.8)	137 (51.7)	376 (54.6)	0 (0.0)	0.89 (0.67-1.18)	0.425
Marital status, Living alone	618 (66.0)	166 (63.8)	452 (66.9)	18 (1.9)	0.87 (0.64-1.18)	0.382
Nursing home resident	66 (7.0)	26 (9.9)	40 (5.9)	13 (1.4)	1.75 (1.05-2.94)	0.317
Occupation				21 (2.2)		
Others	67 (7.2)	17 (6.5)	50 (7.4)		1 (reference)	
Working	202 (21.7)	44 (16.9)	158 (23.5)		0.81 (0.43-1.55)	0.543
Retired	664 (71.2)	200 (76.6)	464 (69.0)		1.26 (0.71-2.25)	0.418
LIFESTYLE FACTORS						
Smoking status				33 (3.5)		
No	323 (35.1)	66 (26.0)	257 (38.5)		1 (reference)	
Current smoker	179 (19.4)	54 (21.3)	125 (18.7)		1.68 (1.10-2.55)	0.015
Previous smoker	419 (45.5)	134 (52.8)	285 (42.7)		1.83 (1.30-2.57)	<0.001
Alcohol status				35 (3.7)		
No alcohol	356 (38.7)	99 (39.1)	257 (38.6)		1 (reference)	
1-7 doses	385 (41.9)	105 (41.5)	280 (42.0)		0.97 (0.70-1.34)	0.870
8-14 doses	105 (11.4)	31 (12.3)	74 (11.1)		1.08 (0.67-1.75)	0.732
> 14 doses	73 (7.9)	18 (7.1)	55 (8.3)		0.84 (0.47-1.51)	0.582
Physically activity				52 (5.4)		
Not physical active	263 (29.2)	74 (29.8)	189 (28.9)		1 (reference)	
Physical activity < 2,5 hr/week	231 (25.6)	64 (25.8)	167 (25.5)		0.97 (0.66-1.45)	0.915
Physical activity ≥ 2,5 hr/week	408 (45.2)	110 (44.4)	298 (45.6)		0.94 (0.66-1.33)	0.735
Body Mass Index, median (IQR)	26.5 (23.2; 30.8)	26.2 (22.9; 29.5)	26.7 (23.3; 31.2)	249 (26.1)	0.97 (0.94-0.99)	0.031
Body Mass Index†				249 (26.1)		
Healthy weight	246 (34.9)	74 (36.1)	172 (34.4)		1 (reference)	
Obese	193 (27.4)	45 (22.0)	148 (29.6)		0.70 (0.45-1.08)	0.114
Overweight	239 (33.9)	74 (36.1)	165 (33.0)		1.04 (0.70-1.53)	0.833
Underweight	27 (3.8)	12 (5.9)	15 (3.0)		1.85 (0.83-4.16)	0.132
ADL dependence*	260 (28.0)	81 (31.2)	179 (26.8)	25 (2.6)	1.23 (0.90-1.69)	0.180
SYMPTOMS						
Feeling unwell	559 (61.2)	173 (67.8)	386 (58.7)	41 (4.3)	1.48 (1.09-2.01)	0.010
Feeling tired	657 (72.6)	190 (75.4)	467 (71.5)	49 (5.1)	1.22 (0.87-1.70)	0.241
Headache	351 (38.3)	99 (38.8)	252 (38.1)	37 (3.9)	1.03 (0.76-1.38)	0.832
Dizziness	346 (37.7)	96 (37.6)	250 (37.8)	37 (3.98)	0.99 (0.73-1.34)	0.973
Confusion	207 (22.6)	58 (22.7)	149 (22.5)	37 (3.89)	1.01 (0.71-1.43)	0.938
Dyspnea	379 (41.4)	171 (67.3)	208 (31.5)	39 (4.1)	4.48 (3.29-6.11)	<0.001
Cough	358 (39.1)	173 (68.1)	185 (28.0)	39 (4.1)	5.49 (4.01-7.52)	<0.001

Expectoration	279 (30.5)	140 (55.1)	139 (21.0)	39 (4.1)	4.61 (3.38-6.28)	<0.001
Sore throat	104 (11.4)	39 (15.4)	65 (9.8)	39 (4.1)	1.66 (1.08-2.54)	0.019
Cold (common cold)	95 (10.4)	45 (17.7)	50 (7.6)	39 (4.1)	2.63 (1.70-4.05)	<0.001
Fever feling at home	612 (64.2)	169 (63.8)	443 (64.3	0 (0.0)	0.97 (0.72-1.31)	0.880
Chest pain	168 (18.4)	71 (28.1)	97 (14.7)	40 (4.2)	2.26 (1.60-3.21)	<0.001
Oedema	79 (8.6)	10 (4.0)	69 (10.4)	39 (4.1)	0.35 (1,17-0.69)	0.002
Nausea	304 (33.2)	76 (30.0)	228 (34.4)	38 (3.9)	0.81 (0.59-1.112)	0.211
Vomiting	190 (20.7)	40 (15.8)	150 (22.6)	38 (3.9)	0.64 (0.43-0.94)	0.023
Loss of appetite	524 (57.2)	149 (58.9)	375 (56.6)	38 (3.9)	1.00 (0.82-1.47)	0.523
Gastrointestinal pain	193 (21.1)	40 (15.8)	153 (23.1)	38 (3.9)	0.62 (0.42-0.91)	0.016
Diarrhoea	134 (14.6)	29 (11.5)	105 (15.8)	38 (3.9)	0.68 (0.44-1.06)	0.095
Muscular pain	344 (37.8)	79 (31.3)	265 (40.3)	44 (4.6)	0.67 (0.49-0.92)	0.013
Back pain	132 (14.5)	33 (13.1)	99 (15.0)	44 (4.6)	0.85 (0.55-1.29)	0.455
CLINICAL ASSESSMENT						
Positive stethoscope findings	329 (36.5)	168 (65.4)	161 (25.0)	52 (5.4)	5.67 (4.15-7.75)	<0.001
Abdominal pain by palpation	192 (22.1)	37 (15.0)	155 (25.0)	86 (9.0)	0.52 (0.35-0.78)	0.002
COMORBIDITIES						
Dementia	32 (3.4)	9 (3.4)	23 (3.3)	0 (0.0)	1.01 (0.46-2.22)	0.964
Neurological diseases	172 (18.0)	53 (20.0)	119 (17.3)	0 (0.0)	1.19 (0.83-1.71)	0.326
Respiratory diseases	269 (28.2)	105 (39.6)	164 (23.8)	0 (0.0)	2.10 (1.55-2.84)	<0.001
Endocrinological diseases	296 (31.0)	80 (30.2)	216 (31.3)	0 (0.0)	0.94 (0.69-1.28)	0.728
Nephrological diseases	252 (26.4)	60 (22.6)	192 (27.9)	0 (0.0)	0.75 (0.54-1.05)	0.101
Cardiovascular diseases	390 (40.9)	116 (43.8)	274 (39.8)	0 (0.0)	1.17 (0.88-1.57)	0.259
Gastrointestinal diseases	100 (10.5)	23 (8.7)	77 (11.2)	0 (0.0)	0.75 (0.46-1.23)	0.260
Rheumatological diseases	118 (12.4)	27 (10.2)	91 (13.2)	0 (0.0)	0.74 (0.47-1.17)	0.205
Cancer diseases	85 (8.9)	26 (9.8)	59 (8.6)	0 (0.0)	1.16 (0.71-1.88)	0.544
Prior pneumonia				100 (10.5)		
No	410 (48.0)	79 (33.3)	331 (53.6)		1 (reference)	
Yes, one time	180 (21.1)	50 (21.1)	130 (21.1)		1.61 (1.07-2.42)	0.022
Yes, more than one time	264 (30.9)	108 (45.6)	156 (25.3)		2.90 (2.05-4.10)	<0.001
SEVERITY ASSESSMENT						
CURB65 ≥3 **	122 (13.0)	29 (11.3)	93 (13.7)	16 (1.7)	0.80 (0.51-1.25)	0.336
Triage***				59 (6.2)		
Green/Blue	183 (20.4)	37 (14.8)	146 (22.6)		1 (reference)	
Yellow	479 (53.5)	126 (50.4)	353 (54.7)		1.40 (0.93-2.13)	0.105
Red/Orange	233 (26.0)	87 (34.8)	146 (22.6)		2.35 (1.50-3.67)	<0.001
VITAL PARAMETERS						
Respiratory rate, median(IQR)	18.0 (16.0; 22.0)	20.0 (18.0; 24.0)	18.0 (16.0; 20.0)	5 (0.5)	1.10 (1.07-1.13)	<0.001
Respiratory rate >20/min	285 (30.0)	124 (47.0)	161 (23.5)	5 (0.5)	2.88 (2.13-3.88)	<0.001
Oxygen saturation % n/min, median (IQR)	96.0 (94.0; 98.0)	95.0 (93.0; 97.0)	97.0 (95.0; 98.0)	4 (0.4)	0.84 (0.80-0.88)	<0.001
Oxygen saturation < 96 %	393 (41.4)	162 (61.1)	231 (33.7)	4 (0.4)	3.09 (2.30-4.14)	<0.001
Heart rate/min, mean (sd)	90.1 (18.3)	93.2 (18.9)	88.9 (18.0)	1 (0.1)	1.01 (1.005-1.02)	0.001
Heart rate <51 or >90/min	460 (48.3)	148 (55.8)	312 (45.3)	1 (0.1)	1.52 (1.14-2.02)	0.003
Systolic blood pressure mmHg, mean (sd)	132.8 (22.5)	134.2 (21.0)	132.2 (23.1)	3 (0.3)	1.003 (0.99-1.01	0.215

Systolic blood pressure <111 or >219 mmHg	156 (16.4)	38 (14.4)	118 (17.2)	3 (0.3)	0.81 (0.54-1.21)	0.314
Diastolic blood pressure mmHg, mean (sd)	74.8 (15.3)	74.2 (13.6)	75.0 (15.8)	3 (0.3)	0.99 (0.98-1.006)	0.483
Diastolic blood pressure ≤60 mmHg	163 (17.1)	40 (15.2)	123 (17.9)	3 (0.3)	0.82 (0.55-1.21)	0.329
Temperature, mean (SD)	37.5 (1.0)	37.6 (1.0)	37.4 (0.9)	5 (0.5)	1.22 (1.05-1.40)	0.006
Fever > 38^{o}C	233 (24.6)	77 (29.3)	156 (22.7)	5 (0.5)	1.40 (1.02-1.93)	0.036
Glascow coma scale <15	31 (3.3)	12 (4.6)	19 (2.8)	5 (0.5)	0.59 (0.28-1.24)	0.168
BLOOD TESTS						
Haematocrit, median (IQR)	38.0 (35.0; 42.0)	38.0 (35.0; 42.0)	39.0 (35.0; 42.0)	260 (27.2)	0.98 (0.95-1.01)	0.465
Haematocrit	268 (38.6)	85 (38.6)	183 (38.6)	260 (27.2)	1.001 (0.72-1.39)	0.994
Haemoglobin mmol/L, median (IQR)	8.0 (7.2; 8.7)	7.9 (7.2; 8.6)	8.0 (7.3; 8.8)	0 (0.0)	0.90 (0.80-1.02)	0.127
Haemoglobin mmol/L	402 (42.1)	118 (44.5)	284 (41.2)	0 (0.0)	1.14 (0.86-1.52)	0.354
Leukocytes 10E9/L, median (IQR)	11.1 (8.3; 14.8)	12.2 (9.5; 15.8)	10.7 (8.0; 14.2)	0 (0.0)	1.05 (1.02-1.07)	<0.001
Leukocytes 10E9/L	670 (70.2)	214 (80.8)	456 (66.2)	0 (0.0)	2.14 (1.52-3.02)	<0.001
Platelets 10E9/L, median (IQR)	240.0 (189.0; 307.8)	260.5 (211.0; 330.8)	232.0 (182.3; 296.0)	10 (1.0)	1.002 (1.001-1.004)	<0.001
Platelets 10E9/L	201 (21.3)	63 (23.9)	138 (20.3)	10 (1.0)	1.23 (0.87-1.72)	0.229
Neutrophils 10E9/L, median (IQR)	8.4 (6.0; 12.2)	9.7 (7.2; 13.0)	8.0 (5.6; 11.6)	10 (1.0)	1.06 (1.03-1.09)	<0.001
Neutrophils 10E9/L	549 (58.2)	187 (71.1)	362 (53.2)	10 (1.0)	2.16 (1.59-2.94)	<0.001
Lymphocytes† 10E9/L, median (IQR)	1.1 (0.7; 1.6)	0.9 (0.6; 1.5)	1.2 (0.8; 1.8)	633 (66.3)	0.98 (0.85-1.12)	0.797
Lymphocytes† 10E9/L	145 (45.2)	53 (55.2)	92 (40.9)	633 (66.3)	1.78 (1.10-2.88)	0.018
Albumin g/L, median (IQR)	39.0 (36.0; 42.0)	39.0 (35.0; 41.0)	39.0 (36.0; 42.0)	7 (0.7)	0.96 (0.93-0.99)	0.029
Albumin g/L	160 (16.9)	39 (14.9)	121 (17.6)	7 (0.7)	0.82 (0.55-1.21)	0.323
Creatinine µmol/L, median (IQR)	84.0 (67.0; 113.0)	81.0 (64.0; 108.0)	86.0 (67.5; 114.0)	0 (0.0)	0.996 (0.993-0.998)	0.003
Creatinine µmol/L	374 (39.2)	106 (40.0)	268 (38.9)	0 (0.0)	1.04 (0.78-1.39)	0.754
Blood urea nitrogen mmol/L, median (IQR)	6.2 (4.4; 8.9)	6.2 (4.5; 8.6)	6.2 (4.4; 9.1)	9 (0.9)	0.99 (0.96-1.02)	0.657
Blood urea nitrogen mmol/L	377 (39.9)	99 (38.1)	278 (40.6)	9 (0.9)	0.90 (0.67-1.20)	0.482
Natrium mmol/L, median (IQR)	137.0 (134.0; 139.0)	137.0 (134.0; 139.0)	137.0 (134.0; 139.0)	0 (0.0)	0.98 (0.95-1.01)	0.394
Natrium mmol/L	432 (45.3)	128 (48.3)	304 (44.1)	0 (0.0)	1.18 (0.89-1.57)	0.245
Prothrombin, median (IQR)	1.1 (1.0; 1.2)	1.1 (1.0; 1.2)	1.1 (1.0; 1.2)	3 (0.3)	1.18 (0.89-1.58)	0.231
Prothrombin	234 (24.6)	65 (24.5)	169 (24.6)	3 (0.3)	0.99 (0.71-1.38)	0.972
Bilirubin µmol/L, median (IQR)	9.0 (6.0; 13.0)	9.0 (6.0; 12.0)	9.0 (6.0; 14.0)	11 (1.1)	0.97 (0.95-0.99)	0.254
Bilirubin µmol/L	152 (16.1)	38 (14.4)	114 (16.8)	11 (1.1)	0.83 (0.55-1.24)	0.369
Glucose mmol/L, median (IQR)	6.7 (5.9; 7.9)	6.9 (6.2; 8.1)	6.6 (5.8; 7.8)	9 (0.9)	1.04 (0.99-1.10)	0.052
Glucose mmol/L	51 (5.4)	19 (7.3)	32 (4.7)	9 (0.9)	1.59 (0.88-2.85)	0.120
C-Reactive Protein mg/L, median (IQR)	95.5 (30.0; 179.3)	125.0 (57.0; 203.5)	82.0 (19.0; 172.0)	0 (0.0)	1.003 (1.001-1.004)	<0.001
C-Reactive Protein mg/L				0 (0.0)		
Low <20mg/L	196 (20.5)	21 (7.9)	175 (25.4)		1 (reference)	
Moderate 21-99 mg/L	291 (30.5)	86 (32.5)	205 (29.8)		3.49 (2.08-5.86)	<0.001
High >=100	467 (49.0)	158 (59.6)	309 (44.8)		4.26 (2.60-6.96)	<0.001
VACCINE AND MEDICAMENTATIONS						
SARS-CoV-2 †	756 (79.2)	222 (83.8)	534 (77.5)	0 (0.0)	1.49 (1.03-2.17)	0.033
Pneumococcal	530 (55.6)	160 (60.4)	370 (53.7)	0 (0.0)	1.31 (0.98-1.75)	0.063
Influenza	635 (66.6)	191 (72.1)	444 (64.4)	0 (0.0)	1.42 (1.04-1.94)	0.025

Analgesics	404 (42.3)	115 (43.4)	289 (41.9)	0 (0.0)	1.06 (0.79-1.41)	0.684
Polypharmacy****	544 (57.0)	163 (61.5)	381 (55.3)	0 (0.0)	1.29 (0.96-1.72)	0.082

Values are numbers (percentages) unless otherwise specified. *ADL dependence: If the patient had one or more dependencies regarding bathing, dressing, toileting, transfer, continence, and feeding. ** CURB65: confusion, uraemia, respiratory rate, blood pressure, age > 65 years. ***Triage: Danish emergency process triage [40] ****Polypharmacy: regular consumption of at least five medications † variables not included in the multivariate model

Table S3: Characteristics of the 954 patients with suspected infection enrolled in the study. It presents the 70 predictors included in the multivariate analysis and randomization of the training set and validation set.

Characteristics	Total, n	Training set, n	Validation set, n	Missings n (%)	p-value
Total of patients	954 (100)	766 (80.3)	188 (19.7)	0 (0.0)	
DEMOGRAPHIC DATA					
Age, median (IQR)	73.0 (59.0; 81.0)	75.0 (63.5; 82.0)	74.0 (60.0; 82.0)	0 (0.0)	0.54
Age ≥75 years	440 (46.1)	348 (45.4)	92 (48.9)	0 (0.0)	0.39
Gender male	513 (53.8)	408 (53.3)	105 (55.9)	0 (0.0)	0.52
Marital status, Living alone	618 (66.0)	488 (65.0)	130 (70.3)	18 (1.9)	0.17
Nursing home resident	66 (7.0)	55 (7.3)	11 (5.9)	13 (1.4)	0.53
Occupation				21 (2.2)	0.62
Others	67 (7.2)	57 (7.6)	10 (5.5)		
Working	202 (21.7)	162 (21.6)	40 (22.0)		
Retired	664 (71.2)	532 (70.8)	132 (72.5)		
LIFESTYLE FACTORS					
Smoking status				33 (3.5)	0.76
No	323 (35.1)	256 (34.5)	67 (37.4)		
Current smoker	179 (19.4)	145 (19.5)	34 (19.0)		
Previous smoker	419 (45.5)	341 (46.0)	78 (43.6)		
Alcohol status				35 (3.7)	0.60
No alcohol	356 (38.7)	283 (38.2)	73 (40.8)		
1-7 doses	385 (41.9)	315 (42.6)	70 (39.1)		
8-14 doses	105 (11.4)	81 (10.9)	24 (13.4)		
> 14 doses	73 (7.9)	61 (8.2)	12 (6.7)		
Physically activity				52 (5.4)	0.76
Not physical active	263 (29.2)	214 (29.4)	49 (28.2)		
Physical activity < 2,5 hr/week	231 (25.6)	189 (26.0)	42 (24.1)		
Physical activity ≥ 2,5 hr/week	408 (45.2)	325 (44.6)	83 (47.7)		
Body Mass Index†				249 (26.1)	0.74

Healthy weight	246 (34.9)	202 (35.8)	44 (31.2)		
Obese	193 (27.4)	154 (27.3)	39 (27.7)		
Overweight	239 (33.9)	187 (33.2)	52 (36.9)		
Underweight	27 (3.8)	21 (3.7)	6 (4.3)		
ADL dependence*	260 (28.0)	203 (27.1)	57 (31.7)	25 (2.6)	0.22
SYMPTOMS					
Malaise	559 (61.2)	458 (62.0)	101 (58.0)	41 (4.3)	0.34
Feeling tired	657 (72.6)	540 (74.0)	117 (66.9)	49 (5.1)	0.06
Headache	351 (38.3)	287 (38.8)	64 (36.0)	37 (3.9)	0.48
Dizziness	346 (37.7)	287 (38.8)	59 (33.1)	37 (3.98)	0.16
Confusion	207 (22.6)	164 (22.2)	43 (24.2)	37 (3.89)	0.57
Dyspnea	379 (41.4)	309 (42.0)	70 (39.1)	39 (4.1)	0.48
Cough	358 (39.1)	294 (39.9)	64 (35.8)	39 (4.1)	0.30
Fever feeling at home	612 (64.2)	464 (64.5)	118 (62.8)	0 (0.0)	0.66
Expectoration	279 (30.5)	224 (30.4)	55 (30.7)	39 (4.1)	0.94
Sore throat	104 (11.4)	86 (11.7)	18 (10.1)	39 (4.1)	0.54
Cold (common cold)	95 (10.4)	81 (11.0)	14 (7.8)	39 (4.1)	0.21
Chest pain	168 (18.4)	134 (18.2)	34 (19.0)	40 (4.2)	0.81
Oedema	79 (8.6)	61 (8.3)	18 (10.1)	39 (4.1)	0.45
Nausea	304 (33.2)	247 (33.4)	57 (32.2)	38 (3.9)	0.76
Vomiting	190 (20.7)	154 (20.8)	36 (20.3)	38 (3.9)	0.88
Loss of appetite	524 (57.2)	424 (57.4)	100 (56.5)	38 (3.9)	0.83
Gastrointestinal pain	193 (21.1)	145 (19.6)	48 (27.1)	38 (3.9)	0.03
Diarrhoea	134 (14.6)	107 (14.5)	27 (15.3)	38 (3.9)	0.79
Muscular pain	344 (37.8)	289 (39.5)	55 (30.9)	44 (4.6)	0.03
Back pain	132 (14.5)	110 (15.0)	22 (12.4)	44 (4.6)	0.36
CLINICAL ASSESSMENT					
Positive stethoscope findings	329 (36.5)	263 (36.5)	66 (36.5)	52 (5.4)	1.00
Abdominal pain by palpation	192 (22.1)	151 (21.7)	41 (23.7)	86 (9.0)	0.58
COMORBIDITIES					
Dementia	23 (3.0)	9 (4.8)	23 (3.3)	0 (0.0)	0.22
Neurological diseases	137 (17.9)	35 (18.6)	119 (17.3)	0 (0.0)	0.82
Pulmonary diseases	212 (27.7)	57 (30.3)	164 (23.8)	0 (0.0)	0.47
Endocrinological diseases	239 (31.2)	57 (30.3)	216 (31.3)	0 (0.0)	0.81
Nephrological diseases	200 (26.1)	52 (27.7)	192 (27.9)	0 (0.0)	0.67

Cardiovascular diseases	303 (39.6)	87 (46.3)	274 (39.8)	0 (0.0)	0.09
Gastrointestinal diseases	81 (10.6)	19 (10.1)	77 (11.2)	0 (0.0)	0.85
Rheumatological diseases	93 (12.1)	25 (13.3)	91 (13.2)	0 (0.0)	0.67
Cancer diseases	66 (8.6)	19 (10.1)	59 (8.6)	0 (0.0)	0.52
Prior pneumonia				100 (10.5)	0.05
No	343 (50.1)	67 (39.6)	331 (53.6)		
Yes, one time	139 (20.3)	41 (24.3)	130 (21.1)		
Yes, more than one time	203 (29.6)	61 (36.1)	156 (25.3)		
SEVERITY ASSESSMENT					
CURB65 ≥3 **	103 (13.6)	19 (10.4)	93 (13.7)	16 (1.7)	0.25
Triage***				59 (6.2)	0.53
Green/Blue	185 (25.6)	48 (27.9)	146 (22.6)		
Yellow	385 (53.3)	94 (54.7)	353 (54.7)		
Red/Orange	153 (21.2)	30 (17.4)	146 (22.6)		
VITAL PARAMETERS					
Respiratory rate >20/min	285 (30.0)	235 (30.8)	50 (26.7)	5 (0.5)	0.27
Oxygen saturation < 96 %	393 (41.4)	324 (42.5)	69 (36.7)	4 (0.4)	0.15
Heart rate <51 or >90/min	460 (48.3)	377 (49.3)	83 (44.1)	1 (0.1)	0.21
Systolic blood pressure <111 or >219 mmHg	156 (16.4)	125 (16.4)	31 (16.6)	3 (0.3)	0.94
Diastolic blood pressure ≤60 mmHg	163 (17.1)	131 (17.1)	32 (17.1)	3 (0.3)	0.99
Fever > 38°C	233 (24.6)	190 (24.9)	43 (23.1)	5 (0.5)	0.61
Glascow coma scale <15	31 (3.3)	23 (3.0)	8 (4.3)	5 (0.5)	0.39
BLOOD TESTS					
Haematocrit	268 (38.6)	218 (39.2)	50 (36.2)	260 (27.2)	0.52
Haemoglobin mmol/L	402 (42.1)	329 (43.0)	73 (38.8)	0 (0.0)	0.31
Leukocytes 10E9/L	670 (70.2)	548 (71.5)	122 (64.9)	0 (0.0)	0.07
Platelets 10E9/L	201 (21.3)	168 (22.2)	33 (17.6)	10 (1.0)	0.17
Neutrophils 10E9/L	549 (58.2)	454 (59.9)	95 (51.1)	10 (1.0)	0.03
Albumin g/L	160 (16.9)	130 (17.1)	30 (16.1)	7 (0.7)	0.76
Creatinine µmol/L	374 (39.2)	303 (39.6)	71 (37.8)	0 (0.0)	0.65
Blood urea nitrogen mmol/L	377 (39.9)	308 (40.5)	69 (37.5)	9 (0.9)	0.46
Natrium mmol/L	432 (45.3)	362 (47.3)	70 (37.2)	0 (0.0)	0.01
Prothrombin	234 (24.6)	186 (24.3)	48 (25.7)	3 (0.3)	0.71
Bilirubin µmol/L	152 (16.1)	119 (15.7)	33 (17.8)	11 (1.1)	0.48

Glucose mmol/L	51 (5.4)	42 (5.5)	9 (4.8)	9 (0.9)	0.71
C-Reactive Protein mg/L				0 (0.0)	0.07
<20 mg/L	196 (20.5)	151 (19.7)	45 (23.9)		
21-99 mg/L	291 (30.5)	226 (29.5)	65 (34.6)		
≥ 100 mg/L	467 (49.0)	389 (50.8)	78 (41.5)		
VACCINE AND MEDICAMENTATIONS					
Pneumococcal	530 (55.6)	414 (54.0)	116 (61.7)	0 (0.0)	0.06
Influenza	635 (66.6)	512 (66.8)	123 (65.4)	0 (0.0)	0.71
Analgesics	404 (42.3)	336 (43.9)	68 (36.2)	0 (0.0)	0.06
Polypharmacy****	544 (57.0)	443 (57.8)	101 (53.7)	0 (0.0)	0.31

Values are numbers (percentages) unless otherwise specified. *ADL dependence: If the patient had one or more dependencies regarding bathing, dressing, toileting, transfer, continence, and feeding. ** CURB65: confusion, uraemia, respiratory rate, blood pressure, age > 65 years. ***Triage: Danish emergency process triage [40] ****Polypharmacy: regular consumption of at least five medications

Figure S1: Performance of the prediction model presented with the area receiver operating characteristic curve

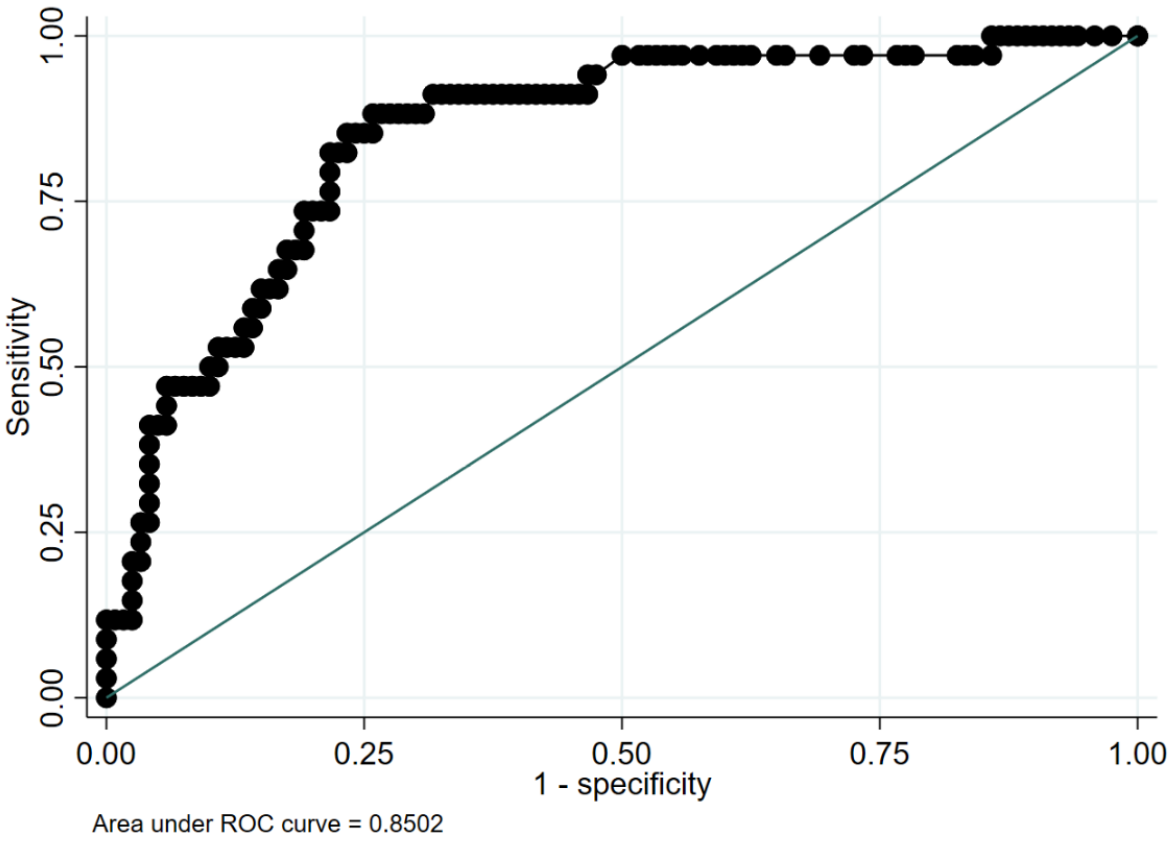
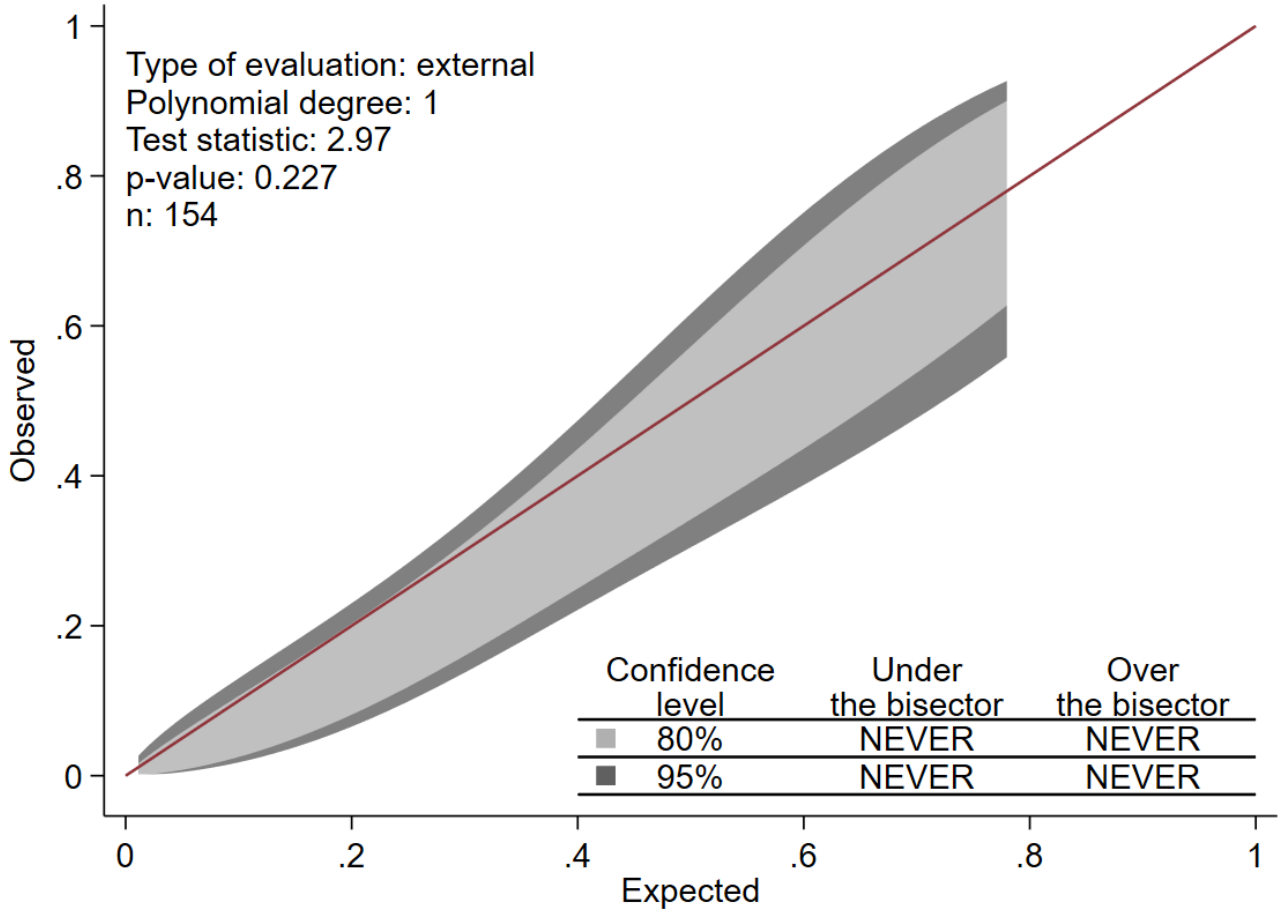


Figure S2: The calibration of the model after recalibration



Formula S1: Based on a lambda result of λ=0.0402856 and a probability threshold of 0.35, the LASSO calculation with characteristics predictive of CAP as follows:

$$
\begin{aligned}
CAP-score = {} & 0.07 \cdot 1_{Unwell=yes} + 0.35 \cdot 1_{Dyspnea=yes} + 0.36 \cdot 1_{Expectoration=yes} + 0.39 \cdot 1_{Cough=yes} \\
& + 0.34 \cdot 1_{Cold=yes} + 0.14 \cdot 1_{Respiratory\ rate\ >20/min=yes} + 0.24 \\
& \cdot 1_{Oxygen\ saturation<96\%=yes} + 0.20 \cdot 1_{Chest\ pain=yes} + 0.56 \cdot 1_{Stethoscope=yes} - 0.12 \\
& \cdot 1_{Previous\ CAP=no} + 0.003 \cdot 1_{Leucocytes<3.5\ or\ >8.8\ 10E9\ /L=yes} + 0.08 \\
& \cdot 1_{Neutrophilocytes>7.5\ 10E9\ /L=yes} - 0.64 \cdot 1_{CRP<20mg\ /L=yes} + 0.53 \cdot 1_{Cough=yes} \cdot 1_{age\geq 75} \\
& - 0.05 . 1_{Edema=yes} \cdot 1_{age\geq 75} + 0.88 \cdot 1_{Glucose>11\ mmol\ /L=yes} \cdot 1_{age\geq 75} + 0.0402856 \\
& \cdot (0.07 + 0.35 + 0.36 + 0.39 + 0.015 + 0.34 + 0.14 + 0.24 + 0.20 + 0.56 + 0.12 \\
& + 0.003 + 0.08 + 0.64 + 0.53 + 0.05 + 0.88) - 1.66192 - \log\left(\frac{0.35}{0.65}\right)
\end{aligned}
$$

For best calibration, 0.07 must be subtracted from the score if the score is between 0.08 and 0.47.

Formula S2: A cutoff value greater than 0 indicates the diagnosis CAP according to our model and can be calculated using the following formula:

$$CAP - score = 0.07 \cdot 1_{Unwell=yes} + 0.35 \cdot 1_{Dyspnea=yes} + 0.36 \cdot 1_{Expectoration=yes} + 0.39 \cdot 1_{Cough=yes} + 0.34 \cdot 1_{Cold=yes} + 0.14 \cdot 1_{Respiratory\ rate\ >20/min=yes} + 0.24 \cdot 1_{Oxygen\ saturation<96\%=yes} + 0.20 \cdot 1_{Chest\ pain=yes} + 0.56 \cdot 1_{Stethoscope=yes} - 0.12 \cdot 1_{Previous\ CAP=no} + 0.003 \cdot 1_{Leucocytes<3.5\ or\ >8.8\ 10E9\ /L=yes} + 0.08 \cdot 1_{Neutrophilocytes>7.5\ 10E9\ /L=yes} - 0.64 \cdot 1_{CRP<20mg\ /L=yes} + 0.53 \cdot 1_{Cough=yes} \cdot 1_{age\geq 75} - 0.05 \cdot 1_{Edema=yes} \cdot 1_{age\geq 75} + 0.88 \cdot 1_{Glucose>11\ mmol\ /L=yes} \cdot 1_{age\geq 75} - 0.842742$$

For best calibration, 0.07 must be subtracted from the score if the score is between 0.08 and 0.47.

Model specification

Besides the high percentage of missings from lymphocytes (66.3%), lymphocytes contributed to a significantly decreased model performance below 80% and a narrower calibration belt (p<0.001), furthermore lymphocytes were missing for 66.3% of the patients. SARS-CoV-2 vaccine was not included in the final model as the vaccine was related to a specific pandemic and did not change any final predictors or values. The inclusion of the BMI had better prediction performance AUC: 0.86 (CI: 0.79-0.93) and yielded more predictors especially related to lifestyle. The predictors that differed from the final model were: Alcohol (8-14 doses/week) 0.01792, level of physical activity under 2,5 hours/week yielded 0.01067, and obesity appeared with a coefficient of -0.93861. In addition, a symptom of diarrhea (-0.17572), muscular pain (-0.00225), gastrointestinal symptoms (-0.807885), sore throat (0.074709 for patients ≥ 75 years old) and the presence of nephrological diseases (-0.18776 for patients ≥ 75 years old) were predictors of CAP in the model constructed including BMI. From a clinical perspective, we chose to exclude the BMI as the final model would be more useful in an acute setting where reliable information about BMI is not always available. From a statistical perspective, BMI had almost 27% of missings, which would be classified as MAR and possibly selected from the population.

References

1. Almirall, J., et al., *Risk Factors for Community-Acquired Pneumonia in Adults: A Systematic Review of Observational Studies.* Respiration, 2017. **94**(3): p. 299-311.
2. Riquelme, R., et al., *Community-acquired pneumonia in the elderly. Clinical and nutritional aspects.* Am J Respir Crit Care Med, 1997. **156**(6): p. 1908-14.
3. Janssens, J.P., *Pneumonia in the elderly (geriatric) population.* Curr Opin Pulm Med, 2005. **11**(3): p. 226-30.
4. Cillóniz, C., et al., *Community-acquired pneumonia in critically ill very old patients: a growing problem.* Eur Respir Rev, 2020. **29**(155).
5. Metlay, J.P., et al., *Influence of age on symptoms at presentation in patients with community-acquired pneumonia.* Arch Intern Med, 1997. **157**(13): p. 1453-9.
6. Laporte, L., et al., *Ten-year trends in intensive care admissions for respiratory infections in the elderly.* Ann Intensive Care, 2018. **8**(1): p. 84.
7. Ravioli, S., et al., *Age- and sex-related differences in community-acquired pneumonia at presentation to the emergency department: a retrospective cohort study.* Eur J Emerg Med, 2022. **29**(5): p. 366-372.
8. Hammond, A., et al., *Predisposing factors to acquisition of acute respiratory tract infections in the community: a systematic review and meta-analysis.* BMC Infect Dis, 2021. **21**(1): p. 1254.
9. Barbagelata, E., et al., *Gender differences in community-acquired pneumonia.* Minerva Med, 2020. **111**(2): p. 153-165.
10. Baik, I., et al., *A prospective study of age and lifestyle factors in relation to community-acquired pneumonia in US men and women.* Arch Intern Med, 2000. **160**(20): p. 3082-8.
11. Heath, G.W., et al., *Exercise and the incidence of upper respiratory tract infections.* Medicine and science in sports and exercise, 1991. **23**(2): p. 152-157.
12. Kim, N.E., et al., *Clinical characteristics and outcomes among older nursing home residents hospitalized with pneumonia.* Arch Gerontol Geriatr, 2021. **95**: p. 104394.
13. Nieman, D.C., et al., *Upper respiratory tract infection is reduced in physically fit and active adults.* British journal of sports medicine, 2011. **45**(12): p. 987-992.
14. Calvillo–King, L., et al., *Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review.* Journal of general internal medicine, 2013. **28**(2): p. 269-282.
15. Clyde, W.A., Jr., *Clinical overview of typical Mycoplasma pneumoniae infections.* Clin Infect Dis, 1993. **17 Suppl 1**: p. S32-6.
16. Gamble, J.M., et al., *Medication transitions and polypharmacy in older adults following acute care.* Ther Clin Risk Manag, 2014. **10**: p. 189-96.
17. Torres, A., et al., *Risk factors for community-acquired pneumonia in adults in Europe: a literature review.* Thorax, 2013. **68**(11): p. 1057-65.
18. Riquelme, R., et al., *Community-acquired pneumonia in the elderly: clinical and nutritional aspects.* Revista médica de Chile, 2008. **136**(5): p. 587-593.
19. Moore, M., et al., *Predictors of pneumonia in lower respiratory tract infections: 3C prospective cough complication cohort study.* Eur Respir J, 2017. **50**(5).
20. Ding, F., et al., *Development and validation of a simple tool composed of items on dyspnea, respiration rates, and C-reactive protein for pneumonia prediction among acute febrile respiratory illness patients in primary care settings.* BMC Med, 2022. **20**(1): p. 360.
21. Nakanishi, M., et al., *Significance of the progression of respiratory symptoms for predicting community-acquired pneumonia in general practice.* Respirology, 2010. **15**(6): p. 969-74.

22. Huijskens, E.G.W., et al., *The value of signs and symptoms in differentiating between bacterial, viral and mixed aetiology in patients with community-acquired pneumonia.* J Med Microbiol, 2014. **63**(Pt 3): p. 441-452.
23. Loubet, P., et al., *Community-acquired pneumonia in the emergency department: an algorithm to facilitate diagnosis and guide chest CT scan indication.* Clin Microbiol Infect, 2020. **26**(3): p. 382.e1-382.e7.
24. Beekman, R., et al., *Validating a clinical prediction score for Legionella-related community acquired pneumonia.* BMC Infect Dis, 2022. **22**(1): p. 442.
25. den Engelsen, C., et al., *Infectious diseases and the use of antibiotics in outpatients at the emergency department of the University Hospital of León, Nicaragua.* Int J Infect Dis, 2009. **13**(3): p. 349-54.
26. Mandell, L.A., *Community-acquired pneumonia: An overview.* Postgrad Med, 2015. **127**(6): p. 607-15.
27. Takase, R., et al., *Clinical Manifestations of Patients with Influenza Differ by Age : A Prospective, Multi-centered Study in the Setouchi Marine Area.* Acta Med Okayama, 2021. **75**(5): p. 567-574.
28. Akhtar, A., et al., *Respiratory-tract infections among geriatrics: prevalence and factors associated with the treatment outcomes.* Therapeutic advances in respiratory disease, 2021. **15**: p. 1753466620971141.
29. Sundhedsstyrelsen. *Sundhedsstyrelsens udmeldinger om alkohol*. 2022 [cited 2022 December 06]; Available from: https://www.sst.dk/da/Viden/Forebyggelse/Alkohol/Alkoholforebyggelse/Sundhedsstyrelsens-udmeldinger-om-alkohol
30. (WHO), W.H.O. [cited 2022 December 6]; Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://apps.who.int/iris/bitstream/handle/10665/337001/9789240014886-eng.pdf
31. Shang, J., et al., *Risk factors for infection in home health care: Analysis of national Outcome and Assessment Information Set data.* Res Nurs Health, 2020. **43**(4): p. 373-386.
32. Guidet, B., et al., *Caring for the critically ill patients over 80: a narrative review.* Ann Intensive Care, 2018. **8**(1): p. 114.
33. Steffens, C., et al., *The Association Between Prescribed Opioid Receipt and Community-Acquired Pneumonia in Adults: a Systematic Review and Meta-analysis.* J Gen Intern Med, 2020. **35**(11): p. 3315-3322.
34. Walters, J.A., et al., *Pneumococcal vaccines for preventing pneumonia in chronic obstructive pulmonary disease.* Cochrane Database Syst Rev, 2017. **1**(1): p. Cd001390.
35. Kraicer-Melamed, H., S. O'Donnell, and C. Quach, *The effectiveness of pneumococcal polysaccharide vaccine 23 (PPV23) in the general population of 50 years of age and older: A systematic review and meta-analysis.* Vaccine, 2016. **34**(13): p. 1540-1550.
36. Liang, C.Y., et al., *Effectiveness of influenza vaccination in the elderly: a population-based case-crossover study.* BMJ Open, 2022. **12**(2): p. e050594.
37. Chalmers, J.D., et al., *Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis.* Thorax, 2010. **65**(10): p. 878-83.
38. Lim, W.S., et al., *BTS guidelines for the management of community acquired pneumonia in adults: update 2009.* Thorax, 2009. **64 Suppl 3**: p. iii1-55.
39. Rosenvinge, F.S. *Antibiotikavejledning for Region Syddanmark*. 06.10.2021 [cited 2022 22 september]; Available from: https://ekstern.infonet.regionsyddanmark.dk/Files/Dokument547684.htm.
40. Plesner, L.L., et al., *The formation and design of the TRIAGE study-baseline data on 6005 consecutive patients admitted to hospital from the emergency department.* Scandinavian journal of trauma, resuscitation and emergency medicine, 2015. **23**(1): p. 1-9.

41. *User Manuel Danish Emergency Process Triage*.
42. Nordberg, M., S. Lethvall, and M. Castrén, *The validity of the triage system ADAPT.* Scandinavian journal of trauma, resuscitation and emergency medicine, 2010. **18**: p. 1-1.
43. Farrohknia, N., et al., *Emergency department triage scales and their components: a systematic review of the scientific evidence.* Scand J Trauma Resusc Emerg Med, 2011. **19**: p. 42.
44. Jones, M., *NEWSDIG: The National Early Warning Score Development and Implementation Group.* Clin Med (Lond), 2012. **12**(6): p. 501-3.
45. Htun, T.P., et al., *Clinical features for diagnosis of pneumonia among adults in primary care setting: A systematic and meta-review.* Sci Rep, 2019. **9**(1): p. 7600.
46. Gong, L., et al., *Clinical profile analysis and nomogram for predicting in-hospital mortality among elderly severe community-acquired pneumonia patients with comorbid cardiovascular disease: a retrospective cohort study.* BMC Pulm Med, 2022. **22**(1): p. 312.
47. Sakakibara, T., et al., *A prediction rule for severe adverse events in all inpatients with community-acquired pneumonia: a multicenter observational study.* BMC pulmonary medicine, 2022. **22**(1): p. 34.
48. Mogensen, C.B., et al., *Ear measurement of temperature is only useful for screening for fever in an adult emergency department.* BMC Emerg Med, 2018. **18**(1): p. 51.
49. Mackowiak, P.A., F.A. Chervenak, and A. Grünebaum, *Defining Fever.* Open Forum Infect Dis, 2021. **8**(6): p. ofab161.
50. Waterer, G.W., L.A. Kessler, and R.G. Wunderink, *Medium-term survival after hospitalization with community-acquired pneumonia.* Am J Respir Crit Care Med, 2004. **169**(8): p. 910-4.
51. Zhao, L.H., J. Chen, and R.X. Zhu, *The relationship between frailty and community-acquired pneumonia in older patients.* Aging Clin Exp Res, 2023. **35**(2): p. 349-355.
52. Kitazawa, T., et al., *Characteristics of pneumonia with negative chest radiography in cases confirmed by computed tomography.* J Community Hosp Intern Med Perspect, 2020. **10**(1): p. 19-24.
53. Huang, Y., et al., *Diagnostic value of blood parameters for community-acquired pneumonia.* Int Immunopharmacol, 2018. **64**: p. 10-15.
54. Alzoubi, O. and A. Khanfar, *Association between neutrophil to lymphocyte ratio and mortality among community acquired pneumonia patients: a meta-analysis.* Monaldi Arch Chest Dis, 2021. **92**(3).
55. Milas, G.P., V. Issaris, and V. Papavasileiou, *Blood urea nitrogen to albumin ratio as a predictive factor for pneumonia: A meta-analysis.* Respir Med Res, 2022. **81**: p. 100886.
56. Kassaw, G., et al., *Outcomes and Predictors of Severe Community-acquired Pneumonia Among Adults Admitted to the University of Gondar Comprehensive Specialized Hospital: A Prospective Follow-up Study.* Infect Drug Resist, 2023. **16**: p. 619-635.
57. Adnan, M., et al., *Prognostic value of five serum markers predicting in-hospital mortality among adults with community acquired pneumonia.* J Infect Dev Ctries, 2022. **16**(1): p. 166-172.
58. Rendón-Ramirez, E.J., et al., *TGF-β Blood Levels Distinguish Between Influenza A (H1N1)pdm09 Virus Sepsis and Sepsis due to Other Forms of Community-Acquired Pneumonia.* Viral Immunol, 2015. **28**(5): p. 248-54.
59. Watanabe, H., et al., *Clinical factors associated with negative urinary antigen tests implemented for the diagnosis of community-acquired pneumococcal pneumonia in adult patients.* Med Princ Pract, 2015. **24**(2): p. 189-94.
60. Zeng, W., et al., *Association of admission blood glucose level and clinical outcomes in elderly community-acquired pneumonia patients with or without diabetes.* Clin Respir J, 2022. **16**(8): p. 562-571.
61. Barmanray, R.D., et al., *In-hospital hyperglycemia but not diabetes mellitus alone is associated with increased in-hospital mortality in community-acquired pneumonia (CAP): a systematic review and meta-analysis of observational studies prior to COVID-19.* BMJ Open Diabetes Res Care, 2022. **10**(4).

62. van der Meer, V., et al., *Diagnostic value of C reactive protein in infections of the lower respiratory tract: systematic review.* Bmj, 2005. **331**(7507): p. 26.
63. van Vugt, S.F., et al., *Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study.* Bmj, 2013. **346**: p. f2450.
64. Ebell, M.H., et al., *Accuracy of Biomarkers for the Diagnosis of Adult Community-acquired Pneumonia: A Meta-analysis.* Acad Emerg Med, 2020. **27**(3): p. 195-206.
65. Ebell, M.H., et al., *Accuracy of Signs and Symptoms for the Diagnosis of Community-acquired Pneumonia: A Meta-analysis.* Acad Emerg Med, 2020. **27**(7): p. 541-553.
66. Division of Nutrition, P.A., and Obesity, National Center for Chronic Disease Prevention and Health Promotion. 3 June 2022 [cited 2023 2 March]; Available from: https://www.cdc.gov/obesity/basics/adult-defining.html.

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	4
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	n/a
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	6+77 + additional file (table S1 and S2)
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	7
Sample size	8	D;V	Explain how the study size was arrived at.	8
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8
	10c	V	For validation, describe how the predictions were calculated.	9
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	8
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n/a
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n/a
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	9
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10-12 (Table 1, Table 2) + additional file (table S2)
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	additional file (table S3)
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	10-14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	10-14(Table 1) + table 2
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	14
	15b	D	Explain how to the use the prediction model.	14 + additional file (formula S6 +S7)
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	13
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	14
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	17
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n/a
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	15+16+ 17
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	15+16+17+18
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	5
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	19

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.