BMJ Open

Diagnostic utility of appetite loss in addition to existing prediction models for community-acquired pneumonia in the elderly: a prospective diagnostic study in acute care hospitals in Japan

Toshihiko Takada,[1,2] Yosuke Yamamoto,[2] Kazuhiko Terada,[3] Mitsuyasu Ohta,[4] Wakako Mikami,[5] Hajime Yokota,[6] Michio Hayashi,[1] Jun Miyashita,[1,2] Teruhisa Azuma,[1] Shingo Fukuma,[1,2] Shunichi Fukuhara[1,2]

To cite: Takada T, Yamamoto Y, Terada K, *et al*. Diagnostic utility of appetite loss in addition to existing prediction models for community-acquired pneumonia in the elderly: a prospective diagnostic study in acute care hospitals in Japan. *BMJ Open* 2017;**7**:e019155. doi:10.1136/bmjopen-2017-019155

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-019155).

Received 14 August 2017
Revised 26 September 2017
Accepted 11 October 2017

For numbered affiliations see end of article.

Correspondence to
Dr Toshihiko Takada;
ttakada@water.ocn.ne.jp

ABSTRACT

Objective Diagnosis of community-acquired pneumonia (CAP) in the elderly is often delayed because of atypical presentation and non-specific symptoms, such as appetite loss, falls and disturbance in consciousness. The aim of this study was to investigate the external validity of existing prediction models and the added value of the non-specific symptoms for the diagnosis of CAP in elderly patients.

Design Prospective cohort study.

Setting General medicine departments of three teaching hospitals in Japan.

Participants A total of 109 elderly patients who consulted for upper respiratory symptoms between 1 October 2014 and 30 September 2016.

Main outcome measures The reference standard for CAP was chest radiograph evaluated by two certified radiologists. The existing models were externally validated for diagnostic performance by calibration plot and discrimination. To evaluate the additional value of the non-specific symptoms to the existing prediction models, we developed an extended logistic regression model. Calibration, discrimination, category-free net reclassification improvement (NRI) and decision curve analysis (DCA) were investigated in the extended model.

Results Among the existing models, the model by van Vugt demonstrated the best performance, with an area under the curve of 0.75(95% CI 0.63 to 0.88); calibration plot showed good fit despite a significant Hosmer-Lemeshow test (p=0.017). Among the non-specific symptoms, appetite loss had positive likelihood ratio of 3.2 (2.0–5.3), negative likelihood ratio of 0.4 (0.2–0.7) and OR of 7.7 (3.0–19.7). Addition of appetite loss to the model by van Vugt led to improved calibration at p=0.48, NRI of 0.53 (p=0.019) and higher net benefit by DCA.

Conclusions Information on appetite loss improved the performance of an existing model for the diagnosis of CAP in the elderly.

Strengths and limitations of this study

- This was the first study to investigate the external validity of existing prediction models for the diagnosis of community-acquired pneumonia in the elderly.
- To evaluate the additional value of appetite loss to existing prediction models, we developed an extended logistic regression model, which was evaluated by net reclassification improvement and decision curve analysis.
- To explore the external validity of our findings, a similar investigation on a larger sample size in a primary care setting should be further conducted.

INTRODUCTION

In elderly patients, community-acquired pneumonia (CAP) could be more severe than that in younger patients and could result in poor prognosis when treatment is delayed.[1] Therefore, several predictors of mortality in elderly patients with CAP have been reported to identify individuals at risk for poor outcomes and to initiate early intervention.[2] However, in the elderly, the diagnosis of CAP is often delayed because of their atypical presentation and underlying comorbidities.[1 3] In particular, the common symptoms of cough, sputum production, fever, chills, rigours and chest pain could be absent and be replaced by non-specific deterioration of their general condition, which can manifest as appetite loss, falls, consciousness disturbance, and so on.[1] Likewise, physical examination is less reliable in the elderly than in younger individuals.[4] For example, asymptomatic elderly patients may have chest examination findings of crackles, which could be age related.[5] In addition to such challenges in the diagnosis of CAP, routine use of chest X-ray (CXR) for all elderly patients with respiratory symptoms is time consuming and might not be cost-effective.

Although several prediction models based on signs and symptoms are available for the diagnosis of CAP,[6–12] none had been developed or validated specifically for elderly patients. Because of the atypical presentation of CAP in the elderly, the performance of these models could be poor. We hypothesised that addition of the non-specific symptoms that are characteristic of CAP in elderly patients may improve the diagnostic performance of these existing models.

Accordingly, we investigated the external validity of these existing models and evaluated the value of adding the non-specific symptoms for the diagnosis of CAP in elderly patients.

MATERIALS AND METHODS

This was a prospective observational study that followed the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) statement for prediction model studies.[13] Written informed consent was obtained from all the subjects.

Setting

This study was conducted between 1 October 2014 and 30 September 2016 at the general medicine departments of three teaching hospitals: Shirakawa Kosei General Hospital in Fukushima, Japan (471-bed capacity); Kimitsu Chuo Hospital in Chiba, Japan (661-bed capacity); and Ashigarakami Hospital in Kanagawa, Japan (296-bed capacity).

Patients

We included ≥65-year-old outpatients who presented with a chief complaint of cough or sputum production. Exclusion criteria were: (1) longer than 1-month duration of cough or sputum production; (2) current intake of antibiotics; (3) did not reside in 'community' (patients who resided in nursing homes or those transferred from another hospital were excluded because epidemiology has been reported to differ among CAP, healthcare-associated pneumonia and hospital-acquired pneumonia).[14]

Data collection

Using a structured collection form, physicians collected data on examination of eligible patients. When patients themselves were unable to answer the questions, physicians collected data from their caregivers. The items in the form included the predictors in existing models: (1) history (cough, sputum production, sore throat, coryza, dyspnoea, fever, chills, night sweats, myalgia, pleurisy, diarrhoea, duration of symptoms, and history of bronchial asthma, chronic obstructive pulmonary disease, chronic heart failure or ischaemic heart disease); and (2) signs (blood pressure, pulse rate, oxygen saturation, respiratory rate, body temperature, diminished breath sounds, wheezes, crackles and rhonchi).[6–11] We added items on appetite loss, falls and consciousness disturbance. Appetite loss was based on self-assessment of 24-hour food intake in proportion to the usual intake.

Outcome measurement

The reference standard for CAP diagnosis was CXR in the posteroanterior and lateral views or in the anteroposterior view for patients who could not stand up. Two certified radiologists (WM and HY) who were blinded to the history and physical examination data independently assessed the CXRs. Any disagreement was resolved by discussion.

Existing prediction models for CAP

Existing models were identified based on a previous external validation study on the prediction models for CAP in the primary care setting.[12] Inclusion criteria for these existing models were (1) use of logistic regression for the diagnosis of CAP in adult patients; (2) predictors were items about history and physical examination; (3) the diagnosis of CAP was made by CXR or CT.

Statistical analysis

The existing models in their original form and without adjustment of intercept and coefficients were externally validated for diagnostic performance by calibration plot and discrimination. The calibration plots were tested with the Hosmer-Lemeshow (HL) test and a p Value <0.05 indicated a lack of good fit.[13] Discrimination was assessed by area under the curve (AUC).

The diagnostic performance of each item in the existing models, as well as that of the non-specific symptoms (appetite loss, falls and consciousness disturbance), was assessed in terms of sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR−), and diagnostic odds ratio (DOR), all of which were reported as point estimates with 95% CI. DOR ranged from zero to infinity and was calculated as the ratio of LR+ to LR−, with higher values indicative of better discriminative performance; a value of 1 indicated that the test did not discriminate between people with and without the disease.[15] The cut-off points for continuous variables were determined according to existing prediction models[16] or the values with the best sensitivity and specificity in the receiver operating characteristic (ROC) curve. The diagnostic performance of the logistic regression model that included appetite loss as the sole predictor was evaluated by calibration plot.

To evaluate the additional value of these non-specific symptoms to the existing prediction models, we developed an extended logistic regression model without changing the coefficients of the existing models. Calibration, discrimination, category-free net reclassification improvement (NRI) and decision curve analysis (DCA) were applied on the extended model. The NRI was the cumulative net proportion of events reclassified correctly plus the net proportion of non-events reclassified correctly.[17] The DCA was a graphical approach to evaluate the prediction models based on the principle that the relative harm of false positives and false negatives can be expressed in terms of a probability threshold. Net benefit was obtained

Table 1 Patients' characteristics The first line "Age (mea±SD) 76.7±7.8 should be written in normal letters, not bold.

Age (mean±SD)	76.7±7.8
Male (%)	64 (58.7)
History of BA/COPD (%)	6 (5.5)
History of IHD/CHF (%)	14 (12.9)
Clinical diagnosis (%)	
CAP	32 (29.4)
Upper respiratory infection	61 (56.0)
Sinusitis	9 (8.2)
Others	7 (6.4)
CAP diagnosed by CXR (%)	24 (22.0)

BA, bronchial asthma; CAP, community-acquired pneumonia; CHF, chronic heart failure; COPD, chronic obstructive pulmonary disease; CXR, chest X-ray; IHD, ischaemic heart disease.

by subtracting the proportion of false positives from the proportion of true positives, weighed by the relative harm of false positive and false negative results.[18] Net benefit was assessed by the probability threshold, which was decided on 10 physicians who worked in the participating hospitals and who were blinded to the results of our research.

To assess the agreement in the CXR interpretations of the two radiologists, Cohen's kappa coefficient (κ) was calculated.[19] Imputation method was not applied for missing values because of the possibility of perfect prediction of categorical data. We did not estimate sample size a priori because there was no available consensus on the adequate sample size among the external validation studies on the prediction models or evaluation studies on the diagnostic performance of the added predictors.[13] Statistical analyses were performed with a commercial software program (STATA, V.14.2 SE; StataCorp, College Station, Texas).

RESULTS

Patient characteristics

A total of 109 patients were included in the study population. The baseline characteristics of the patients are listed in table 1; mean age was 76.7 years (SD 7.8) and 58.7% were men. CAP was diagnosed by CXR in 22% of the patients. There was moderate agreement (κ=0.58) between the two radiologists who read the CXRs.

External validation of existing models

Six existing prediction models for CAP diagnosis were externally validated in our cohort. The formula for each model is presented in the online supplementary table 1.

The diagnostic performance of each predictor item in the history and physical examination is presented in table 2. For each model, a calibration plot was drawn by comparing the predicted probability with the observed probability in our data set (figure 1). All existing models demonstrated significant p Values in the HL test, indicating poor fit (table 3). On the other hand, the calibration plot for the model by van Vugt *et al*[11] visually demonstrated good fit between the predicted probability and the observed prevalence of CAP. Discrimination was highest for the model by Singal *et al*[7] (AUC=0.76; 95% CI 0.65 to 0.87), followed by the models by van Vugt *et al*[11] (AUC=0.75; 95% CI 0.63 to 0.88); Diehr *et al*[6] (AUC=0.75; 95% CI 0.64 to 0.86); and Heckerling *et al*[8] (AUC=0.75; 95% CI 0.64 to 0.86) (table 3).

Diagnostic performance of the non-specific symptoms

The sensitivity, specificity, LR+ and LR− for appetite loss, falls and consciousness disturbance are listed in table 2. The cut-off point for appetite loss was determined at 50% by the ROC curve. The AUC of appetite loss for CAP diagnosis was 0.76 (95% CI 0.65 to 0.87). The calibration plot of the model for appetite loss showed better fit than the model by van Vugt *et al*,[11] which showed the best calibration among the existing models (figure 2). A non-significant HL test suggested good fit of the model for appetite loss (table 3).

Added value of the non-specific symptoms

In the evaluation of the diagnostic performance of each non-specific symptom, falls and consciousness disturbance seemed less useful; therefore, we investigated the added value of appetite loss to the model by van Vugt *et al*,[11] which showed the best performance among the existing models.

The formula for the extended model included the following: −5.258, −0.135 for appetite, +0.446 for dyspnoea, +0.698 for absence of runny nose, +0.596 for diminished vesicular breathing, +1.404 for crackles, +0.961 for tachycardia and +0.980 for temperature >37.8°C. The extended model showed improved performance on the calibration plot, especially in the subgroup with high probability of CAP (figure 2), and good fit by the HL test (p=0.48; χ^2=8.6). Compared with the original model, the extended model tended to have an improved AUC, but this change was not significant (0.77 (95% CI 0.65 to 0.89) vs 0.75 (95% CI 0.63 to 0.88); p=0.48). The NRI of the extended model was 0.53 (95% CI 0.08 to 0.97; p=0.019), which meant that 53% of patients correctly obtained increased probability of CAP.

Figure 3 illustrates the decision curves for the model by van Vugt *et al*[11] and the extended model. Among the 10 physicians, the mean threshold for ordering CXR for the diagnosis of CAP in the elderly was 22.5% (SD 7.1%; range 10%–30%). At a threshold of 10%–30%, a higher net benefit was more frequently observed in the extended model than in the original model. At a threshold of 20%, the net benefit of the extended model increased by 0.02 compared with the original model.

Table 2 Diagnostic performance of the predictors of CAP in the elderly

	CAP+ (n=27)	CAP− (n=82)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Cut-off point
Cough (%)	100	98.8	100 (87.2 to 100)	1.2 (0.3 to 6.6)	1.0 (1.0 to 1.0)	0	–	–
Sputum (%)	88.9	74.1	88.9 (70.8 to 97.6)	25.9 (16.8 to 36.9)	1.2 (1.0 to 1.4)	0.4 (0.1 to 1.3)	2.8 (0.8 to 9.6)	–
Sore throat (%)	22.2	40.2	22.2 (8.6 to 42.3)	59.8 (48.3 to 70.4)	0.6 (0.3 to 1.2)	1.3 (1.0 to 1.7)	0.4 (0.2 to 1.1)	–
Rhinorrhea (%)	34.6	60.5	34.6 (17.2 to 55.7)	39.5 (28.8 to 51.0)	0.6 (0.3 to 1.0)	1.7 (1.1 to 2.4)	0.3 (0.1 to 0.9)	–
Dyspnoea (%)	40.7	34.2	40.1 (22.4 to 61.2)	65.9 (54.6 to 76.0)	1.2 (0.7 to 2.1)	0.9 (0.6 to 1.3)	1.3 (0.6 to 3.2)	–
Fever (%)	74.1	35.4	74.1 (53.7 to 88.9)	64.6 (53.3 to 74.9)	2.1 (1.5 to 3.0)	0.4 (0.2 to 0.8)	5.2 (2.0 to 13.5)	–
Chill (%)	29.6	27.2	29.6 (13.8 to 50.2)	71.6 (60.5 to 81.1)	1.0 (0.5 to 2.1)	0.98 (0.7 to 1.3)	1.1 (0.4 to 2.7)	–
Night sweats (%)	33.3	24.4	33.3 (16.5 to 54.0)	75.6 (64.9 to 84.4)	1.4 (0.7 to 2.6)	0.9 (0.7 to 1.2)	1.5 (0.6 to 3.9)	–
Myalgia (%)	7.4	18.3	7.4 (0.9 to 24.3)	81.7 (71.6 to 89.4)	0.4 (0.1 to 1.7)	1.1 (1.0 to 1.3)	0.4 (0 to 1.5)	–
Pleural pain (%)	7.4	7.4	7.4 (0.9 to 24.3)	92.6 (84.6 to 97.2)	1.0 (0.2 to 4.7)	1.0 (0.9 to 1.1)	1.0 (0 to 4.7)	–
Diarrhoea (%)	7.4	7.4	7.4 (0.9 to 24.3)	92.6 (84.6 to 97.2)	1.0 (0.2 to 4.7)	1.0 (0.9 to 1.1)	1.0 (0 to 4.7)	–
Duration of symptoms (days)	6.1	7.2	42.3 (23.4 to 63.1)	62.5 (51.0 to 73.1)	1.1 (0.7 to 1.9)	0.9 (0.6 to 1.3)	1.2 (0.5 to 3.0)	7 days from the model by Melbye *et al*[9]
Coma (%)	15.4	6.1	15.4 (4.4 to 34.9)	93.9 (86.3 to 98.0)	2.5 (0.7 to 8.7)	0.9 (0.8 to 1.1)	2.8 (0.7 to 10.6)	–
Fall (%)	11.1	1.2	11.1 (2.4 to 29.2)	98.8 (93.4 to 100.0)	9.1 (1.0 to 84.0)	0.9 (0.8 to 1.0)	10.1 (1.4 to ∞)	–
Appetite loss (%)	48.8	80.1	66.7 (46.0 to 83.5)	79.3 (68.9 to 87.4)	3.2 (2.0 to 5.3)	0.4 (0.2 to 0.7)	7.7 (3.0 to 19.7)	50% determined by the ROC curve
SBP (mm Hg)	129.9	130.1	3.7 (0.1 to 19.0)	100 (95.6 to 100)	–	1.0 (0.9 to 1.0)	–	90 mm Hg from CURB65[16]
DBP (mm Hg)	72.5	73.9	22.2 (8.6 to 42.3)	84.1 (74.4 to 91.3)	1.4 (0.6 to 3.3)	0.9 (0.7 to 1.2)	1.4 (0.5 to 4.4)	60 mm Hg from CURB65[16]
Pulse rate (/min)	89.3	82.8	29.6 (13.8 to 50.2)	80.5 (70.3 to 88.4)	1.5 (0.7 to 3.2)	0.9 (0.7 to 1.1)	1.7 (0.7 to 4.6)	100/min from the model by van Vugt *et al*[11]
SpO_2 (%)	93.7	96.4	51.9 (31.9 to 71.3)	91.5 (83.2 to 96.5)	6.1 (2.7 to 13.5)	0.5 (0.4 to 0.8)	11.5 (4.0 to 33.4)	95% as normal value
Respiratory rate (/min)	21.3	19.0	3.7 (0.1 to 19.0)	100 (95.5 to 100)	–	1.0 (0.9 to 1.0)	–	30/min from CURB65[16]
Body temperature (□)	37.6	36.7	29.6 (13.8 to 50.2)	91.5 (83.2 to 96.5)	3.5 (1.4 to 8.7)	0.8 (0.6 to 1.0)	4.5 (1.5 to 13.6)	38°C from the model by Hopstaken *et al*[10]
Decreased respiratory sound (%)	18.5	4.9	18.5 (6.3 to 38.1)	95.1 (87.8 to 98.6)	3.8 (1.1 to 13.0)	0.9 (0.7 to 1.0)	4.4 (1.2 to 16.4)	–
Crackle (%)	48.2	14.8	48.1 (28.7 to 68.1)	85.2 (75.6 to 92.1)	3.3 (1.7 to 6.2)	0.6 (0.4 to 0.9)	5.3 (2.1 to 14.0)	–
Wheeze (%)	3.7	3.7	3.7 (0.1 to 19.0)	96.3 (89.6 to 99.2)	1.0 (0.1 to 9.2)	1.0 (0.9 to 1.1)	1.0 (0 to 7.4)	–
Rhonchi (%)	3.7	6.2	3.7 (0.1 to 19.0)	93.8 (86.2 to 98.0)	0.6 (0.1 to 4.9)	1.0 (0.9 to 1.1)	0.6 (0 to 4.0)	–

CAP, community-acquired pneumonia; CURB65, confusion, urea >7 mmol/L, respiratory rate ≥30 breaths/min, low blood pressure (systolic <90 mm Hg or diastolic ≤60 mm Hg); DBP, diastolic blood pressure; DOR, diagnostic OR; LR+, positive likelihood ratio; LR−, negative likelihood ratio; ROC, receiver operating characteristic; SBP, systolic blood pressure; SpO_2, oxygen saturation.

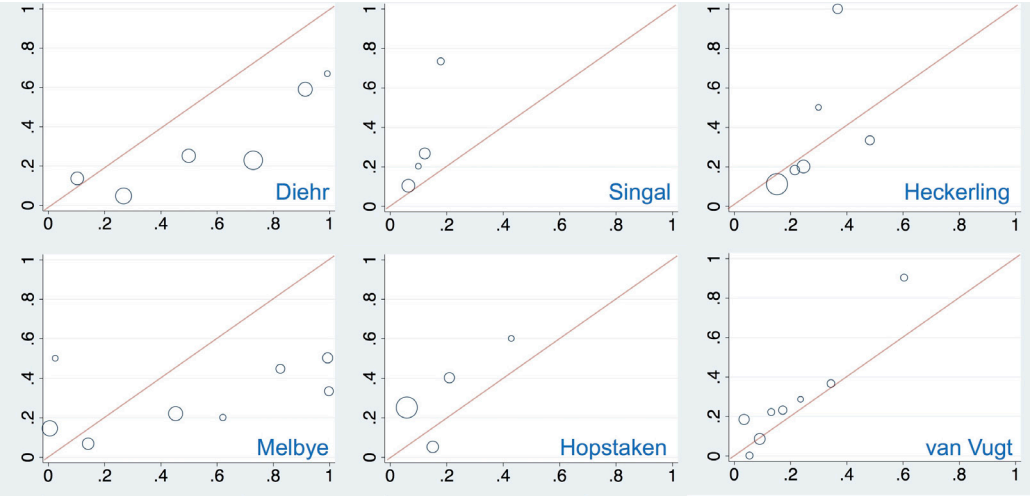

Figure 1 Calibration plots of existing models. The x-axis represents the predicted proportion and the y-axis represents the observed proportion, with a reference for perfect calibration (red line). The size of the circle indicates the number of patients. The model by van Vugt visually demonstrated better fit than the other models.

DISCUSSION

Our study assessed the external validity of existing models for the diagnosis of CAP in an elderly population at acute care hospitals in Japan. These existing models had been validated in a meta-analysis of individual data of patients who were at least 18 years old.[12] Although the existing models were supposed to have poor external validity in the elderly population, the model by van Vugt *et al*[11] showed relatively good calibration and discrimination. In the evaluation of the diagnostic performance of the non-specific symptoms of CAP in the elderly, falls and consciousness disturbance showed high specificity but low sensitivity, leading to clinically irrelevant LR+ and LR−. On the other hand, information about appetite showed good performance for the diagnosis of CAP in the elderly. Nurse-assessed food consumption has been reported to predict bacteraemia with a sensitivity of 0.92 and negative predictive value of 0.98.[20] Therefore, appetite loss could be an important predictor of infection. In the assessment of the added value of appetite loss to the model by van Vugt *et al*,[11] calibration was improved, especially in the population with high probability; however, the improvement in discrimination was not significant. NRI was significant as more than 50% of the patients were correctly assessed to have increased probabilities. Results of the DCA showed higher net benefit of the extended model at a threshold appetite loss of 10%–30%. For example, applying the extended model at a threshold appetite loss of 20% could correctly detect CAP in two additional patients per 100 patients.

Table 3 Comparison of diagnostic performance among existing models and appetite loss

	Calibration (HL test)		Discrimination	
Model	χ^2	**p Value**	**AUC**	**95% CI**
Diehr *et al*[6]	134.9	<0.001	0.75	0.64 to 0.86
Singal *et al*[7]	39.2	<0.001	0.76	0.65 to 0.87
Heckerling *et al*[8]	19.8	0.003	0.75	0.64 to 0.86
Melbye *et al*[9]	6104.2	<0.001	0.64	0.51 to 0.77
Hopstaken *et al*[10]	48.2	<0.001	0.54	0.42 to 0.66
van Vugt *et al*[11]	18.7	0.017	0.75	0.63 to 0.88
Appetite loss	1.91	0.928	0.76	0.65 to 0.87

AUC, area under the curve; HL, Hosmer-Lemeshow.

The strength of our study was that this was the first study that evaluated the external validity of existing models for CAP in the elderly. Based on our results, the model by van Vugt *et al*[11] can be applied in the elderly population despite its relatively poor calibration. To the best of our knowledge, this was the first study to investigate and identify the value of information on appetite loss, independently and in addition to the model by van Vugt *et al*,[11] for the diagnosis of CAP in the elderly. Because routine use of CXR might not be time efficient or cost-effective, estimating the diagnostic probability using information on appetite loss could be an effective strategy in elderly patients with respiratory symptoms. The current study was conducted in hospital settings where CXR was easily available, but three of the six existing models were derived in the hospital setting.[6–8] Therefore, we believe that there was a need to correctly estimate the diagnostic probability of CAP in both hospital and primary care settings.

This study had several limitations. First, our sample size was relatively small. Because this study was conducted in acute care hospitals, many patients were ineligible because they were already prescribed antibiotics in the primary care setting. Therefore, a similar investigation on a larger sample size in a primary care setting should be further explored. Second, the frequency of CAP was relatively higher than that previously reported,[12] probably because the number of patients with CAP was higher in

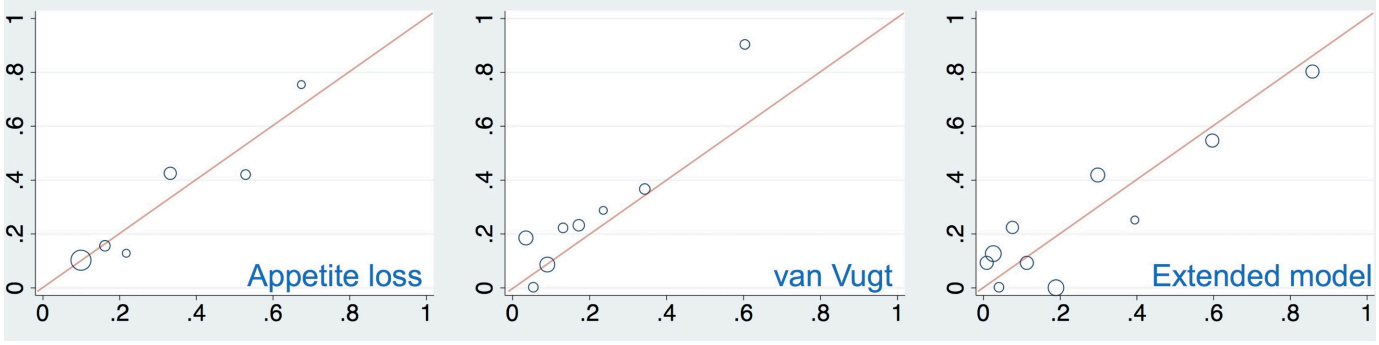

Figure 2 Comparison of the calibration plots among the model of appetite loss, the model by van Vugt and the extended model. The extended model comprised the model by van Vugt with additional information on appetite loss.

the acute care setting than in the primary care setting. Therefore, the external validity of our results should be applied carefully in the primary care setting. Healthcare in Japan is a free-access system that allows people to be examined and treated at the medical institutions (ie, clinics, secondary hospitals and university hospitals) of their choice.[21] Therefore, the differences in patient characteristics between the hospital and primary care settings might be less prominent in Japan than those in other countries where patients have to see their primary care physicians first. Third, although CT has been reported to be superior in terms of diagnostic accuracy,[22] CXR was used as the reference standard of CAP in this study because performing CT on all patients was not feasible and ethical due to difficult access, higher radiation exposure and cost.

In conclusion, the model by van Vugt *et al*[11] demonstrated relatively high performance among the existing models for the diagnosis of CAP in the elderly. Information about appetite loss independently demonstrated diagnostic utility for CAP in the elderly and improved the performance of the model by van Vugt *et al*.[11]

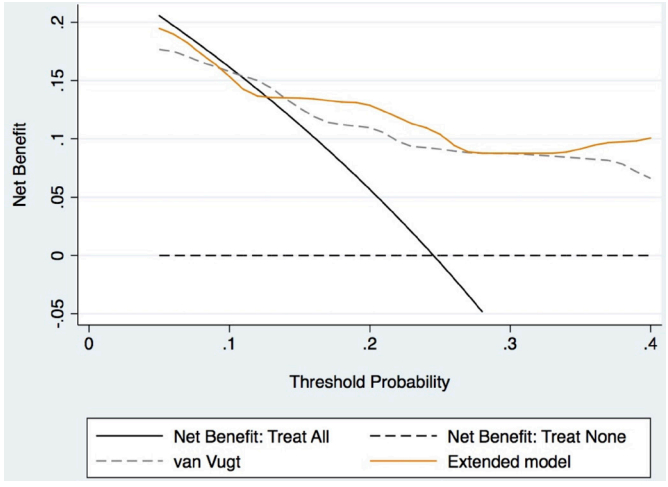

Figure 3 Decision curve for the model by van Vugt and the extended model. The net benefits for the strategy 'CXR for all patients' (black line), for the strategy 'No CXR' (black dashed line), for the model by van Vugt (grey dashed line) and for the extended model (yellow line) are shown. At a threshold of 10%–30%, the extended model demonstrated higher net benefit than the original model.

Author affiliations
[1]Department of General Medicine, Shirakawa Satellite for Teaching and Research (STAR), Fukushima Medical University, Fukushima, Japan
[2]Department of Healthcare Epidemiology, School of Public Health in the Graduate School of Medicine, Kyoto University, Kyoto, Japan
[3]Department of General Medicine, Kimitsu Chuo Hospital, Kisarazu, Japan
[4]Department of Internal Medicine, Ashigarakami Hospital, Kanagawa, Japan
[5]Department of Radiology, Keiyu Hospital, Kanagawa, Japan
[6]Department of Radiology, Chiba University Hospital, Chiba, Japan

Contributors TT had full access to all data in the study; he takes responsibility for the integrity of the data and accuracy of the data analysis and wrote the first draft. YY designed the study, interpreted the data and drafted the paper. KT, MO, WM and HY collected and interpreted the data. MH, JM and TA designed the study, and collected and interpreted the data. S Fukuma and S Fukuhara supervised the research and revised the work critically for important intellectual content.

Funding This work was supported by a Grant-in-Aid for Epidemiological Research of St Luke's Life Science Institute.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Kyoto University Graduate School and Faculty of Medicine, Shirakawa Kosei General Hospital, Kimitsu Chuo Hospital, and Ashigarakami Hospital.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Marrie TJ. Community-acquired pneumonia in the elderly. *Clin Infect Dis* 2000;31:1066–78.
2. Cataudella E, Giraffa CM, Di Marca S, *et al*. Neutrophil-to-lymphocyte ratio: an emerging marker predicting prognosis in elderly adults with community-acquired pneumonia. *J Am Geriatr Soc* 2017;65:1796–801.
3. Berman P, Hogan DB, Fox RA. The atypical presentation of infection in old age. *Age Ageing* 1987;16:201–7.

4. Thiem U, Heppner HJ, Pientka L. Elderly patients with community-acquired pneumonia: optimal treatment strategies. *Drugs Aging* 2011;28:519–37.
5. Kataoka H, Matsuno O. Age-related pulmonary crackles (rales) in asymptomatic cardiovascular patients. *Ann Fam Med* 2008;6:239–45.
6. Diehr P, Wood RW, Bushyhead J, *et al*. Prediction of pneumonia in outpatients with acute cough-a statistical approach. *J Chronic Dis* 1984;37:215–25.
7. Singal BM, Hedges JR, Radack KL. Decision rules and clinical prediction of pneumonia: evaluation of low-yield criteria. *Ann Emerg Med* 1989;18:13–20.
8. Heckerling PS, Tape TG, Wigton RS, *et al*. Clinical prediction rule for pulmonary infiltrates. *Ann Intern Med* 1990;113:664–70.
9. Melbye H, Straume B, Aasebø U, *et al*. Diagnosis of pneumonia in adults in general practice. relative importance of typical symptoms and abnormal chest signs evaluated against a radiographic reference standard. *Scand J Prim Health Care* 1992;10:226–33.
10. Hopstaken RM, Muris JW, Knottnerus JA, *et al*. Contributions of symptoms, signs, erythrocyte sedimentation rate, and C-reactive protein to a diagnosis of pneumonia in acute lower respiratory tract infection. *Br J Gen Pract* 2003;53:358–64.
11. van Vugt SF, Broekhuizen BD, Lammens C, *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.
12. Schierenberg A, Minnaard MC, Hopstaken RM, *et al*. External validation of prediction models for pneumonia in primary care patients with lower respiratory tract infection: an individual patient data meta-analysis. *PLoS One* 2016;11:e0149895.
13. Moons KG, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
14. Kollef MH, Shorr A, Tabak YP, *et al*. Epidemiology and outcomes of health-care-associated pneumonia: results from a large US database of culture-positive pneumonia. *Chest* 2005;128:3854–62.
15. Glas AS, Lijmer JG, Prins MH, *et al*. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003;56:1129–35.
16. Lim WS, van der Eerden MM, Laing R, *et al*. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58:377–82.
17. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11–21.
18. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
20. Komatsu T, Onda T, Murayama G, *et al*. Predicting bacteremia based on nurse-assessed food consumption at the time of blood culture. *J Hosp Med* 2012;7:702–5.
21. Toyabe S. Non-specialized inpatient care provided by university hospitals in Japan. *Risk Manag Healthc Policy* 2008;1:23–9.
22. Syrjälä H, Broas M, Suramo I, *et al*. High-resolution computed tomography for the diagnosis of community-acquired pneumonia. *Clin Infect Dis* 1998;27:358–63.