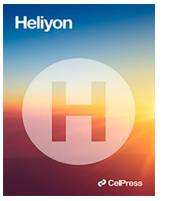
Heliyon 10 (2024) e23227

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Review article

A systematic review of tools for predicting complications in patients with influenza-like illness

Tania Marx [a,*], Nada Khelifi [a], Isabelle Xu [a], Laurie Ouellet [a], Annie Poirier [a], Benoit Huard [a], Myriam Mallet [a], Frédéric Bergeron [b], Maurice Boissinot [c], Michel G. Bergeron [c], Simon Berthelot [a,d]

[a] *Axe Santé des Populations et Pratiques Optimales en Santé, Centre de Recherche du CHU de Québec-Université Laval, Québec, Qc, Canada*
[b] *Bibliothèque-Direction des Services-conseils, Université Laval, Québec, Qc, Canada*
[c] *Centre de Recherche en Infectiologie de l'Université Laval, Axe Maladies Infectieuses et Immunitaires, Centre de Recherche du CHU de Québec-Université Laval, Québec, Qc, Canada*
[d] *Department of Family and Emergency Medicine, Université Laval, Québec, Qc, Canada*

ARTICLE INFO

Keywords:
Influenza-like illness
Risk stratification tool
Outpatient
Hospitalization
Mortality

ABSTRACT

Objective: To identify tools that predict the risk of complications for patients presenting to an outpatient clinic or an emergency department (ED) with influenza-like illness.
Methods: We searched Medline, Embase, Cochrane Library and CINAHL from inception to July 2023. We included articles reporting on the derivation or validation of a score or algorithm used to stratify the risk of hospitalization or mortality among patients with influenza-like illness in the ED or outpatient clinic.
Results: Twelve articles reporting on eight scores and six predictive models were identified. For predicting the need for hospitalization, the area under the curve (AUC) of the PMEWS and the CURB-65 ranged respectively from 0.76 to 0.94, and 0.65 to 0.88. The Community Assessment Tool had an AUC of 0.62. For predicting inpatient mortality, AUC was 0.66 for PMEWS and 0.79 for CURB-65, 0.79 for the SIRS criteria and 0.86 for the qSOFA score. Two scores were developed without external validation during the Covid-19 pandemic. The CovHos score and the Canadian Covid discharge score had an AUC ranged from 0.70 to 0.91. The predictive models performed adequately (AUC from 0.76 to 0.92) but will require external validation for clinical use. Tool diversity and study population heterogeneity precluded meta-analysis.
Conclusion: Although the CURB, PMEWS and qSOFA scores appear to predict accurately the risk of complications of influenza-like illness, none were reliable enough to justify their widespread ED use. Refinement of an existing tool or development of a new tool to optimize the management of these patients is needed.

1. Introduction

Epidemic waves of respiratory disease such as COVID-19 and seasonal influenza place a heavy burden on public healthcare resources and constitute a major health issue. In the USA, seasonal influenza is estimated to cause 9.3 million to 49 million illnesses per

* Corresponding author.
E-mail address: tania.marx@univ-fcomte.fr (T. Marx).

https://doi.org/10.1016/j.heliyon.2023.e23227
Received 25 July 2023; Received in revised form 22 November 2023; Accepted 29 November 2023
Available online 3 December 2023

year depending on virulence and vaccination [1]. Based on mathematical models, influenza-associated mortality worldwide is estimated to range typically from 4.0 to 8.8 per 100,000 individuals overall and 51.3 to 99.4 per 100,000 among persons aged 75 or more [2]. The morbidity associated with epidemic influenza translates to increased hospital admissions and use of healthcare services and hence pressure on national health systems. A disease burden analysis based on 2015 data estimated that annual influenza epidemics resulted in 3.7 million medical doctor office visits, 0.65 million emergency department visits and 247,000 hospitalizations in the USA. Even though influenza is usually a self-limiting illness, it has a major socioeconomic impact through occupation of medical resources, worker absenteeism and losses of productivity. The total economic burden of the 2015 influenza epidemic was valued about $11.2 billion US [3].

Since emergency departments feel the impact of respiratory epidemic waves directly, ED care providers must be able to stratify) these patients according to their risk of complications and thus assign the most appropriate care resources accordingly. Correct identification of patients who can be safely discharged or redirected to an outpatient clinic and those who require hospital admission is crucial from both public health and economic perspectives. Routine clinical judgement alone has been shown to be a poor predictor of disease severity, which may be overestimated or underestimated [4]. Influenza management guidelines therefore now recommend the use of a data-based severity assessment tool as an adjunct to clinical judgement [5–7].

The objective of this systematic review was to identify risk stratification tools that appear to be reliable for predicting the risk of complications in patients presenting to an outpatient clinic or an ED with influenza-like illness.

2. Methods

A systematic review was conducted in accordance with PRISMA guidelines (preferred reporting items for systematic reviews and meta-analyses) [8]. The PICOS format was specified as follows:

Population: adults or pediatric patients presenting with influenza-like illness at an ED or outpatient clinic.

Intervention: scale, algorithm, or prediction models for risk evaluation.

Comparator: none.

Outcome: hospitalization or mortality

Study design: observational studies or interventional trials evaluating the performance of risk stratification tools.

The study protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO - CRD42019138802).

2.1. Search strategy and study selection

A systematic literature search was conducted in databases (Medline, Embase, Cochrane Library, CINAHL) from their inception to July 13, 2021. The results were updated on July 20, 2023. The search strategy is provided as a supplementary file and was limited to publications in English and French (Table S1). Additional publications were identified in the bibliographies of the retrieved research articles and systematic reviews that were included in the study.

After removing duplicates, all abstracts were screened independently by four reviewers (S.B., A.P., T.M., B.H.) for eligibility using the Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Full texts of articles thus selected were then assessed for relevance by two independent reviewers (A.P. and T.M.). An article was retained for the study if it reported on the management of influenza-like illness or acute respiratory infections in an outpatient setting (e.g., ED, clinic, primary care practice, health advice line), AND on the performance of a scale, score or algorithm used to stratify patients according to their risk of complications. Conflicting opinions of A.P. and T.M. were resolved through discussion, and a third opinion (S.B.) was sought when necessary. An article was excluded if its subject matter was an untested algorithm based on expert consensus OR an investigation of composite outcome (e.g., mortality or admission to intensive care without differentiation) or of outcomes other than those initially selected OR focused exclusively on hospitalized patients OR if the study design and results did not allow tool performance evaluation. Letters to editors, case reports, conference abstracts, reviews or guidelines for influenza-like illness management were also excluded.

2.2. Data extraction and quality assessment

Two independent reviewers (N.K. and I.X.) extracted data from the selected articles using a pretested standardized form based on the checklist for critical appraisal and data extraction for systematic reviews of prediction modelling studies, aka CHARMS [9]. A third reviewer (T.M.) resolved inconsistencies or disagreements between N.K. and I.X. The following data were extracted: authors, journal, year of publication, region where study was conducted, period of data collection, study design, participant description, outcome(s) to be predicted, predictor type and method of measurement, method of handling missing data, model development or validation and performance.

Study quality was assessed using the prediction model risk of bias assessment tool (PROBAST) which rates study methodology and applicability to the review question as “high”, “low” or “unclear” risk of bias based on a set of questions and a scoring guide [10].

2.3. Statistical analysis

A descriptive synthesis was obtained based on the CHARMS. For each study, 2 × 2 tables were made with the number of true positives, true negatives, false positives, and false negatives according to the different thresholds studied for the scores found. Forest

plots for these thresholds were produced with 95% confidence intervals using Review Manager (RevMan) Version 5.4.1 (Cochrane Collaboration, 2020) to visualize the study sensitivities and specificities reported and rank the studies by sensitivity [11] to assess the heterogeneity between them and thus determine if a meta-analysis was feasible [12–14].

3. Results

A total of 6584 articles were retrieved from the databases by the systematic search. After removing 856 duplicates, 5728 articles were screened for relevance and 119 of these were selected for examination of the full text. More than 50% of the full text were excluded due to the subject matter not dealing with any tool or algorithm predicting the risk of complications, or because the population studied differed from our focus. This reduced the number to 9 articles. Fig. 1 shows the flow diagram for inclusion of articles in this study and the reasons of exclusions. Sixteen potentially relevant articles were found by searching the bibliographies of included research articles and relevant systematic reviews identified by the selection process. After full-text reading, only three articles did meet the inclusion criteria. The 12 studies retained for the present review report on eight different scores and six predictive models.

3.1. Characteristics of the included studies

The characteristics of the included studies are summarized in Table 1.

All were published between 2007 and 2022. Five were conducted in North America [16,18,20,24,26], one in South America [25], four in Europe [15,17,22,23] and two in Taiwan [19,21]. All studies were focused on emergency departments except for two: one was conducted in a clinical health setting [20] and the second with a network of general practitioners [23]. Half of the articles were multicenter and four articles were prospective observational studies [16–18,22]. External validation of scores were reported in four of the articles [15–17,19], predictive model development was reported in six [18,20,21,23–25] but only one with external validation [21].

3.2. Participants

Eight studies involved adults only [16,19,21–26]. In one case all participants were aged 15 years or older [15]. The other three included children [17,18,20]. For eight studies, only patients with a positive influenza test [18,19,21] or a positive Covid-19 test [20,22–24,26] were included. One study included patients with acute respiratory complaints [16] and another included patients for which influenza or Covid-19 was suspected, based on any of the following conditions: 1) fever; 2) influenza-like illness (two or more symptoms among cough, sore throat, rhinorrhea, limb or joint pain, headache, vomiting or diarrhea) [17]; or 3) clinical symptoms enabled the prediction of Covid-19 including cough, fever, headache, sore throat, shortness of breath [25]. One study included patients diagnosed with community-acquired pneumonia used as a proxy for pandemic influenza [15].

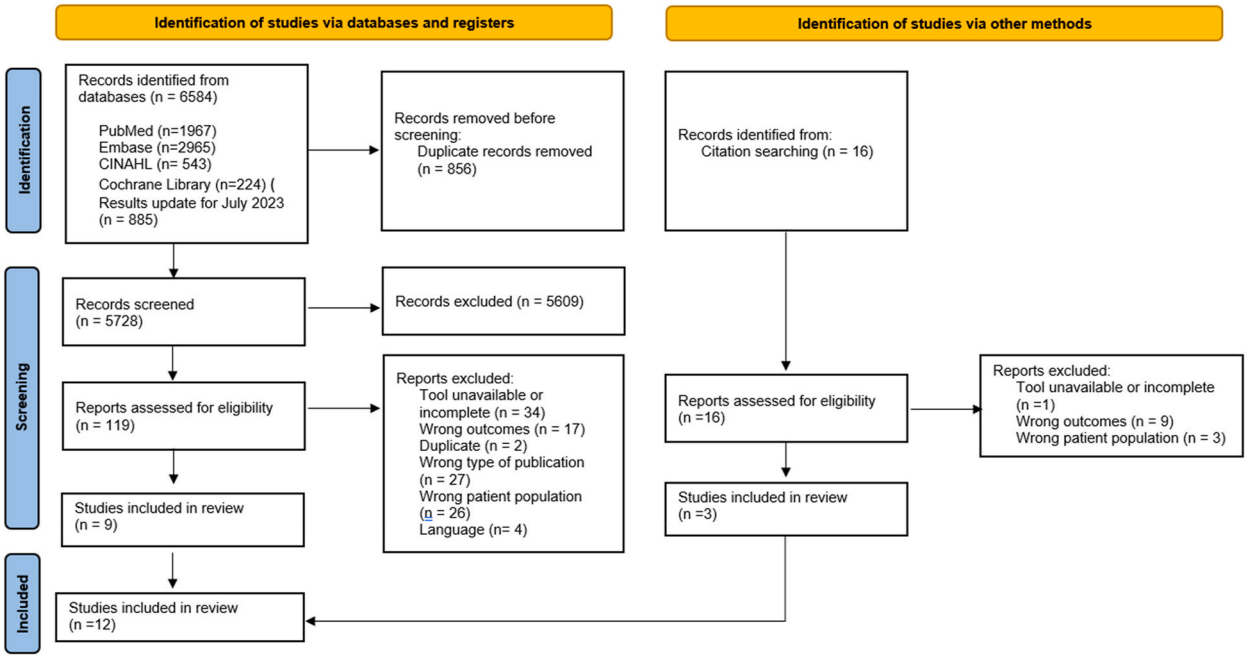


Fig. 1. PRISMA flow diagram.

Table 1
Characteristics of included studies.

Authors, year, journal	Region	Data collection period	Study design	Participants	Outcomes	Candidate predictors	Method for measurement	Handling of missing data	Sample size (N participants/N outcomes)	Model development or model evaluation	Model performance (95% CI)	
Challen et al. (2007) BMC Health Serv Res [15].	Tertiary care ED Manchester, United Kingdom	February to December 2005	Retrospective observational cohort	Adults (>15 years old) with community acquired pneumonia Age ≥65 years: 52.5%	Hospitalization Mortality	PMEWS* (10 criteria based on age, physiological data, history, social isolation, performance status of limited activity, 10 criteria) CURB-65 (5 criteria based on physiological data, age, laboratory)	Electronic medical records	37 excluded	Hospitalization: 241/164 Mortality: 186/42	External validation Hospitalization PMEWS: AUC = 0.94 (0.92, 0.97) CURB-65: AUC = 0.88 (0.84, 0.92) Mortality PMEWS: AUC = 0.66 (0.57, 0.75) CURB-65: AUC = 0.79 (0.71, 0.86)	PMEWS ≥2 Hospitalization sens: 0.99 (0.97; 1.00) spe: 0.55 (0.43; 0.66) NPV: 0.95 (0.89.1.02) PPV: 0.82 (0.77; 0.87) Mortality sens: 1.00 (1.00; 1.00) spe: 0.06 (0.02; 0.10) NPV: 1.00 (1.00; 1.00) PPV: 0.24 (0.17; 0.30) PMEWS ≥3 Hospitalization sens: 0.94 (0.90; 0.97) spe: 0.71 (0.61; 0.81) NPV: 0.85 (0.76; 0.93) PPV: 0.87 (0.82; 0.92) Mortality sens: 1.00 (1.00; 1.00) spe: 0.14 (0.08; 0.19) NPV: 1.00 (1.00; 1.00) PPV: 0.25 (0.19; 0.32)	CURB-65 ≥ 2 Hospitalization sens: 0.66 (0.59; 0.73) spe: 0.94 (0.88; 0.99) NPV: 0.56 (0.48; 0.65) PPV: 0.96 (0.92; 0.99) Mortality sens: 0.95 (0.89; 1.02) spe: 0.41 (0.33; 0.49) NPV: 0.97 (0.92; 1.01) PPV: 0.32 (0.24; 0.40) CURB-65 ≥ 3 Hospitalization sens: 0.40 (0.33; 0.48) spe: 1.00 (1.00; 1.00) NPV: 0.44 (0.37; 0.51) PPV: 1.00 (1.00; 1.00) Mortality sens: 0.74 (0.60; 0.87) spe: 0.64 (0.56; 0.71) NPV: 0.90 (0.84; 0.96) PPV: 0.36 (0.26; 0.46)

(*continued on next page*)

Table 1 (*continued*)

Rodriguez-Noriega et al. (2010) PLoS One [16]	Tertiary care ED Guadalajara, Mexico	April 2009 to August 2009 Pandemic H1N1	Prospective observational cohort	Adults with acute respiratory symptoms Median age: 29 (IQR 22–41) years	Hospitalization	ILI-score (16 criteria based on clinical signs and symptoms, history, laboratory, radiology)	Standardized questionnaire	No missing data	Hospitalization:1840/233	External validation (description)	ILI-score ≥7 Sens: 0.94 (0.90; 0.97) Spe: 0.58 (0.55; 0.60) NPV: 0.98 (0.98; 0.99) PPV: 0.24 (0.22; 0.27)	ILI-score ≥16 Sens: 0.49 (0.42; 0.55) Spe: 0.51 (0.49; 0.54) NPV: 0.87 (0.85; 0.89) PPV: 0.13 (0.11; 0.15)	
Challen et al. (2012) Emerg Med J [17]	3 EDs Sheffield and Manchester, United Kingdom	September 2009 to February 2010 Second wave of the 2009H1N1 pandemic	Prospective observational cohort	Children and adults suspected of pandemic influenza (H1N1) Age ≤16 years: 72%	Hospitalization	PMEWS (10 criteria based on age, physiological data, history, social isolation, performance status of limited activity) CURB-65 (5 criteria based on physiological data, age, laboratory) CAT for swine flu hospital pathway (7 criteria based on physiological data, clinical signs)	Clinical assessment form	No missing data	Hospitalization: 178/44 (adults)	External validation (for adults) CURB-65 AUC = 0.65 (0.54, 0.76) PMEWS AUC = 0.76 (0.66, 0.86) CAT for swine flu hospital pathway AUC = 0.62 (0.51, 0.72)	CURB-65 ≥ 2 Sens: 0.18 (0.07; 0.30) Spe:0.93 (0.89; 0.97) NPV:0.78 (0.71; 0.84) PPV: 0.47 (0.23; 0.71) CURB-65 ≥ 3 Sens: 0.02 (−0.02; 0.07) Spe: 0.99 (0.98; 1.01) NPV: 0.76 (0.69; 0.82) PPV: 0.50 (−019; 1.19)	PMEWS ≥2 Sens: 0.80 (0.68; 0.91) Spe: 0.40 (0.31; 0.48) NPV: 0.85 (0.77; 0.94) PPV: 0.30 (0.22; 0.38) PMEWS ≥3 Sens: 0.73 (0.60; 0.86) Spe: 0.55 (0.47; 0.64) NPV: 0.86 (0.79; 0.93) PPV: 0.35 (0.25; 0.44)	CAT any criterion positive Sens: 0.36 (0.22; 0.51) Spe: 0.79 (0.72; 0.86) NPV: 0.79 (0.72; 0.86) PPV: 0.36 (0.22; 0.51)

(*continued on next page*)

Table 1 (*continued*)

Chen et al. (2013) Am J Emerg Med [18]	2 tertiary care ED, and 1 community hospital ED Baltimore, USA	December 2007 to May 2009	Prospective observational cohort	Children and adults with positive influenza test results Median age: 10 (IQR 4–26) years	Hospitalization	Clinical tree decision with 5 nodes (underlying illness, age, influenza viral load level, vaccination history)	Electronic medical records	No number, surrogate splits	Hospitalization: 146/56	Development: recursive partitioning algorithm used a 2-stage procedure to derive binary trees. Internal validation: the minimal relative error rule by cost-complexity parameter using internal 10-fold cross-validation AUC = 0.84 (0.77; 0.90)	Sens: 0.83 (0.73; 0.90) Spe: 0.76 (0.63; 0.86) NPV: 0.87 (0.80; 0.95) PPV: 0.68 (0.56; 0.79)	
Chu et al. (2020) BMC Infect Dis [19].	Tertiary care ED Taoyuan, Taiwan	January 2010 to December 2016	Retrospective observational cohort	Adults with positive influenza test results Mean age: 48 years (SD 19)	Mortality	qSOFA score (3 criteria based on clinical signs) SIRS criteria (4 criteria based on clinical signs, laboratory)	Electronic medical records	No missing data	Mortality: 3561/95	External validation qSOFA score AUC = 0.86 SIRS criteria AUC = 0.79	qSOFA ≥1 Sens: 0.82 (0.74; 0.90) Spe: 0.77 (0.75; 0.78) NPV: 0.99 (0.99; 1.00) PPV: 0.09 (0.07; 0.11) qSOFA ≥2 Sens: 0.26 (0.17; 0.35) Spe: 0.96 (0.96; 0.97) NPV: 0.98 (0.97; 0.98) PPV: 0.17 (0.11; 0.23)	SIRS ≥2 Sens: 0.79 (0.71; 0.87) Spe: 0.24 (0.23; 0.26) NPV: 0.98 (0.97; 0.99) PPV: 0.03 (0.02; 0.03) SIRS ≥3 Sens: 0.55 (0.45; 0.65) Spe: 0.65 (0.63; 0.66) NPV: 0.98 (0.98; 0.99) PPV: 0.04 (0.03; 0.05)
Jehi et al. (2020) PLoS One [20]	Clinic health setting Cleveland, USA	March 2020 to June 2020	Retrospective observational cohort	Children and adults with positive Covid-19 test results Median age: 65 (IQR 52–77) years	Hospitalization	Predictive model for hospitalization risk (demographic variables, symptoms, history, laboratory, social characteristic, medication history), nomogram to estimate the patient probability of hospitalization	Electronic medical records	Missing data, multivariate imputation by chained equation	Development cohort: 2852/582 Validation cohort: 1684/376	Development: shrinkage and selection operator logistic regression algorithm AUC = 0.90 (0.89, 0.91) Internal validation: 10 fold cross-validation method, 1000 bootstrap resamples AUC = 0.81 (0.79, 0.84)	Risk 10% Sens: 0.77 Spe: 0.73 NPV: 0.92 PPV: 0.45 Risk 30% Sens: 0.52 Spe: 0.92 NPV: 0.90 PPV: 0.65 Risk 50% Sens: 0.39 Spe: 0.96 NPV: 0.85 PPV: 0.75	

(*continued on next page*)

Table 1 (*continued*)

Cheong et al. (2021) BMC infect Dis [21].	2 tertiary care ED Taoyuan, Taiwan	2010 to 2016	Retrospective observational cohort	Adults with positive influenza test results Mean age: 51 (SD 19) years	Mortality	Predictive model for mortality (demographic variables, triage category, baseline vital signs, history, laboratory nomogram)	Electronic medical records	2876, excluded	Mortality: 1680/72	Development: multiple logistic regression and stepwise and backward regression AUC = 0.76 External validation: mortality AUC = 0.77	NA
Salvatore et al. (2021) Cureus [22]	Tertiary care ED Bologna, Italy	March to April 2020	Prospective observational cohort	Adults with positive Covid-19 test results Mean age: 62 (SD 19) years	Hospitalization Mortality within 30 days	CovHos score (5 criteria based on demographic variables, laboratory and alveolar-to-arterial oxygen gradient)	Electronic medical records	NA	Hospitalization: 667/465 for derivation cohort 309/228 for validation cohort Mortality: 667/108 for derivation cohort	Development: multivariable logistic regression model Hospitalization/ AUC = 0.91 (0.88, 0.93) Mortality AUC = 0.82 (0.78, 0.87) Internal validation: Hospitalizatino AUC = 0.90 (0.88, 0.93)	Derivation cohort: Hospitalization CovHos ≥12 Sens: 0.85 Spe: 0.82 Mortality CovHos ≥22 Sens: 0.79 Spe: 0.77 Validation cohort: Hospitalization CovHos ≥12 Sens: 0.82 Spe: 0.74
Herings et al. (2021) BMJ Open [23]	264 general practitioners, Netherlands	April 2020 to January 2021	Retrospective observational cohort	Adults with positive Covid-19 test results, registered in the Covid database	Hospitalization and mortality	Predictive model (6 criteria based on demographic variables, comorbidities, social characteristic)	Electronic medical records	No missing data	Hospitalization and mortality: 4057/1979 for derivation cohort, 1729/94 for validation cohort	Development: Lasso regression Internal validation: AUC = 0.91 (0.88, 0.94)	NA
Webb et al. (2022) PLoS ONE [24]	32 urgent care facilities, 23 ED, and 16 community drive-up testing sites Utah, USA	March to October 2020	Retrospective observational cohort	Adults with positive Covid-19 test results Mean age: 40 (SD 16) years	Hospitalization (within 14 days of testing) Mortality (within 28 days)	Predictive model (demographic variables, symptoms, comorbidities, history, social characteristic)	Electronic medical records	NA	Hospitalization: 16,030/990 for derivation cohort 6786/429 for validation cohort Mortality: 16,030/73 for derivation cohort 6786/20 for validation cohort	Development: multivariable logistic regression model Hospitalization/ AUC = 0.82 (0.81, 0.84) Mortality/AUC = 0.91 (0.83, 0.94) Internal validation: Hospitalization/ AUC = 0.80 (0.78, 0.82) Mortality/AUC = 0.80 (0.69, 0.90)	Score ≥6 Sens: 0.71 Spe: 0.76 NPV: 0.97 PPV: 0.17

(*continued on next page*)

Table 1 (*continued*)

De Freitas et al. (2022) J Clin Med [25]	Tertiary care ED Sao Paulo, Brazil	March to August 2020	Retrospective observational cohort	Adults with respiratory syndrome during the first wave of the Covid-19 pandemic	Hospitalization	Predictive model (11 criteria based on demographic variables, symptoms, baseline vital signs, comorbidities)	Electronic medical records	Missing data imputed using nearest neignbors	Hospitalization: 7336/740	Development: tree model, random forest and Lasso regression AUC = 0.89 to 0.93 Internal validation: machine learning approaches by 10-fold cross-validation AUC = 0.92 (0.90; 0.94)	NA
Brooks et al. (2022) JACEP Open [26]	49 ED across 8 provinces, Canada	March 2020 to September 2021	Retrospective observational cohort	Unvaccinated adults with positive Covid-19 test results who were discharged from the ED Mean age: 46 (SD 16) years	Hospitalization and mortality (within 72 h of ED discharge)	CCEDRRN Covid discharge score (7 criteria based on demographic variables, baseline vital signs, history, arrival mode, pregnancy or not and respiratory distress)	Electronic medical records	Multiple implementations for missing data	Hospitalization and mortality: 11,487/404 for derivation cohort 3818/131 for validation cohort	Development: multiple logistic regression with restricted cubic splines and fast step-down procedure AUC = 0.70 (0.68, 0.73) Internal validation: 1000 bootstrap samples and 5 multiple imputations AUC = 0.71 (0.67, 0.75)	Validation cohort: Score ≤3 Sens: 0.89 (0.84; 0.95) Spe: 0.39 (0.37; 0.40) NPV: 0.99 (0.98; 0.99) PPV: 0.05 (0.04; 0.06) Score ≥9 Sens: 0.15 (0.09; 0.24) Spe: 0.96 (0.95; 0.96) NPV: 0.97 (0.96; 0.97) PPV: 0.11 (0.06; 0.15)

Definition: ED = emergency department; PMEWS = pandemic medical early warning score; ILI = influenza-like illness; CAT = community assessment tool; CCEDRRN = Canadian COVID-19 Emergency Department Rapid Response Network; IQR: interquartile range; SD: standard derivation; sens = sensibility; spe = specificity; NPV = negative predictive value; PPV = positive predictive value; AUC = area under the curve; NA = not applicable.

3.3. Outcomes

Hospitalization was the primary outcome in seven of the twelve studies [15,16,18,20,22,24,25] and a secondary outcome in one study [17]. Hospitalization was measured on the day of enrolment in the study, except in three case where it was considered any time within three days [20,26] or within 14 days [24] of enrolment. In-hospital mortality was recorded in five studies, as a primary outcome in one [19] and as a secondary outcome in the others [15,21,22,24]. Performance of risk-stratification tools in terms of both hospitalizations and mortality was assessed in two study [23,26]. It was specified in only one study that outcome assessors were blind to the results of the risk stratification tool [19].

3.4. Scores

In four studies, external validation results of six influenza-like illness risk stratification scores were reported. The pandemic medical early warning score (PMEWS), which assigns points for abnormal vital signs, age ≥65, social isolation, chronic disease and functional limitations, was validated in a population of adults with community-acquired pneumonia as a proxy for pandemic influenza [15], where it was a good predictor of the need for hospitalization (area under the receiver operating characteristic curve 0.94 or 0.92–0.97) but was less successful in predicting inpatient mortality (AUC 0.66 or 0.57–0.75). The same team later validated the PMEWS score secondarily in an adult population presenting to the ED of three hospitals with suspected influenza, obtaining an AUC of 0.76 (0.66–0.86) for discrimination between hospitalized and non-hospitalized patients [17]. The performance of the PMEWS score at different thresholds (≥2 or ≥ 3) is presented in Table 1 and Fig. 2.

The CURB-65 score, which assigns points for age and physiological and laboratory data, also was assessed on adult patients with community-acquired pneumonia [15]. This score was not as good a predictor of the need for hospitalization as the PMEWS was (AUC 0.88 or 0.84–0.92). However, it seemed to perform better at predicting inpatient mortality (AUC 0.79 or 0.71–0.86). In the later study by this team [17] the CURB-65 was less successful at predicting hospitalization of adults with suspected influenza (AUC 0.65 or 0.54–0.76). CURB-65 score performance at different thresholds (≥2 or ≥ 3) is presented in Table 1 and Fig. 2.

The community assessment tool (CAT) consisting of seven physiological or clinical signs was developed in United Kingdom for the swine flu hospital pathway in early 2009. The performance of this score for adults with suspected pandemic influenza also was evaluated along with PMEWS and CURB-65. Its AUC for discriminating between hospitalized and non-hospitalized patients was 0.62 or 0.51–0.72 [17].

The influenza-like illness score (ILI score), which consists of 16 items including clinical signs and laboratory and radiology results, was developed originally to support the decision to hospitalize elderly patients with pneumonia or influenza. It has been validated in adults with respiratory symptoms in the ED [16]. Its performance at predicting the risk of hospitalization is shown in Table 1 and Fig. 2.

The systemic inflammatory response syndrome (SIRS) criteria were developed to define sepsis. They assign points for clinical signs and white blood cell count. The qSOFA score was developed by the Third International Consensus Definitions for Sepsis (Sepsis-3) as a new sepsis screening tool for use outside the intensive care unit in the ED [27]. Prediction of in-hospital mortality of patients with influenza by SIRS had an AUC of 0.79 compared to 0.86 for qSOFA [19]. Based on these data, the accuracy of the SIRS criteria was estimated to be best at a cut-off score ≥3. However, the sensitivity of SIRS criteria ≥3 was poorer than a qSOFA score ≥1, which ensured better sensitivity than a qSOFA score ≥2, which is the cut-off point defined by Sepsis-3 [19].

Two studies developed a Covid-19 risk stratification score but without external validation. The first concerned the CovHos score which was created to give a tool to assist clinicians in stratifying patients based on the severity at their arrival at the ED and in predicting the need for hospitalization and mortality within 30 days [22]. The CovHos score, based on 5 criteria (age, sex, alveolar to arterial oxygen gradient, neutrophils/lymphocytes ratio and CRP) identified a cut-off of 12 points in predicting hospitalization. The AUC was 0.91 (0.88–0.93), the sensitivity was 85% and the specificity was 82%. A cut-off of 22 points was defined to predict mortality with an AUC of 0.82 (0.78–0.87). The performances were lower with a sensitivity of 79% and a specificity of 77%.

The second study concerned the CCEDRRN Covid discharge score developed by the Canadian Covid-19 Rapid Response Network (CCEDRRN) [26]. This score was based on 7 variables (age, sex, temperature, arrival mode, pregnant or not, respiratory distress and respiratory rate) to identify patients at risk for hospitalization or death within 72 h of ED discharge. The AUC was 0.70 (0.68–0.73). The sensitivity of using a score of 3 was 0.89 (0.84–0.95); the specificity of using a score of 9 was 0.96 (0.95–0.96).

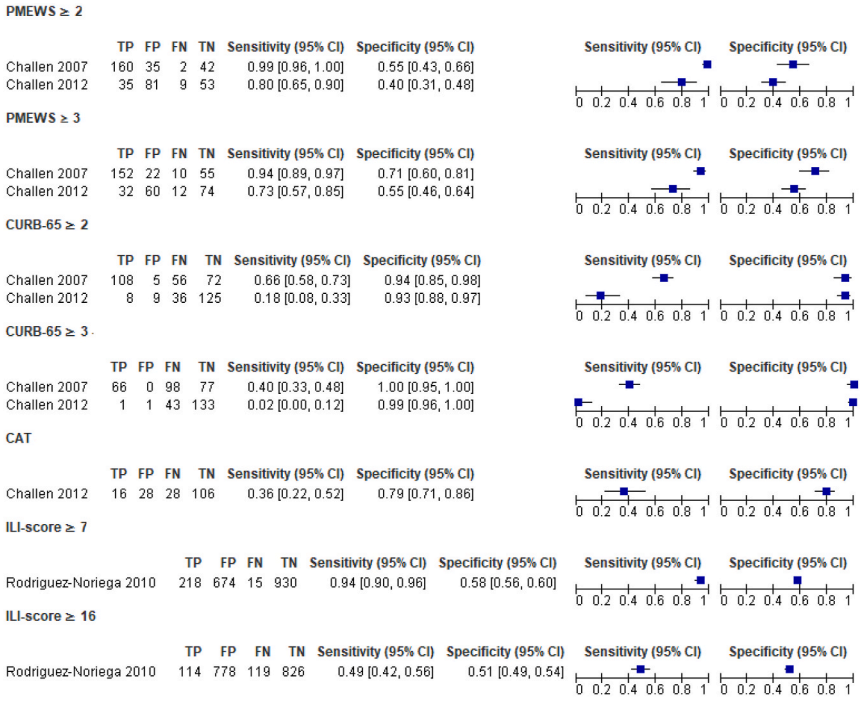
The results are reported in Table 1 and Fig. 2.

According to the results of our systematic review, a meta-analysis was not feasible. In fact, only the two Challen et al. studies (2007, 2012) assessed the same scores (PMEWS and CURB-65) and outcome (hospitalization). Given the significant heterogeneity of the populations studied, a meta-analysis was deemed inappropriate and was not performed [12–14]. Similarly, because studies that used mortality as an outcome measure were on different scores, their results could not be combined in a meta-analysis for the present purpose.

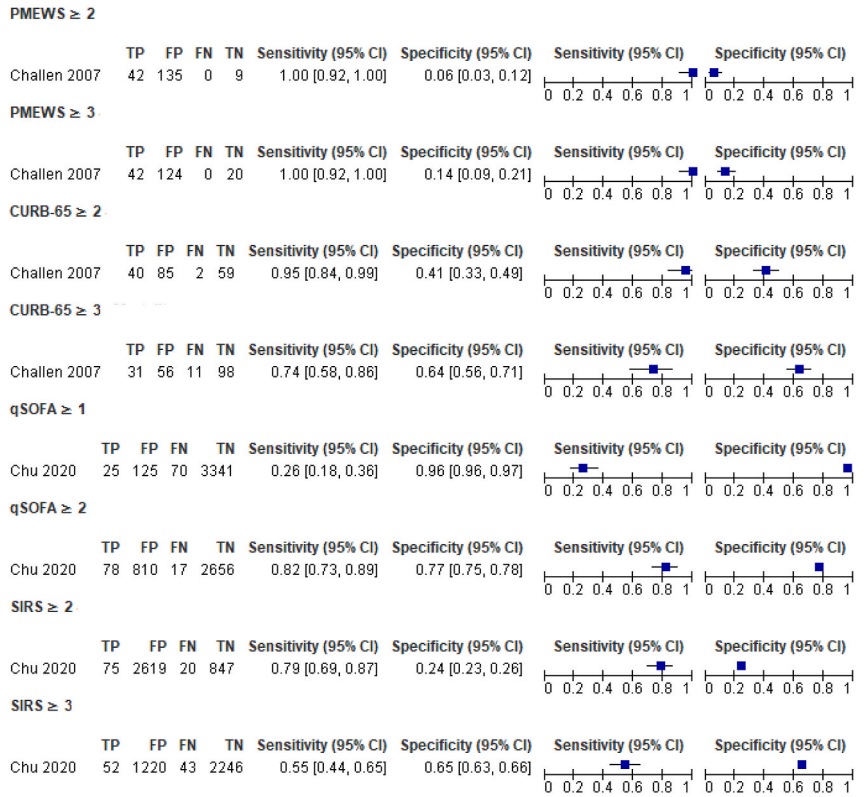
3.5. Predictive models

In six studies, rules for estimating the probability of complications in patients with influenza-like illness or Covid-19 symptoms were derived from clinical results (Table 1). A decision tree algorithm derived by recursive partitioning was developed to determine the need to hospitalize patients that test positive for influenza [18]. The performance of this algorithm was evaluated by internal 10-fold cross validation. Predictors of hospitalization included chronic illness (e.g., diabetes), age, influenza vaccination status, and influenza viral load. The AUC was 0.84 (0.77–0.90), the sensitivity was 0.83 (0.73–0.90) and the specificity was 0.76 (0.63–0.86).

Hospitalization



Mortality



(caption on next page)

Fig. 2. Forest plot showing sensitivity and specificity of the scores with external validation selected to predict hospitalization and mortality among patients with influenza-like illness.

A risk stratification tool for predicting in-hospital mortality in ED patients diagnosed with influenza has been derived using a two-stage-modeling method, in the first stage fitting a multiple logistic regression model with age, sex, vital signs, history, and complete blood count as independent variables and using this score with other laboratory data to construct a backward regression model. A nomogram was then derived to calculate a score (up to 550 points) using hypothermia, tachypnea, low systolic blood pressure, diabetes mellitus, leukocytosis, leukopenia, high percentage of segmented neutrophils and high concentration of C-reactive protein as risk factors. External validation of the model yielded an AUC of 0.77 [21].

The subject of two studies was the development of a statistical model to predict hospitalization risk for patients diagnosed with COVID-19 [20,25]. For Jehi et al. a least absolute shrinkage and selection operator logistic regression algorithm was run to identify the most predictive risk factors, followed by an internal 10-fold cross validation of performance. The hospitalization risk was thus found to be increased by age, being male, smoking (present or past), diabetes, hypertension, chronic lung disease, poor socioeconomic status, shortness of breath, diarrhea, and certain medications such as non-steroidal anti-inflammatory drugs and immunosuppressive treatment. The model discrimination was excellent with an AUC of 0.81 (0.79–0.84) in the validation cohort [20]. For DeFreitas et al., the algorithm was created with a decision tree and lasso regression, followed by a machine learning approaches by 10-fold cross validation of performance. Including age, sex, baseline vital signs, duration of symptoms, presence of hypertension or diabetes mellitus, the model had a high discriminatory value (AUC 0.92 (0.90–0.94)) [25].

Two predictive models were derived to predict hospitalization and mortality for patients with Covid-19 test results [23]. For Herings et al. least absolute shrinkage and selection operator regression analysis was used to select predictors in the model and to estimate and shrink regression coefficients. The variables selected into the model were age, sex, chronic comorbidity score, neighborhood deprivation score and wave of Covid-19 pandemic. The AUC was 0.91 (0.88–0.94) [23]. Webb et al. used a multivariable logistic regression model to identified predictors for hospitalization and mortality [24]. Age, sex, communities color, comorbidities and presence of dyspnea are included in the model which has an excellent discrimination (AUC ranged from 0.82 to 0.91). A score of 6 has a sensitivity and specificity of 71% and 76% respectively.

The results are reported in Table 1.

3.6. Risk of bias assessment

PROBAST is used to assess both the risk of bias and the applicability of a study across the four domains (participants, predictors, outcome, and analysis).

Only one study raised a major concern about applicability, namely the Rodriguez-Noriega study [16], which also had a high risk of bias. The Cheong et al. study [21] might be somewhat biased (unclear) because of the large number of excluded patients. The clinical prediction rules carry a high risk of bias since their models were not validated externally and therefore warrant such a rating according to analysis by PROBAST [10]. Details of the risk of bias assessment are presented in Table 2.

4. Discussion

The studies that we examined in this systematic review propose fourteen different tools that were developed to predict the risk of hospitalization or death of ED or outpatient clinic patients with influenza-like illness. Eight tools were scores and six were predictive models. Among the scores, the PMEWS seemed to be the most accurate in predicting the need for hospitalization of adults, with a sensitivity ranging from 0.80 to 0.99 for a score $\geq$2. The predictive model developed by Chen et al. was as good as the PMEWS at predicting hospitalization but was derived from a population composed mostly of pediatric patients and its risk of bias was considered high based on its validation. CURB-65 or qSOFA scores $\geq$2 appeared to be the most accurate predictors of death, whereas the predictive model of Cheong et al. was adequate based on the PROBAST but less reliable than the scores. Although these scores and models need more robust validation using larger populations, they already could support clinical judgment by drawing attention to risk factors for complications of influenza-like illness.

Severity scores currently available to support triage decisions for patients with influenza-like illness were developed for pandemic influenza (e.g., PMEWS, CAT), pneumonia (e.g., CURB-65) or sepsis (e.g., qSOFA, SIRS). Provisional guidelines have long recommend the use of severity assessment tools to help frontline physicians manage pandemic influenza [5]. CURB-65 is simple, consisting of only five criteria, namely confusion, urea, respiratory rate, blood pressure, and age over 65 years. However, this score performed poorly in young adults [4]. Predictive factors of severe influenza are well defined in the literature and include age, comorbidities, respiratory failure, hypoxia, and abnormal vital signs or laboratory analyses [28]. Factors such as hypoxia and comorbidities added as refinements might improve the performance of CURB-65 in predicting the risk of complications of influenza-like illness.

United Kingdom Department of Health guidelines on surge capacity also recommend the use of the PMEWS for triaging patients with influenza-like illness. This score includes more risk factors than CURB-65 does [29] and is simple and easily applied in the ED since it does not involve any laboratory tests but relies instead on physiological abnormalities apparent in triage to determine the risk of complications and the appropriate level of care [15].

Designed to help clinicians identify patients that may be referred to secondary care, the community assessment tool or CAT is applicable when high demand for healthcare forces tightening of the criteria for admitting patients to hospitals in affected areas [28].

Table 2
PROBAST results.

Study	ROB				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Challen 2007	+	+	+	+	+	+	+	+	+
Rodriguez-Noriega 2010	+	?	-	-	+	?	?	-	?
Challen 2012	+	+	+	+	+	+	+	+	+
Chen 2013	+	+	+	+	+	+	+	-	+
Chu 2020	+	+	+	+	+	+	+	+	+
Jehi 2020	+	+	+	+	+	+	+	-	+
Cheong 2021	+	+	+	?	+	+	+	?	+
Salvatore 2021	+	+	+	?	+	+	+	-	+
Herings 2021	+	+	+	+	+	+	+	-	+
Webb 2022	+	+	+	?	+	+	+	-	+
De Freitas 2022	+	+	+	+	+	+	+	-	+
Brooks 2022	+	+	+	+	+	+	+	-	+

Definition: PROBAST = Prediction model Risk Of Bias ASsessment Tool; ROB = risk of bias.
+ indicates low ROB/low concern regarding applicability.
− indicates high ROB/high concern regarding applicability.
? indicates unclear ROB/unclear concern regarding applicability.

The CAT decision tool identifies patients with any severity criteria (e.g., respiratory failure, altered vital signs) as admissible. Its use is not recommended when EDs are working at their usual staffing levels and can apply their routine care procedures [30].

Other scores such as SIRS and qSOFA have been developed to screen for sepsis and predict in-hospital mortality. Although they have been validated quite robustly as predictors of adverse outcomes in patients with sepsis, their application to patients with influenza requires further study. Chu et al. has demonstrated that the predictive performance of the qSOFA may be better than that of the SIRS in the prognostication of patients with influenza, but this finding may be related to the inclusion of a significant number of patients with influenza-induced sepsis or a secondary bacterial infection in their cohort. A patient who presents to ED triage with a qSOFA score as low as 1 may require hospitalization [19].

During a pandemic, hospital beds may become scarce, and suitable decision aids could help front-line healthcare providers recognize patients that need to be admitted. Some tools initially developed to assess severity and assist triage of patients with community-acquired pneumonia have been validated secondarily for pandemics or seasonal epidemics. However, chronic overcrowding of EDs requires a battery of solutions, which may include the use of risk-stratification tools to redirect patients safely to other care settings or to self-care at home. For example, British guidelines recommend that patients with influenza-related pneumonia and a CURB-65 score of 0 or 1 be sent home with instructions for self-care [5]. Similarly, since the PMEWS is based on readily observable clinical characteristics, future research could explore whether a low score justifies discharging patients to their homes or redirecting them to care settings other than the ED without seeing the ED physician. The US Centers for Disease Control and Prevention propose an expert-based management algorithm with an assessment of severity based on age and co-morbidities to predict adverse outcome for patients with influenza [31]. In this systematic review, we have identified several risk-stratification tools that may be considered for predicting hospitalization or mortality of patients with influenza. The risk factors used most frequently in these tools are comorbidities, age, symptoms (e.g., fever, cough, dyspnea) vital signs, and vaccination history. However, most of the tools also include laboratory analyses, which limits their applicability during triage in the ED or in an outpatient clinic. Chen et al. found that viral load could be predictive of the need for hospitalization in patient populations with milder influenza and fewer comorbidities. Such a measurement would undoubtedly improve the reliability of risk stratification scores based solely on clinical features and represents a huge opportunity for developers of rapid bedside diagnostic devices for respiratory viral infections in general [18].

Before listing the limitations of this review, let us mention that our literature search was structured to find prognostic tools that predict complications in patients with influenza-like illness and was not limited to confirmed influenza. This corresponds more closely to clinical practice in the ED, which is focused on symptoms presented by patients on arrival at the ED and not on a specific presumed or confirmed acute respiratory infection. We therefore included studies that assess tools intended to predict adverse outcomes in patients with influenza-like illness, including COVID-19. This allowed us to select a larger number of studies, but with a more heterogeneous population. On the other hand, 26 articles were excluded because they were based on populations that differed too much from those we sought to study. As a result, tools developed for critical care unit patients were excluded. We also excluded 14 studies of tools used to predict the need to place already hospitalized patients in intensive care. Finally, the diversity of scores and high heterogeneity of the study populations of the articles included in our review did not allow us to perform a meta-analysis.

5. Conclusion

This systematic review allowed us to identify nine different tools that have been developed to predict the risk of hospitalization or death in ED or ambulatory clinic patients with influenza-like illness. The PMEWS seems to be the best predictor of the need for hospitalization in the adult population, while the CURB-65 seems to be the best predictor of mortality. Although some of these scores have interesting features that may support clinical judgment, none of them appears to perform well enough to justify its widespread adoption in the ED. There is a need to refine an existing tool or develop a new one to optimize the management of these patients.

Data availability statement

Data associated with this study has no been deposited into a publicly available repository. However, data will be made available on request from the corresponding author.

CRediT authorship contribution statement

Tania Marx: Writing - review & editing, Writing - original draft, Validation, Formal analysis. **Nada Khelifi:** Writing - review & editing, Formal analysis. **Isabelle Xu:** Writing - review & editing, Formal analysis. **Laurie Ouellet:** Writing - review & editing, Formal analysis. **Annie Poirier:** Writing - review & editing, Formal analysis. **Benoit Huard:** Writing - review & editing, Formal analysis. **Myriam Mallet:** Writing - review & editing, Funding acquisition. **Frédéric Bergeron:** Writing - review & editing, Data curation. **Maurice Boissinot:** Writing - review & editing, Conceptualization. **Michel G. Bergeron:** Writing - review & editing, Funding acquisition, Conceptualization. **Simon Berthelot:** Writing - review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Simon Berthelot reports article publishing charges was provided by Quebec Ministry of the Economy and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e23227.

References

[1] D.Y. Gaitonde, F.C. Moore, M.K. Morgan, Influenza: diagnosis and treatment, Am Fam Physician. 15 déc 100 (12) (2019) 751–758.
[2] A.D. Iuliano, K.M. Roguski, H.H. Chang, D.J. Muscatello, R. Palekar, S. Tempia, et al., Estimates of global seasonal influenza-associated respiratory mortality: a modelling study, The Lancet. 31 mars 391 (10127) (2018) 1285–1300.
[3] W.C.W.S. Putri, D.J. Muscatello, M.S. Stockwell, A.T. Newall, Economic burden of seasonal influenza in the United States, Vaccine. 22 juin 36 (27) (2018) 3960–3966.
[4] A. Singanayagam, J.D. Chalmers, A.T. Hill, Severity assessment in community-acquired pneumonia: a review, QJM Mon J Assoc Physicians. juin 102 (6) (2009) 379–388.
[5] British Infection Society, British Thoracic Society, Health protection agency. Pandemic flu: clinical management of patients with an influenza-like illness during an influenza pandemic. Provisional guidelines from the British infection society, British thoracic society, and health protection agency in collaboration with the department of health, Thorax. janv 62 (Suppl 1) (2007) 1–46.
[6] M. Patel, A. Dennis, C. Flutter, Z. Khan, Pandemic (H1N1) 2009 influenza, BJA Br J Anaesth. févr 104 (2) (2010) 128–142.
[7] N. Qualls, A. Levitt, N. Kanade, N. Wright-Jegede, S. Dopson, M. Biggerstaff, et al., Community mitigation guidelines to prevent pandemic influenza — United States, 2017, MMWR Recomm Rep. 21 avr 66 (1) (2017) 1–32.
[8] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, et al., The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ 372 (2021) n71.
[9] K.G.M. Moons, J.A.H. de Groot, W. Bouwmeester, Y. Vergouwe, S. Mallett, D.G. Altman, et al., Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist, PLoS Med. 11 (10) (2014), e1001744.
[10] K.G.M. Moons, R.F. Wolff, R.D. Riley, P.F. Whiting, M. Westwood, G.S. Collins, et al., PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration, Ann Intern Med. 1 janv 170 (1) (2019), W1-W33.
[11] D.L. Schriger, D.G. Altman, J.A. Vetter, T. Heafner, D. Moher, Forest plots in reports of systematic reviews: a cross-sectional study reviewing current practice, Int. J. Epidemiol. 39 (2) (2010) 421–429.
[12] K.W. Kim, J. Lee, S.H. Choi, J. Huh, S.H. Park, Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-Part I. General guidance and tips, Korean J. Radiol. 16 (6) (2015) 1175–1187.
[13] J. Campbell, M. Klugar, S. Ding, D. Carmody, S. Hakonsen, Y. Jadotte, The Systematic Review of Studies of Diagnostic Test Accuracy, Joanna Briggs Inst Rev Man, 2015, pp. 1–46.
[14] Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: analysing and presenting results. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 10. 210apr. J.-C.(The Cochrane Collaboration)..
[15] K. Challen, J. Bright, A. Bentley, D. Walter, Physiological-social score (PMEWS) vs. CURB-65 to triage pandemic influenza: a comparative validation study using community-acquired pneumonia as a proxy, BMC Health Serv Res. 1 mars 7 (2007) 33.
[16] E. Rodriguez-Noriega, E. Gonzalez-Diaz, R. Morfin-Otero, G.F. Gomez-Abundis, J. Briseño-Ramirez, H.R. Perez-Gomez, et al., Hospital triage system for adult patients using an influenza-like illness scoring system during the 2009 pandemic-Mexico, PloS One. 14 mai 5 (5) (2010), e10658.
[17] K. Challen, S.W. Goodacre, R. Wilson, A. Bentley, M. Campbell, C. Fitzsimmons, et al., Evaluation of triage methods used to select patients with suspected pandemic influenza for hospital admission, Emerg Med J EMJ. mai 29 (5) (2012) 383–388.
[18] K.F. Chen, Y.H. Hsieh, C.A. Gaydos, A. Valsamakis, R.E. Rothman, Derivation of a clinical prediction rule to predict hospitalization for influenza in EDs, Am J Emerg Med. mars 31 (3) (2013) 529–534.
[19] S.E. Chu, C.J. Seak, T.H. Su, C.H. Chaou, H.J. Tseng, C.H. Li, Prognostic accuracy of SIRS criteria and qSOFA score for in-hospital mortality among influenza patients in the emergency department, BMC Infect Dis. 29 mai 20 (1) (2020) 385.
[20] L. Jehi, X. Ji, A. Milinovich, S. Erzurum, A. Merlino, S. Gordon, et al., Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19, PLoS One 15 (8) (2020), e0237419.

[21] C.W. Cheong, C.L. Chen, C.H. Li, C.J. Seak, H.J. Tseng, K.H. Hsu, et al., Two-stage prediction model for in-hospital mortality of patients with influenza infection, BMC Infect Dis. 19 mai 21 (1) (2021) 451.
[22] V. Salvatore, A. Gianstefani, G. Farina, I. Carletti, N. Carpentieri, A.L. Tinuper, et al., CovHos, a new score to predict the need of hospitalization for coronavirus disease 2019 (COVID-19) patients at the emergency department, Cureus 13 (10) (oct 2021), e18717.
[23] R.M.C. Herings, K.M.A. Swart, BAM van der Zeijst, AA van der Heijden, van der Velden K, E.G. Hiddink, et al., Development and validation of an algorithm to estimate the risk of severe complications of COVID-19: a retrospective cohort study in primary care in The Netherlands, BMJ Open 11 (12) (2021), e050059.
[24] B.J. Webb, N.M. Levin, N. Grisel, S.M. Brown, I.D. Peltan, E.S. Spivak, et al., Simple scoring tool to estimate risk of hospitalization and mortality in ambulatory and emergency department patients with COVID-19, PLoS One 17 (3) (2022), e0261508.
[25] V.M. De Freitas, D.M. Chiloff, G.G. Bosso, J.O.P. Teixeira, IC. de G. Hernandes, P. Padilha M do, et al., A machine learning model for predicting hospitalization in patients with respiratory symptoms during the COVID-19 pandemic, J Clin Med. 5 août 11 (15) (2022) 4574.
[26] S.C. Brooks, R.J. Rosychuk, J.J. Perry, L.J. Morrison, H. Wiemer, P. Fok, et al., Derivation and validation of a clinical decision rule to risk-stratify COVID-19 patients discharged from the emergency department: the CCEDRRN COVID discharge score, J Am Coll Emerg Physicians Open. déc 3 (6) (2022), e12868.
[27] M. Singer, C.S. Deutschman, C.W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, et al., The third international consensus Definitions for sepsis and septic shock (Sepsis-3), JAMA. 23 févr 315 (8) (2016) 801–810.
[28] A. Singanayagam, A. Singanayagam, V. Wood, J.D. Chalmers, Factors associated with severe illness in pandemic 2009 influenza a (H1N1) infection: implications for triage in primary and secondary care, J Infect. oct 63 (4) (2011) 243–251.
[29] Department of Health, Pandemic Flu: Managing Demand and Capacity in Health Care Organisations, 2009. Disponible sur: https://www.hsj.co.uk/download?ac=1205655.
[30] Department of Health Gateway., Swine flu clinical package, Disponible sur: https://webarchive.nationalarchives.gov.uk/ukgwa/20120106065601, 2009.
[31] US Centers for disease control and Prevention, Influenza Season Triage Algorithm for adults (older than 18 years of age) with influenza-like illness, Disponible sur: https://www.cdc.gov/h1n1flu/clinicians/pdf/adultalgorithm.pdf, , 2009.