



OPEN

Predicting major clinical events among Canadian adults with laboratory-confirmed influenza infection using the influenza severity scale

Henrique Pott^{1,2✉}, Jason J. LeBlanc^{1,3}, May ElSherif¹, Todd F. Hatchette^{1,3}, Shelly A. McNeil^{1,4}, Melissa K. Andrew^{1,5} & the Serious Outcomes Surveillance (SOS) Network of the Canadian Immunization Research Network (CIRN)*

We developed and validated the Influenza Severity Scale (ISS), a standardized risk assessment for influenza, to estimate and predict the probability of major clinical events in patients with laboratory-confirmed infection. Data from the Canadian Immunization Research Network's Serious Outcomes Surveillance Network (2011/2012–2018/2019 influenza seasons) enabled the selecting of all laboratory-confirmed influenza patients. A machine learning-based approach then identified variables, generated weighted scores, and evaluated model performance. This study included 12,954 patients with laboratory-confirmed influenza infections. The optimal scale encompassed ten variables: demographic (age and sex), health history (smoking status, chronic pulmonary disease, diabetes mellitus, and influenza vaccination status), clinical presentation (cough, sputum production, and shortness of breath), and function (need for regular support for activities of daily living). As a continuous variable, the scale had an AU-ROC of 0.73 (95% CI, 0.71–0.74). Aggregated scores classified participants into three risk categories: low (ISS < 30; 79.9% sensitivity, 51% specificity), moderate (ISS ≥ 30 but < 50; 54.5% sensitivity, 55.9% specificity), and high (ISS ≥ 50; 51.4% sensitivity, 80.5% specificity). ISS demonstrated a solid ability to identify patients with hospitalized laboratory-confirmed influenza at increased risk for Major Clinical Events, potentially impacting clinical practice and research.

Keywords Influenza, Risk adjustment, Major clinical events, Outcomes

Influenza is a respiratory viral infection that affects millions worldwide yearly. The impact of influenza can vary depending on several factors, such as the virus, the host, and contextual factors like the degree of match achieved between vaccine and circulating strains, vaccine coverage, and pre-existing population immunity^{1–3}. Despite this variability, influenza remains a significant burden on people's health worldwide, with approximately one billion cases annually, of which 3–5 million are severe and 290,000–650,000 result in influenza-related deaths^{4,5}.

Although most cases have a benign course, some are at increased risk of adverse clinical outcomes, including children < 5 years, older adults, and those with a high comorbidity burden^{2,6,7}. On the other hand, influenza vaccination and the timely use of antivirals have proved effective in attenuating these outcomes^{2,3,8,9}. Understanding the benefits of these interventions requires a comprehensive evaluation in relation to illness severity, a gap in existing knowledge.

Influenza severity ranges from mild illness treated at home without any intervention or seen on an outpatient basis to more severe illness, including the need for ventilatory support, intensive care unit (ICU) admission, or death. A recent review revealed several popular tools employed for assessing the severity of influenza and

¹Canadian Centre for Vaccinology, Dalhousie University, Halifax, Canada. ²Department of Medicine, Universidade Federal de São Carlos, Rod. Washington Luis, km 235, São Carlos, SP 13656-905, Brazil. ³Department of Pathology, Dalhousie University, Halifax, Canada. ⁴Department of Medicine (Infectious Diseases), Dalhousie University, Halifax, Canada. ⁵Department of Medicine (Geriatrics), Dalhousie University, Halifax, Canada. *A list of authors and their affiliations appears at the end of the paper. ✉email: henriquepott@ufscar.br

community-acquired pneumonia, such as the Pneumonia Severity Index (PSI), CURB-65, Acute Physiology And Chronic Health Evaluation II (APACHE II), Sequential Sepsis-related Organ Failure Assessment (SOFA), and quick SOFA (qSOFA)⁷. Although these tools are used to estimate the severity of influenza, they are not specific to it.

Thus, there is an unmet need for a standardized risk assessment for influenza, particularly to characterize and estimate the probability of experiencing adverse clinical outcomes by score or a predetermined risk level and to adjust studies assessing the effect of interventions on these outcomes. In addition to being robust, this tool must be simple enough to allow its application in retrospective and prospective studies. Such a tool would enable public health systems to establish proper surveillance and evaluate the effectiveness of public health protocols tailored by risk severity.

Here, we aimed to develop a scale that can identify patients at risk of severe influenza outcomes, thus helping to guide preventive and therapeutic interventions.

Methods

Data source

The Canadian Immunization Research Network (CIRN) is a nationwide group of top vaccine experts working on vaccine safety, effectiveness, and acceptance (<https://cirnetwork.ca/>). They also focus on the implementation and evaluation of vaccination programs. CIRN plays a key role in providing research insights that help shape public health decisions related to vaccinations, ultimately benefiting the health of Canadians. The CIRN Serious Outcomes Surveillance (SOS) Network, established in 2009, aims to understand the impact of influenza and assess how effective seasonal flu vaccines are. As such, hospitalized patients meeting a broad definition of acute respiratory illness who have been tested for influenza are enrolled, either as test-positive influenza cases or test-negative controls. The SOS Network actively monitors influenza cases at multiple hospitals across several Canadian provinces each season^{3,10–14}, gathering data from different sites depending on available resources. This study used pooled data from the CIRN SOS Network database.

Participants

We used data from the 2011/2012 to 2018/2019 influenza seasons, selecting all patients with laboratory-confirmed influenza infection. The present analyses used data collected during the initial assessment of the patient within the hospital, which reflects the patient's condition immediately after being hospitalized. All hospitalized patients across the full range of illness severity were included in the present analyses, including those with and without supplemental oxygen and those requiring ventilatory support and ICU admission. There were no other data filters, and we kept all cases with missing values.

The study adhered to the guidelines outlined in the Declaration of Helsinki. The Research Ethics Boards approved the protocol, including data and sample collection and medical record review at all participating institutions (ClinicalTrials.gov Identifier: NCT1517191).

Definition of influenza infection

Nasopharyngeal swab samples from all participating subjects underwent reverse transcription polymerase chain reaction (RT-PCR) influenza testing¹⁵. Subjects were classified as “laboratory-confirmed cases” if they tested positive for influenza or “negative controls” if they tested negative. Only laboratory-confirmed influenza cases were included in the present analyses.

Data collection

Demographic and clinical data collection followed a standardized CIRN SOS Network protocol described elsewhere^{13,16}. A broad set of variables from the SOS dataset were fed into model development. Demographic data included sex and age. Health-related data included: smoking status, clinical symptoms and signs (feverishness, nasal congestion, headache, abdominal pain, malaise, cough, diarrhea, weakness, shortness of breath, vomiting, dizziness, sore throat, nausea, muscle aches, arthralgia, prostration, seizures, myalgia, sneezing, conjunctivitis, sputum production, chest pain, encephalitis, nose bleed, altered consciousness, chills, and anorexia), function i.e. degree of dependence on activities of daily living (transferring, ambulating, need for assistive devices to ambulate, balance, bathing, toileting, handling medications, dressing, eating, handling finances), need for regular support for activities of daily living, need for additional support for activities of daily living, sensory disturbances (vision, hearing, and speech), bladder and bowel dysfunction, appetite disturbances, and comorbidities (ischemic heart disease, cardiac arrhythmias, valvular disease, congestive heart failure, hypertension, peripheral vascular disease, cerebrovascular disease, dementia, other noncognitive neurological disorders, hemiplegia/paraplegia, chronic pulmonary disease, pulmonary vascular disease, rheumatological disease, peptic ulcer disease, liver disease, diabetes mellitus, solid tumor, any type of metastatic cancer, HIV/AIDS, hypothyroidism, lymphoma, coagulopathy, blood loss anemia, deficiency anemia, alcohol abuse, drug abuse, obesity, involuntary weight loss, fluid and electrolyte disorders, edema, any psychiatric disease, depression, and peripheral skin ulcers). Influenza vaccination status was deemed “vaccinated” for those who received a current season flu vaccine more than 14 days before the onset of symptoms and “unvaccinated” if otherwise. Data collection was done by on-site study monitors who obtained the data for each patient based on the best possible source, including a review of patient charts or medical records and interviews with patients, family members, and healthcare team members where required. Influenza vaccination status was verified using medical records or registries where available or through contact with the immunizing health care professional.

Outcomes of interest

The outcome of interest in this sample of patients was defined as the occurrence of a Major Clinical Event (MCE). We chose this outcome as it is a specific and measurable health event for which all study subjects were at risk at the time of hospitalization. MCE was defined as a composite outcome of the need for supplemental oxygen therapy, admission to an intermediate care unit, need for non-invasive or invasive ventilation, admission to an intensive care unit, or death. To assess the diagnostic accuracy of the scale in predicting MCE outcome, two steps were taken: (1) assessing the scale's overall diagnostic performance as a continuous variable and (2) grouping scores into three risk categories (low, moderate, and high) based on sensitivity and specificity values. The population's risk level must be considered when deciding on the test's minimum sensitivity and specificity levels. For low-risk individuals, a minimum specificity of 50% and maximum sensitivity should be chosen to ensure that those at low risk are identified without reducing the detectability of those at higher risk. For high-risk individuals, a minimum sensitivity of 50% and maximum specificity should be selected to ensure that those at higher risk are detected without reducing the screening of those at lower risk. The remaining scores were classified as moderate risk.

Development workflow

Supplementary material Fig. 1 describes the workflow scheme.

Data preprocessing

We used a data-splitting approach to validate our findings. As per the studies conducted by Dobbin and Simon¹⁷ and Nguyen et al.¹⁸, a train-test splitting ratio of around 30% is considered reasonable. To provide an unbiased evaluation of the model fit on the training dataset while fine-tuning model hyperparameters, we also held out about 15% of the training set as a validation set. We randomly divided the total sample into three sets—a training set (60%), a validation set (10%), and a test set (30%), each stratified based on their MCE status to ensure an equal balance between groups. We then transformed the raw data into a valuable and efficient format: missing values were kept at an "Unknown" level, categorical variables were converted into dummy variables, and the continuous variable (age) was centralized and standardized. The variable imbalances between sets were evaluated using standardized mean differences and differences in proportion. A strict threshold of 0.05 was applied to indicate significant imbalances between the groups¹⁹.

Variable selection

We utilized the training and validation sets for variable selection. We modeled the outcome as a function of all predicting variables using Random Forest, which generated a list of importance rankings based on the Gini index. Next, we applied the Random Forest algorithm and tenfold cross-validation to calculate the Area Under the Receiver Operating Characteristic Curve (AU-ROC). We progressively included variables from the list, starting with the highest rank, until we reached a saturation threshold of AU-ROC variance $\leq 1\%$ for two consecutive iterations.

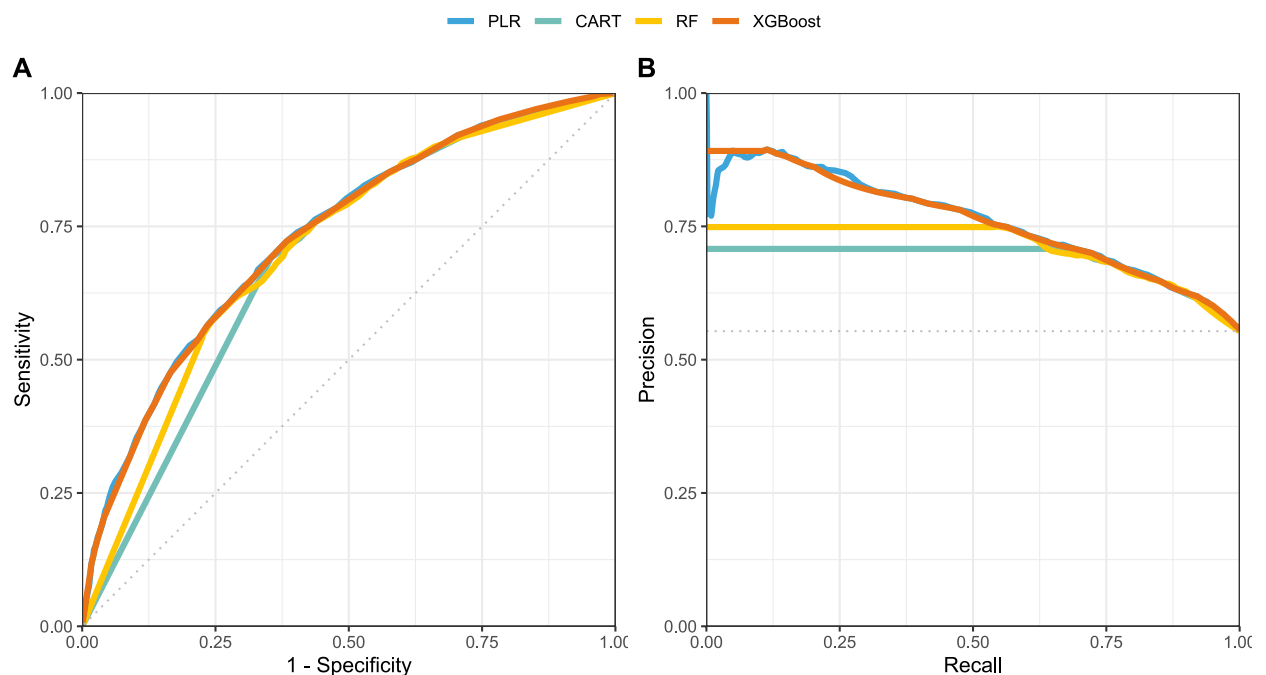


Figure 1. The (A) ROC and (B) PR curves of the models in predicting the occurrence of MCE on the test set.

Generating weighted scores

Logistic Regression was applied to generate weighted scores, modeling MCE as a function of the chosen variables. Each variable's β coefficient was divided by the lowest β coefficient and rounded to the nearest value. The sum of scores for each category gave the total score normalized to a range of 0–100 for practicality, with 100 representing the highest risk for MCE and 0 denoting a zero risk.

Model evaluation

We evaluated the weighted scores' ability to predict MCE via four methods: Penalized Logistic Regression (PLR), Classification and Regression Trees (CART), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost). We employed a stratified ten-fold cross-validation for each predictive model to select the ideal hyperparameter combination using grid search, then trained each model individually.

Evaluation metrics

We developed Receiver Operating Characteristic (ROC), Precision-Recall (PR), and Gain curves for each prediction model to evaluate the performance of the four algorithms on the test dataset. We employed six traditional metrics (AU-ROC, Area Under the Precision-Recall Curve [AU-PRC], sensitivity, specificity, precision, and F1 score) to generate predicted classes. Then, we determined point-estimated metrics by cross-tabulating the observed and predicted classes. We selected the best model based on the best performance in the ROC and PR spaces²⁰.

Data analysis

All analyses were conducted using was performed in R (version 4.2.1) using RStudio IDE (RStudio 2022.02.1.461 "Prairie Trillium" Release).

Ethics approval and participation consent

All participants provided informed consent for data, sample collection, and medical record screening per the local Research Ethics Boards' requirements. The Research Ethics Boards approved the protocol of participating institutions (ClinicalTrials.gov Identifier: NCT01517191).

Results

The original dataset enrolled 24,068 participants; 12,954 (53.8%) had laboratory-confirmed influenza infections. Supplementary Table 1 presents the overall characteristics of the study population and indicates no significant differences (above the 0.05 threshold) among the three datasets (training, validation, and test sets).

Variable selection and weighting

Supplementary Figs. 3 and 4 indicate that a list of importance rankings based on the Gini index was observed for the predicting variables and the cumulative ten-fold cross-validated AU-ROC estimates until reaching the saturation threshold for two consecutive iterations. This led us to select ten variables from four domains: demographic (age and sex), health history (smoking status, chronic pulmonary disease, diabetes mellitus, and influenza vaccination status), clinical presentation (cough, sputum production, and shortness of breath), and function (need for regular support for activities of daily living). Table 1 displays the resulting scale and its weighted scores. The scale displayed a right-skewed distribution (Supplementary Fig. 2A), with a median value of 39 across all participants (1st–3rd quartile, 23–59), and no differences between women and men (Supplementary Fig. 2B).

Comparing performance across different models

Table 2 reveals similar accuracy of predictive models on the training set, with minor disparities in AU-ROC. We chose not to exclude any models before assessing their performance on the test set. Figure 1 displays the ROC and PR curves of the prediction models on the test set, with individual ROC and gain curves in the Supplementary material (Figs. 5–8). Results show that all models had acceptable discrimination performance, ranging from 69.2 to 73.1% in AU-ROC, similar to those in their training set. Examining the ROC and PR curves, the Penalized Logistic Regression and eXtreme Gradient Boosting models demonstrate the best overall performance when using the ISS as a continuous variable.

Finding the optimal cutoff points

When tested on the set, the ISS attained a remarkable AU-ROC of 0.73 (95% CI, 0.71–0.74). When applying the Youden index to determine the optimal score threshold, it revealed a value of ≥ 37 , leading to a sensitivity of 0.70 (95% CI, 0.68–0.72) and a specificity of 0.63 (95% CI, 0.60–0.65). To interpret the scores simpler, Supplementary Table 2 provides a conversion table that maps the cutoff values to their respective predicted risks and performance metrics. For further simplification, the scores were aggregated into three categories: low risk (ISS < 30; sensitivity 79.9% [95% CI, 78–81.7%], specificity 51% [95% CI, 48.6–53.3%]), moderate risk (ISS ≥ 30 but < 50; 54.5% sensitivity, 55.9% specificity) and high risk (ISS ≥ 50 ; sensitivity 51.4% [95% CI, 49.3–53.6%], specificity 80.5% [95% CI, 78.7–82.4%]; Fig. 2).

Discussion

This study demonstrates the development and assessment of accuracy for the Influenza Severity Scale. This 10-item scale was designed to differentiate patients infected with influenza by their risk of experiencing major clinical events. The ISS is based on simple, easily accessible data and discriminates accurately. This emphasizes the

Domain	Parameter	Level	Score
Demographic	Age group	< 66 years	0
		≥ 66 years	5
	Sex	Male	0
		Female	2
Health history	Smoking status	Never smoked	0
		Former smoker	5
		Current smoker	7
		Unknown	9
	Chronic Pulmonary Diseases	No	0
		Yes	20
	Diabetes mellitus	No	0
		Yes	4
Vaccination in the current influenza season	Yes	0	
	No	1	
Clinical presentation	Cough	No	0
		Yes	4
	Sputum production	No	0
		Yes	7
	Shortness of breath	No	0
		Yes	30
Function	Require regular support for activities of daily living	Unknown	0
		No	7
		Yes	18
Total score			100
Risk stratification			
Low risk			≤ 30
Moderate risk			> 30 and < 50
High risk			≥ 50

Table 1. The Influenza Severity Scale (ISS).

Evaluation metric	Penalized Logistic Regression	Classification and regression trees	Random Forest	eXtreme Gradient Boosting
Train set				
AU-ROC*	0.730	0.690	0.709	0.728
Test set				
AU-ROC*	0.731	0.692	0.707	0.730
AU-PRC**	0.759	0.687	0.711	0.760
Accuracy	0.673	0.670	0.666	0.668
Specificity	0.563	0.562	0.527	0.522
Sensitivity	0.762	0.758	0.778	0.785
Precision	0.684	0.682	0.671	0.671
F1 score	0.721	0.718	0.721	0.723

Table 2. Performance of the models in predicting the occurrence of MCE on the train and test sets.

*AU-ROC, Area Under the Receiver Operating Characteristic curve. **AU-PRC, Area Under the Precision-Recall curve.

utility of the ISS for both retrospective and prospective studies in accurately assessing and predicting influenza-related outcomes.

Adams and colleagues reviewed common severity assessments for influenza and community-acquired pneumonia⁷. They examined 118 studies focusing on influenza, which included evaluations of tools such as PSI, CURB 65, APACHE II, SOFA, and qSOFA^{21–35}. The clinical outcomes studied in these assessments involved mortality rates (overall, in ICU, in hospital), ICU admissions, mechanical ventilation needs, length of hospital stays, and total hospitalizations. Despite the extensive findings, none of the assessments were specifically designed for influenza or considered other critical outcomes like admission to intermediate care units or the need for oxygen therapy or non-invasive ventilation. Additionally, significant differences were observed in the clinical parameters and research endpoints across the studies reviewed. These variations made it challenging to combine results and

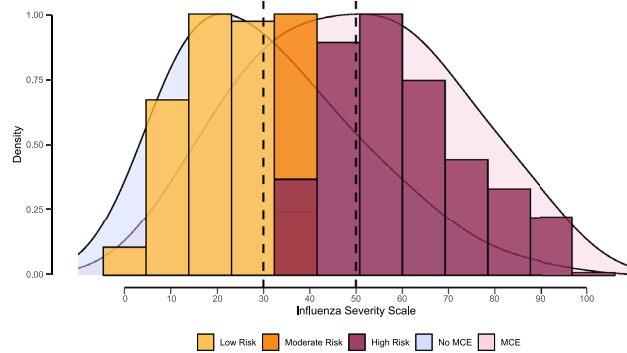


Figure 2. The histogram of the ISS scores by risk categories overlaps a density plot by the outcome MCE. Sturges' Rule determined the optimal number of bins.

perform a meta-analysis accurately. Consequently, it hindered the authors from making precise evaluations of diagnostic performance and conducting direct comparisons between different assessment tools.

Nevertheless, although PSI and CURB-65 are generally reliable in predicting 30-day mortality rates for community-acquired pneumonia in different clinical settings, some studies suggest that they may not be effective in predicting mortality rates for influenza pneumonia cases. For example, in a study conducted by Riquelme et al.³⁶, these pneumonia scoring systems were ineffective in predicting the survival rate of low-risk patients with the H1N1 2009 influenza pandemic. Another study revealed that these scoring systems are still inefficient for the influenza pandemic because they cannot accurately predict the need for intensive care services³⁷. The study revealed promising findings for SMRT-CO in identifying low-risk patients, with an AU-ROC of 0.826³⁷.

In contrast, other studies found that the AU-ROC values for predicting mortality in patients with influenza A were 0.777 for CURB-65 and 0.560 for PSI²². This study proposed the FluA-p score as a novel approach to predict mortality in patients with influenza A-related pneumonia, achieving an AU-ROC of 0.908²². However, despite its high AU-ROC, the FluA-p score relies on laboratory variables as risk parameters, similar to the SOFA score. This reliance may present challenges when incorporating it into environments with limited resources or assessing severity in epidemiological studies that often lack access to laboratory data.

When analyzing medical research data with categorical outcomes, it's crucial to consider performance metrics. While the ROC curve is commonly used to assess test performance, dealing with imbalanced datasets can distort results. Combining ROC and Precision-Recall (PR) curves along with their respective AUC measurements (AU ROC and AU PR) is recommended to address this issue. Surprisingly, PR curves are often overlooked in diagnostic performance studies in influenza patients. To fill this gap, we assessed ISS's discriminative ability using ROC and PR curves. The ISS demonstrated strong discriminative performance in the test set, achieving AU ROC and AU PR values exceeding 70%. Remarkably, these results were achieved without using any laboratory parameters and by including patients from various sites to minimize bias stemming from local practices.

Estimating disease severity involves assessing the probability of significant clinical events among those individuals who are at risk but have not yet experienced any at the start of the observation period. This probability is determined by a set of parameters, some of which can be modified while others cannot. The severity assessment must consider modifiable and non-modifiable parameters, with the latter being the most important. Modifiable parameters may not always lead to a decrease in the risk of clinical events, but they can indicate how an intervention, disease progression, or host response will affect the baseline risk. Even if the modifiable parameters do not indicate a high risk, this does not necessarily mean the individual's baseline risk has changed. Instead, it reflects the natural history of the disease and how it interacts with the host. Therefore, individuals must be stratified based on modifiable and non-modifiable risks. Our tool considers this, allowing us to gain insight into the progression of the disease and the host's response to the infection. At the same time, it recognizes and respects the individual's particularities and dynamic responses.

To ensure the most effective risk management strategies, ISS utilizes a 3-threshold system to differentiate between low- and high-risk patients when measuring the likelihood of major clinical events. These thresholds were specifically chosen to ensure a high sensitivity and specificity rate exceeding 80%. This 3-tier system facilitates risk-based treatment protocols and permits a more centralized and cost-effective care distribution. For medical personnel, this allows for treatments and services to be tailored to the specific needs of patients, enabling higher-quality care and more successful patient outcomes. During epidemiological studies, the use of ISS enables the identification of the disease severity and its association with a particular strain. It can also be used as an adjustment measure when assessing the effectiveness of interventions.

Our study has several limitations that must be considered. Firstly, the data we used was obtained from hospital surveillance, which inherently limits our ability to account for specific institutional protocols. Our data consisted of the initial assessment of the patient within the hospital, which reflects the patient's condition immediately after being hospitalized. While using this approach, we might have lost the ability to capture the scale's performance in predicting the need for hospitalization. However, we could still track all significant clinical events that occurred afterward. Also, we have not had lower tract samples, meaning we may not have identified some people with severe disease, thus skewing the results. Moreover, our tool's reliance on non-modifiable parameters limits its capacity to be used as an evolutionary or responsive variable during hospitalization, as it does not concentrate

on physiological parameters. Unsurprisingly, the absence of modifiable parameters, such as vital signs, laboratory, and radiological variables, impacted the discriminatory performance of our tool. Future studies should include these variables to enhance its performance while ensuring a balance with non-modifiable parameters. Some rare characteristics are likely underrepresented in our cohort, leading to an insufficient data sample to identify them as significant factors influencing ISS. Nonetheless, they could be relevant in other settings, also warranting further validation. Lastly, ISS comprises ten variables and individual scores, which can be challenging to remember in clinical practice. Ideally, future work could focus on developing automated means of collecting and calculating the ISS to support its use in research and clinical settings. Despite these drawbacks, our dataset was sizeable, multi-centric, and included a wide variety of people who are generally not included in these kinds of studies, such as individuals on the extremes of age with or without comorbidities.

In summary, the ISS is a tool used to assess the severity of influenza infection. It considers the patient's symptoms and non-modifiable parameters to predict the likelihood of MCE without requiring lab tests other than confirming influenza infection. This can help direct protocols and policies to those at greater risk of MCE, such as older adults, women, smokers, CPD patients, diabetics, and those not vaccinated against influenza. It also highlights the importance of considering multiple factors (and their intersection) contributing to a person's risk rather than individual factors considered singly.

Additionally, the tool can raise awareness of *important symptoms* that predict worse outcomes, namely (in order of importance) shortness of breath, sputum production, and coughing, as well as *important clinical features* such as underlying health conditions and functional status. Notably, the most important clinical factors identified here were underlying chronic lung disease, shortness of breath, and baseline functional impairment with the requirement for support in activities of daily living. These are relevant and readily identifiable factors that impact clinical prognostication and decision-making. For example, a 70-year-old patient with baseline functional impairment who presents with shortness of breath is at high risk for MCE from influenza, and this can be communicated to the patient and family early in their admission.

We believe ISS will be essential for public health systems to monitor the effects of public health protocols on clinical outcomes and establish efficient surveillance measures. Further research is necessary to explore the utility of ISS in different clinical settings and its capacity to predict mortality. Ultimately, the ISS may prove to be a valuable metric for assessing and improving influenza-related health outcomes, contributing to the betterment of public health.

Data availability

The datasets generated and/or analysed during the current study are not publically available due to the confidential nature of the data obtained from patients, however, datasets are available through the corresponding author on reasonable request.

Received: 16 October 2023; Accepted: 17 July 2024

Published online: 08 August 2024

References

- Lina, B. *et al.* Complicated hospitalization due to influenza: Results from the Global Hospital Influenza Network for the 2017–2018 season. *BMC Infect. Dis.* **20**, 465. <https://doi.org/10.1186/s12879-020-05167-4> (2020).
- Andrew, M. K. *et al.* Age differences in comorbidities, presenting symptoms, and outcomes of influenza illness requiring hospitalization: A worldwide perspective from the global influenza hospital surveillance network. *Open Forum Infect. Dis.* **10**, 244. <https://doi.org/10.1093/ofid/ofad244> (2023).
- Pott, H. *et al.* Vaccine Effectiveness of non-adjuvanted and adjuvanted trivalent inactivated influenza vaccines in the prevention of influenza-related hospitalization in older adults: A pooled analysis from the Serious Outcomes Surveillance (SOS) Network of the Canadian Immunization Research Network (CIRN). *Vaccine* <https://doi.org/10.1016/j.vaccine.2023.08.070> (2023).
- Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. *Lancet* **391**, 1285–1300. [https://doi.org/10.1016/S0140-6736\(17\)33293-2](https://doi.org/10.1016/S0140-6736(17)33293-2) (2018).
- WHO. WHO launches new global influenza strategy. <https://www.who.int/news/item/11-03-2019-who-launches-new-global-influenza-strategy> (2019).
- Uyeki, T. M., Hui, D. S., Zambon, M., Wentworth, D. E. & Monto, A. S. Influenza. *Lancet* **400**, 693–706. [https://doi.org/10.1016/S0140-6736\(22\)00982-5](https://doi.org/10.1016/S0140-6736(22)00982-5) (2022).
- Adams, K. *et al.* A literature review of severity scores for adults with influenza or community-acquired pneumonia—implications for influenza vaccines and therapeutics. *Hum. Vaccin. Immunother.* **17**, 5460–5474. <https://doi.org/10.1080/21645515.2021.1990649> (2021).
- Liu, J. W., Lin, S. H., Wang, L. C., Chiu, H. Y. & Lee, J. A. Comparison of antiviral agents for seasonal influenza outcomes in healthy adults and children: A systematic review and network meta-analysis. *JAMA Netw. Open* **4**, e2119151. <https://doi.org/10.1001/jamanetworkopen.2021.19151> (2021).
- Minozzi, S. *et al.* Comparative efficacy and safety of vaccines to prevent seasonal influenza: A systematic review and network meta-analysis. *EClinicalMedicine* **46**, 101331. <https://doi.org/10.1016/j.eclinm.2022.101331> (2022).
- Andrew, M. K. *et al.* The importance of frailty in the assessment of influenza vaccine effectiveness against influenza-related hospitalization in elderly people. *J. Infect. Dis.* **216**, 405–414. <https://doi.org/10.1093/infdis/jix282> (2017).
- Andrew, M. K. *et al.* Influenza surveillance case definitions miss a substantial proportion of older adults hospitalized with laboratory-confirmed influenza: A report from the Canadian Immunization Research Network (CIRN) Serious Outcomes Surveillance (SOS) Network. *Infect. Control Hosp. Epidemiol.* **41**, 499–504. <https://doi.org/10.1017/ice.2020.22> (2020).
- ElSherif, M. L. *et al.* to 2015 to characterize the burden of respiratory syncytial virus disease in canadian adults >=50 years of age hospitalized with acute respiratory illness. *Open Forum Infect Dis.* **10**(ofad315), 2023. <https://doi.org/10.1093/ofid/ofad315> (2012).
- Nichols, M. K. *et al.* Influenza vaccine effectiveness to prevent influenza-related hospitalizations and serious outcomes in Canadian adults over the 2011/12 through 2013/14 influenza seasons: A pooled analysis from the Canadian Immunization Research Network (CIRN) Serious Outcomes Surveillance (SOS Network). *Vaccine* **36**, 2166–2175. <https://doi.org/10.1016/j.vaccine.2018.02.093> (2018).

14. McNeil, S. A. *et al.* Interim estimates of 2014/15 influenza vaccine effectiveness in preventing laboratory-confirmed influenza-related hospitalisation from the Serious Outcomes Surveillance Network of the Canadian Immunization Research Network, January 2015. *Euro Surveill* **20**, 21024. <https://doi.org/10.2807/1560-7917.es2015.20.5.21024> (2015).
15. Wang, R. & Taubenberger, J. K. Methods for molecular surveillance of influenza. *Expert Rev. Anti Infect. Ther.* **8**, 517–527. <https://doi.org/10.1586/eri.10.24> (2010).
16. Nichols, M. K. *et al.* The impact of prior season vaccination on subsequent influenza vaccine effectiveness to prevent influenza-related hospitalizations over 4 influenza seasons in Canada. *Clin. Infect. Dis.* **69**, 970–979. <https://doi.org/10.1093/cid/ciy1009> (2019).
17. Dobbin, K. K. & Simon, R. M. Optimally splitting cases for training and testing high dimensional classifiers. *BMC Med. Genom.* **4**, 31. <https://doi.org/10.1186/1755-8794-4-31> (2011).
18. Nguyen, Q. H. *et al.* Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Math. Probl. Eng.* **2021**, 4832864. <https://doi.org/10.1155/2021/4832864> (2021).
19. Austin, P. C. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* **28**, 3083–3107. <https://doi.org/10.1002/sim.3697> (2009).
20. Davis, J. & Goadrich, M. *The Relationship Between Precision-Recall and ROC Curves* (Springer, 2006).
21. Adeniji, K. A. & Cusack, R. The Simple Triage Scoring System (STSS) successfully predicts mortality and critical care resource utilization in H1N1 pandemic flu: A retrospective analysis. *Crit. Care* **15**, R39. <https://doi.org/10.1186/cc10001> (2011).
22. Chen, L., Han, X., Li, Y. L., Zhang, C. & Xing, X. FluA-p score: A novel prediction rule for mortality in influenza A-related pneumonia patients. *Respir. Res.* **21**, 109. <https://doi.org/10.1186/s12931-020-01379-z> (2020).
23. Choi, W. I. *et al.* Clinical characteristics and outcomes of H1N1-associated pneumonia among adults in South Korea. *Int. J. Tuberc. Lung Dis.* **15**, 270–275 (2011).
24. Jain, S. *et al.* Influenza-associated pneumonia among hospitalized patients with 2009 pandemic influenza A (H1N1) virus—United States, 2009. *Clin. Infect. Dis.* **54**, 1221–1229. <https://doi.org/10.1093/cid/cis197> (2012).
25. Justel, M. *et al.* IgM levels in plasma predict outcome in severe pandemic influenza. *J. Clin. Virol.* **58**, 564–567. <https://doi.org/10.1016/j.jcv.2013.09.006> (2013).
26. Miller, A. C. *et al.* Influenza A 2009 (H1N1) virus in admitted and critically ill patients. *J. Intensive Care Med.* **27**, 25–31. <https://doi.org/10.1177/0885066610393626> (2012).
27. Nicolini, A., Ferrera, L., Rao, F., Senarega, R. & Ferrari-Bravo, M. Chest radiological findings of influenza A H1N1 pneumonia. *Rev. Port Pneumol.* **18**, 120–127. <https://doi.org/10.1016/j.rppneu.2011.12.008> (2012).
28. Oh, W. S. *et al.* A prediction rule to identify severe cases among adult patients hospitalized with pandemic influenza A (H1N1) 2009. *J. Korean Med. Sci.* **26**, 499–506. <https://doi.org/10.3346/jkms.2011.26.4.499> (2011).
29. Papadimitriou-Oliveris, M. *et al.* Predictors of mortality of influenza virus infections in a Swiss Hospital during four influenza seasons: Role of quick sequential organ failure assessment. *Eur. J. Intern. Med.* **74**, 86–91. <https://doi.org/10.1016/j.ejim.2019.12.022> (2020).
30. Talmor, D., Jones, A. E., Rubinson, L., Howell, M. D. & Shapiro, N. I. Simple triage scoring system predicting death and the need for critical care resources for use during epidemics. *Crit. Care Med.* **35**, 1251–1256. <https://doi.org/10.1097/01.CCM.0000262385.95721.CC> (2007).
31. Yang, S. Q. *et al.* Influenza pneumonia among adolescents and adults: A concurrent comparison between influenza A (H1N1) pdm09 and A (H3N2) in the post-pandemic period. *Clin. Respir. J.* **8**, 185–191. <https://doi.org/10.1111/crj.12056> (2014).
32. Zhou, J. *et al.* A functional variation in CD55 increases the severity of 2009 pandemic H1N1 influenza A virus infection. *J. Infect. Dis.* **206**, 495–503. <https://doi.org/10.1093/infdis/jis378> (2012).
33. Zhu, L. *et al.* High level of neutrophil extracellular traps correlates with poor prognosis of severe influenza A infection. *J. Infect. Dis.* **217**, 428–437. <https://doi.org/10.1093/infdis/jix475> (2018).
34. Pereira, J. M. *et al.* Severity assessment tools in ICU patients with 2009 influenza A (H1N1) pneumonia. *Clin. Microbiol. Infect.* **18**, 1040–1048. <https://doi.org/10.1111/j.1469-0691.2011.03736.x> (2012).
35. Zimmerman, O. *et al.* C-reactive protein serum levels as an early predictor of outcome in patients with pandemic H1N1 influenza A virus infection. *BMC Infect. Dis.* **10**, 288. <https://doi.org/10.1186/1471-2334-10-288> (2010).
36. Riquelme, R. *et al.* Predicting mortality in hospitalized patients with 2009 H1N1 influenza pneumonia. *Int. J. Tuberc. Lung Dis.* **15**, 542–546. <https://doi.org/10.5588/ijtld.10.0539> (2011).
37. Commons, R. J. & Denholm, J. Triage pandemic flu: Pneumonia severity scores are not the answer. *Int. J. Tuberc. Lung Dis.* **16**, 670–673. <https://doi.org/10.5588/ijtld.11.0446> (2012).

Acknowledgements

The authors thank the dedicated Serious Outcomes Surveillance Network monitors, whose tremendous efforts made this study possible. The results of this study have been accepted for presentation at the Options for the Control of Influenza XII—2024.

Author contributions

HP and MKA conceived of the study. HP conducted the analyses and wrote the initial draft. MKA, SAM, JL, TFH and ME provided critical insight into interpretation of analyses and contributed manuscript revisions. All authors approved the final manuscript.

Funding

Funding for this study was provided by the Public Health Agency of Canada and the Canadian Institutes of Health Research to the Canadian Immunization Research Network and during some seasons through an investigator-initiated Collaborative Research Agreement with GlaxoSmithKline Biologicals SA. Additional funding was obtained through a grant from the Foundation for Influenza Epidemiology under the auspices of the Fondation de France. The authors are solely responsible for the final content and interpretation. The authors received no financial support or other forms of compensation related to the development of the manuscript, and the funders were not involved in analyses, interpretation of findings, or manuscript writing.

Competing interests

HP reports grant funding from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001. TFH reports grants from Pfizer and GSK outside the submitted work. SAM reports grants and payments from Pfizer, GSK, Merck, Novartis, and Sanofi outside the submitted work. MK Andrew reports grant funding from the GSK group of companies, Pfizer and Sanofi Pasteur, and honoraria for

past ad-hoc advisory activities from Pfizer, Sanofi, and Seqirus, all unrelated to the present manuscript. JL and ME report no conflicts of interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-67931-9>.

Correspondence and requests for materials should be addressed to H.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

the Serious Outcomes Surveillance (SOS) Network of the Canadian Immunization Research Network (CIRN)

Guy Boivin⁶, Sylvie Trottier⁶, Francisco Diaz-Mitoma⁷, Chris Verschoor⁷, Grant Stiver⁸, William Bowie⁸, Karen Green⁹, Allison McGeer⁹, Jennie Johnstone¹⁰, Mark Loeb¹⁰, Kevin Katz¹¹, Phillippe Lagacé-Wiens¹², Bruce Light¹², Anne McCarthy¹³, Andre Poirier¹⁴, Jeff Powis¹⁵, David Richardson¹⁶, Makeda Semret¹⁷, Stephanie Smith¹⁸, Geoff Taylor¹⁸, Daniel Smyth¹⁹, Louis Valiquette²⁰ & Duncan Webster²¹

⁶Centre Hospitalier Universitaire de Québec, Quebec, QC, Canada. ⁷Health Sciences North Research Institute, Sudbury, ON, Canada. ⁸University of British Columbia, Vancouver, BC, Canada. ⁹Mount Sinai Hospital, Toronto, ON, Canada. ¹⁰McMaster University, Hamilton, ON, Canada. ¹¹North York General Hospital, Toronto, ON, Canada. ¹²St. Boniface Hospital, Winnipeg, MB, Canada. ¹³The Ottawa Hospital, Ottawa, ON, Canada. ¹⁴Centre Intégré Universitaire de santé et services sociaux, Quebec, QC, Canada. ¹⁵Michael Garron Hospital, Toronto, ON, Canada. ¹⁶William Osler Health System, Brampton, ON, Canada. ¹⁷McGill University, Montreal, QC, Canada. ¹⁸University of Alberta Hospital, Edmonton, AB, Canada. ¹⁹The Moncton Hospital, Moncton, NB, Canada. ²⁰Université de Sherbrooke, Sherbrooke, QC, Canada. ²¹Horizon Health, Saint John, NB, Canada.