A Systematic Review of Clinical Decision Rules for the Diagnosis of Influenza

Mark H. Ebell, MD, MS
Anna Afonso, BS

Department of Epidemiology and Biostatistics, College of Public Health, The University of Georgia, Athens, Georgia

ABSTRACT

PURPOSE In this study, we assessed whether multivariate models and clinical decision rules can be used to reliably diagnose influenza.

METHODS We conducted a systematic review of MEDLINE, bibliographies of relevant studies, and previous meta-analyses. We searched the literature (1962-2010) for articles evaluating the accuracy of multivariate models, clinical decision rules, or simple heuristics for the diagnosis of influenza. Each author independently reviewed and abstracted data from each article; discrepancies were resolved by consensus discussion. Where possible, we calculated sensitivity, specificity, predictive value, likelihood ratios, and areas under the receiver operating characteristic curve.

RESULTS A total of 12 studies met our inclusion criteria. No study prospectively validated a multivariate model or clinical decision rule, and no study performed a split-sample or bootstrap validation of such a model. Simple heuristics such as the so-called fever and cough rule and the fever, cough, and acute onset rule were each evaluated by several studies in populations of adults and children. The areas under the receiver operating characteristic curves were 0.70 and 0.79, respectively. We could not calculate a single summary estimate, however, as the diagnostic threshold varied among studies.

CONCLUSIONS The fever and cough, and the fever, cough, and acute onset heuristics have modest accuracy, but summary estimates could not be calculated. Further research is needed to develop and prospectively validate clinical decision rules to identify patients requiring testing, empiric treatment, or neither.

Ann Fam Med 2011;9:69-77. doi:10.1370/afm.1192.

INTRODUCTION

Accurate diagnosis of influenza is important for several reasons. If the probability of disease exceeds the treatment threshold or is below the testing threshold, no further testing is needed. If office-based testing is performed, its interpretation depends on the pretest probability of disease. And, although a systematic review found that neuraminidase inhibitors are of only modest benefit in patients with undifferentiated influenza-like illness, greater benefit was seen in patients who actually had laboratory-confirmed influenza.[1] Accurate diagnosis is also helpful because it enables a more accurate prognosis, implicitly rules out other diagnoses, and guides patient education; however, 2 previous meta-analyses[2,3] showed that individual findings on the history and physical examination have only modest accuracy for the clinical diagnosis of influenza (Table 1). These studies did find that certain combinations of variables, such as the combination of fever plus cough, the combination of fever, cough, and acute onset,[3] and the combination of fever plus presentation within 3 days,[2] had positive likelihood ratios for influenza between 4.0 and 5.4. These results suggest that clinical decision rules (CDRs) that integrate data from several clinical findings and are developed using multivariate methods might be helpful.

Economic analyses have shown that diagnostic testing is cost-effective only when the pretest probability of influenza is low or intermediate,

Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Mark H. Ebell, MD, MS
Department of Epidemiology and Biostatistics
College of Public Health
The University of Georgia
N122A Coverdell Building
Athens, GA 30602
ebell@uga.edu

whereas empiric therapy may be appropriate for patients seeking care within 36 hours of symptom onset if the pretest probability is high.[4-6] CDRs to accurately identify patients who are at low, moderate, or high risk of influenza could thus identify patients for whom testing or empiric therapy may be appropriate, and others who need neither. Such identification would help physicians avoid unnecessary testing and treatment, and potentially reduce health care costs. The goal of this study was therefore to identify and evaluate the accuracy and validity of existing CDRs for the diagnosis of influenza.

Table 1. Results From a Previous Systematic Review of the History and Physical Examination for the Diagnosis of Influenza

	Likelihood Ratio for Influenza	
Clinical Finding	**Positive**	**Negative**
Rigors	7.2	0.9
Sweating	2.9	0.6
Being confined to bed	2.5	0.5
Inability to cope with daily activities	2.3	0.4
Fever (subjective)	1.7	0.5
Absence of systemic symptoms	1.5	0.4
Headache	1.3	0.6
Cough	1.3	0.4
Myalgia	1.3	0.6
Nasal congestion	1.2	0.7
Chills	1.1	0.7
Sore throat	1.1	0.9
Sputum	1.1	0.9

Note: Data from Ebell et al.[2]

METHODS

We undertook a systematic review of studies reported between 1962 and 2010 evaluating combinations of symptoms and CDRs for the diagnosis of influenza. As this was a secondary literature review, institutional review board approval was not required. We defined a CDR as a point score, equation, or algorithm developed using multivariate methods. We limited our search to CDRs using elements of the history and physical examination, including vital signs. We performed a multipronged search of the relevant medical literature. To be included in our review, a study had to (1) provide data on the accuracy of a combination of symptoms or CDR using elements of the history and physical examination in patients with respiratory tract infection, (2) enroll patients prospectively using a cohort (not case-control) design, and (3) use an adequate reference standard. We defined an adequate reference standard as any reference laboratory test for the diagnosis of influenza. As point-of-care tests are not sufficiently sensitive[3] to be appropriate reference standards, we excluded studies taking that approach.

Table 2. Characteristics of Included Studies

Study and Year	Population and Season(s)	N	Influenza Prevalence, %
Govaert et al,[13] 1998	Average-risk, unvaccinated primary care patients older than 60 y with ILI during 1991-1992 flu season	1,838 (logistic regression); 1,791 (heuristics)	6.6
Carrat et al,[23] 1999	Primary care patients older than 1 y with ILI, RTI syndrome, and/or fever >38°C without signs of other infections during 1995-1996 flu season	600	26
Monto et al,[11] 2000	Adolescent or older with fever and ≥2 other symptoms (headache, myalgia, cough, or sore throat) during fall or winter of 1994-1998	3,744	66
Boivin et al,[12] 2000	Patients with ILI (fever ≥37.8°C and ≥2 other symptoms) seeking care at an outpatient clinic during 1998-1999 flu season	100	79
van Elden et al,[10] 2001	Primary care patients with fever (≥38°C), ≥1 constitutional symptom (malaise, headache, myalgia, chills), and ≥1 respiratory symptom (coryza, sneeze, cough, sore throat, hoarseness) seeking care within 48 hr of onset during 1997-1998 flu season	81	51
Zambon et al,[19] 2001	Persons aged 12 y or older seen within 48 hr of onset of ILI as part of a multicenter international clinical trial of an influenza drug	1,033	77 (any test positive); 67 (all 3 tests positive)
Walsh et al,[21] 2002	Inpatients older than age 65 y or with underlying cardiopulmonary conditions with a respiratory diagnosis during 1999-2000 flu season	332	18
Friedman and Attia,[20] 2004	Children (0-17 y) seen in the ED with suspected ILI (fever and coryza, cough, headache, sore throat, or muscle aches) during 2002 flu season	128	35
Senn et al,[7] 2005	Persons with ILI seeking care at an outpatient clinic during 1999-2000 flu season	201	52
Stein et al,[8] 2005	Consecutive adults with RTI in past 3 wk with cough, sinus pain, congestion/rhinorrhea, sore throat, or fever seen in the ED during 2002 flu season	258	21
Ohmit and Monto,[24] 2006	Study 1 (zanamivir): children aged 5-12 y with fever and duration of illness <36 hr during flu season Study 2 (oseltamivir): children aged 1-12 y with fever and cough or nasal symptoms, and duration <48 hr during flu season	Study 1: 468 Study 2 (1-4 y): 255 Study 2 (5-12 y): 221	Study 1: 74 Study 2: 67
van den Dool et al,[22] 2008	Adult inpatients in the general medicine, pulmonology, and infectious disease wards of a tertiary care hospital during 2006-2007 flu season	264	8.7

DIF = direct immunofluorescence; ED = emergency department; ELISA = enzyme-linked immunosorbent assay; ILI = influenza-like illness; NPV = negative predictive value; PCR = polymerase chain reaction; RSV = respiratory syncytial virus; RTI = respiratory tract infection.

Our initial search of PubMed used the following strategy: (influenza[Title/Abstract]) AND (diagnosis[Title/Abstract]) AND (multivariate[Title/Abstract] OR logistic[Title/Abstract] OR "prediction model"[Title/Abstract] OR "decision model"[Title/Abstract] OR "decision rule"[Title/Abstract] OR "clinical model"[Title/Abstract] OR "clinical rule"[Title/Abstract]) LIMITED TO abstracts, human.

This search yielded 45 studies, of which 7 appeared potentially relevant and were reviewed in more detail.[7-13] Next, we searched the Clinical Queries feature of PubMed using parameters for "Clinical prediction guide (narrow)" and "influenza." This search yielded 181 articles, of which 5 were possibly relevant,[14-18] but on closer review, all dealt with prognosis of patients with influenza rather than diagnosis. Next, we used the "Related articles" feature of PubMed's Clinical Queries service to search for studies indexed using similar key words to a particularly relevant study, that of Stein and colleagues.[8] This process yielded 244 articles, of which 9 were potentially relevant, and 5 had not been found using other search strategies.[2,3,19-21] A search of the references of the 12 articles deemed potentially relevant identified 10 articles for closer review, and 3 additional studies for inclusion.[22-24] Finally, we searched Google Scholar using the terms "influenza clinical decision," but did not identify any new studies among the first 200 returned search results. The Cochrane Controlled Trials Register was not searched as it is limited to studies of therapy.

Studies were abstracted in parallel and discrepancies were resolved by consensus. Study design elements (eg, size of the study, reference standard) were evaluated for each study to assess its validity. We queried authors where data were missing for multivariate models, such as the cutoff for an abnormal score, but had limited success. The search was initially performed on January 26, 2010, and was repeated on July 8, 2010, as part of the revision of this article. One potential study[25] was identified during that final search, but was not included because it did not meet the inclusion criterion of prospective data collection.

RESULTS

Our preliminary search identified 12 studies that on review of the abstract appeared to meet the inclusion criteria. We excluded several articles after review of the full publication because they used a case-control design,[26] did not gather original data,[2] or provided information only on the predictive accuracy of a white blood cell count,[9] which we did not consider a CDR. This process left 9 studies for the systematic review that met our inclusion criteria. A review of bibliographies of the 12 studies initially deemed relevant identified an additional 10 articles that appeared to fit inclusion criteria, of which 3 were included based on a review of the full article. We ultimately included a total of 12 studies in the systematic review (Table 2).

Country	Reference Standard	Comment
Netherlands	Increase in antibody titer	Any symptoms reported during study period were included; part of randomized trial of vaccination.
France	DIF and ELISA; if disagreement, further testing with PCR and culture	Separate models created for H1N1 and H3N2 subtypes, and for all influenza.
North America, Europe, southern hemisphere	Culture or increase in antibody titer (in some studies PCR or immunofluorescence)	High prevalence of influenza. Data from 6 randomized trials of zanamivir vs placebo.
Canada	PCR and culture	–
Netherlands	PCR and culture	Only PCR results were used for model development.
Europe, North America	PCR, culture, or increase in antibody titer	Reference standard was abnormal finding on all 3 tests.
United States	Culture, rapid antigen test, increase in antibody titer, or PCR	A model with 2 variables (temperature >38°C and absence of dyspnea) reportedly did not discriminate well; results not reported by authors.
United States	Culture	–
Switzerland	Culture	Cutoff for logistic model presumably probability of influenza >50%.
United States	PCR	Validated previously developed fever and cough rule.[11]
United States	Culture and/or 4-fold increase in antibody titer (or PCR in zanamivir trial only)	Validated previously developed fever and cough rule.[11] RSV infection excluded. Unable to calculate sensitivity, specificity, NPV from data for fever and cough rule.
Netherlands	PCR	Included all patients regardless of symptoms; validated previously developed fever and cough rule.[11]

Table 3. Quality Assessment of Included Studies Based on the QUADAS Tool for Studies of Diagnostic Accuracy[27]

Criterion	Govaert et al,[13] 1998	Carrat et al,[23] 1999	Monto et al,[11] 2000	Boivin et al,[12] 2000	van Elden, et al,[10] 2001	Zambon et al,[19] 2001	Walsh et al,[21] 2002	Friedman and Attia,[20] 2004	Senn et al,[7] 2005	Stein et al,[8] 2005	Ohmit and Monto,[24] 2006	van den Dool et al,[22] 2008
1. Was the spectrum of participants representative of the patients who will receive the test in practice?	Y[a]	Y [a]	Y[a]	Y[a]	Y [a]	Y[a]	N[a]	Y[a]	Y [a]	Y[a]	Y[a]	N[a]
2. Were selection criteria clearly described?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y
3. Was the reference standard likely to classify the target condition correctly?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4. Was the period between performance of the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the 2 tests?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
5. Did the whole sample or a random selection of the sample receive verification using the reference standard?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
6. Did participants receive the same reference standard regardless of the index test result?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
7. Was the reference standard independent of the index test?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8. Was the execution of the index test described in sufficient detail to permit its replication?	Y[b]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
9. Were the index test results interpreted without knowledge of the results of the reference test?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10. Were the reference standard results interpreted without knowledge of the results of the index test?	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]	U[c]
11. Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?	U	U	U	U	U	U	U	U	U	U	U	U
12. Were uninterpretable, indeterminate, or intermediate test results reported?	U	Y	Y	Y	Y	Y	Y	U	Y	Y	U	Y
13. Were withdrawals from the study explained?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

N = no; NA = not applicable; QUADAS = Quality Assessment of Diagnostic Accuracy; U = unknown; Y = yes.

[a] Participants were classified as being representative of patients who will receive the test in practice if they were undifferentiated outpatients with symptoms of respiratory tract infection. Participants were classified as not being representative if they were a subset of inpatients.

[b] No temperature cutoff was given for definition of fever.

[c] Although not explicitly stated, it is unlikely that diagnostic laboratories verifying the presence of influenza had access to the clinical data of each patient.

The prevalence of influenza in the included studies ranged from 6.6% to 79%, and studies had widely varying inclusion criteria. Some included only children,[20] whereas others included only older adults.[13,21] Most were conducted in the outpatient setting, although 1 study was limited to emergency department patients[20] and 2 studies were limited to inpatients.[21,22] One study[13] prospectively followed a large group of elderly patients through the flu season and recorded symptoms of those who sought care (many of the patients never had any respiratory symptoms during the study period).

The studies were generally of good quality (Table 3), in part because of our inclusion criteria. Although a number of studies developed multivariate models, none of the studies included any kind of prospective validation of these models, such as a split-sample or bootstrap procedure (Table 4).

Variables common to 2 or more studies included fever, cough, headache, and vaccination status. Some studies used variables that were difficult to generalize, such as the week of the study,[7] or used definitions that were not reproducible, such as "increased influenza activity."[10] A number of studies did not report details of the multivariate model such as the constant, ß coefficients, odds ratios, and the cutoff for defining a positive test,[7,8,10,12,13] while others did not report the accuracy

Table 4. Multivariate Models, Point Scores, and Symptom Combinations Reported by Included Studies

Study	Model	Accuracy
Govaert et al,[13] 1998	Variable: OR	Not reported
	Cough: 5.25	
	Fever: 2.18	
	Vaccinated: 0.56	
	Constant: 0.041	
Govaert et al,[13] 1998	Symptom count (fever, cough, acute onset, malaise, rigors or chills, myalgia, headache, sore throat)	Not reported other than data shown
	Score (n): % influenza	
	0 (1,155): 3.2	
	1-2 (145): 6.2	
	≥3 (491): 15	
Carrat et al,[23] 1999	Variable: OR	Goodness of fit: $P = .98$
	All influenza	
	Temperature >38.2˚C: 2.45	
	Rhinorrhea: 1.83	
	Temperature >38.9˚C, respiratory signs, and stiffness or myalgia	PPV: 40%; NPV: 80%
	Temperature >37.7˚C and cough or sore throat	PPV: 30%; NPV: 86%
	Any 3 of temperature >37.7˚C, cough, chills, moderate/severe fatigue, cervical or dorsal pain, pharyngitis, and another case at patient's home	PPV: 27%; NPV: 91%
Monto et al,[11] 2000	Variable: OR (95% CI)	Accuracy of multivariate model not reported
	Fever >37.7˚C: 3.26 (2.75-3.87)	
	Cough: 2.85 (2.21-3.68)	
	Nasal congestion: 1.98 (1.54-2.54)	
	Age ≥55 y: 1.60 (1.18-2.16)	
	Weakness: 1.54 (1.07-2.22)	
	Onset >36 hr: 1.53 (1.24-1.90)	
	Loss of appetite: 1.43 (1.10-1.86)	
	Sex, male: 1.27 (1.08-1.50)	
	Sore throat: 0.72 (0.57-0.91)	
Boivin et al,[12] 2000	Variable: OR (95% CI)	PPV: 87%; NPV: 39%
	Cough: 6.68 (1.4-34.1)	
	Temperature ≥38˚C: 3.06 (1.35-8.02)	
van Elden et al,[10] 2001	Period of increased influenza activity, cough, headache at onset, feverishness at onset, and not vaccinated	PPV: 75%; NPV: 80%
	During an outbreak, abrupt onset (<5 days), temperature >38˚C, and at least 1 of cough, coryza, headache, retrosternal pain, or myalgia	PPV: 52%
	At least 4 of sudden onset, contact with influenza, fever, cough, chills, malaise, myalgia, or hyperemic mucous membranes of the nose and throat (≥6 required if not in outbreak)	PPV: 54%; NPV: 85%
Senn et al,[7] 2005	Week of consultation (49-50 vs ≥51), duration of symptoms (≤48 hr vs >48 hr), temperature >37.8˚C, and cough	AUC = 0.74
		Sensitivity: 80%; specificity: 59%
		LR+: 1.95; LR–: 0.34
		PPV: 67%; NPV: 73%
		LR+/LR– = 5.7
Ohmit and Monto,[24] 2006	Study 1, zanamivir—variable: OR (95% CI)	Accuracy not reported; only variables significant at $P < .05$ level included
	Age, y: 1.11 (1.00-1.23)	
	Fever: 2.67 (1.66-4.30)	
	Cough: 5.19 (2.66-10.10)	
	Myalgia: 0.61 (0.38-0.99)	
	Sore throat: 0.41 (0.24-0.70)	
	Study 2, oseltamivir (age 1-4 y)—variable: OR (95% CI)	
	Myalgia: 2.32 (1.22-4.39)	
	Study 2, oseltamivir (age 5-12 y)—variable: OR (95% CI)	
	Cough: 10.94 (2.90-40.80)	
	Headache: 2.24 (1.15-4.37)	

AUC = area under the receiver operating characteristic curve; CI = confidence interval; LR+ = positive likelihood ratio, LR– = negative likelihood ratio; NPV = negative predictive value; OR = odds ratio; PPV = positive predictive value.

Table 5. Accuracy of Simple Heuristics for the Clinical Diagnosis of Influenza

Rule and Study[a]	Heuristic	Sensitivity, %	Specificity, %	PPV, %	NPV, %	LR+	LR–	LR+/ LR–
Fever and cough rule								
Stein et al,[8] 2005	Fever (≥37.8˚C) and cough	40	92	58	84	5.1	0.7	7.3
Govaert et al,[13] 1998	Fever (>38˚C) and cough	30	94	26	95	5.0	0.74	6.8
Boivin et al,[12] 2000[b]	Fever (≥37.8˚C) and cough	78	55	87	39	1.7	0.4	4.3
Monto et al,[11] 2000	Fever (≥37.8˚C) and cough	64	67	79	49	1.94	0.54	3.6
van den Dool et al,[22] 2008	Fever (>38.3˚C) and cough	42	90	26	95	4.2	0.64	6.5
Fever, cough, and acute onset rule								
Stein et al,[8] 2005	Fever (≥37.8˚C), cough, and duration ≤48 hr	75	89	65	93	6.5	0.3	21.7
Walsh et al,[21] 2002	Fever (≥38˚C), cough, and duration <7 days	78	73	47	91	2.9	0.3	9.7
Govaert et al,[13] 1998	Fever (>38˚C), cough, and acute onset	27	95	30	95	5.9	0.76	7.8
Monto et al,[11] 2000	Fever (≥37.8°C), cough, and acute onset	63	68	77	51	1.95	0.54	3.6
Cough, headache, and pharyngitis rule								
Friedman and Attia,[20] 2004[c]	Cough, headache, and pharyngitis	80	78	77	81	3.7	0.26	14.2

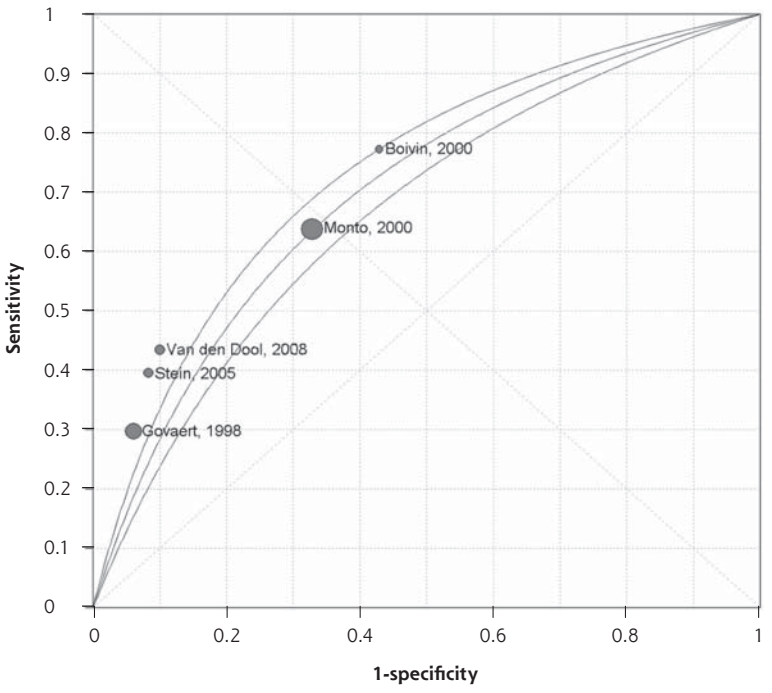
ILI = influenza-like illness; LR– = negative likelihood ratio; LR+ = positive likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; RTI = respiratory tract infection.

[a] Study population and prevalence of influenza are shown in Table 2.

[b] Odds ratio in logistic regression model was 6.7 for cough and 3.1 for temperature greater than 38˚C.

[c] Odds ratio in logistic regression model was 7.2 for cough, 4.3 for headache, and 3.9 for pharyngitis.

Figure 1. Summary ROC curve for the fever and cough rule.



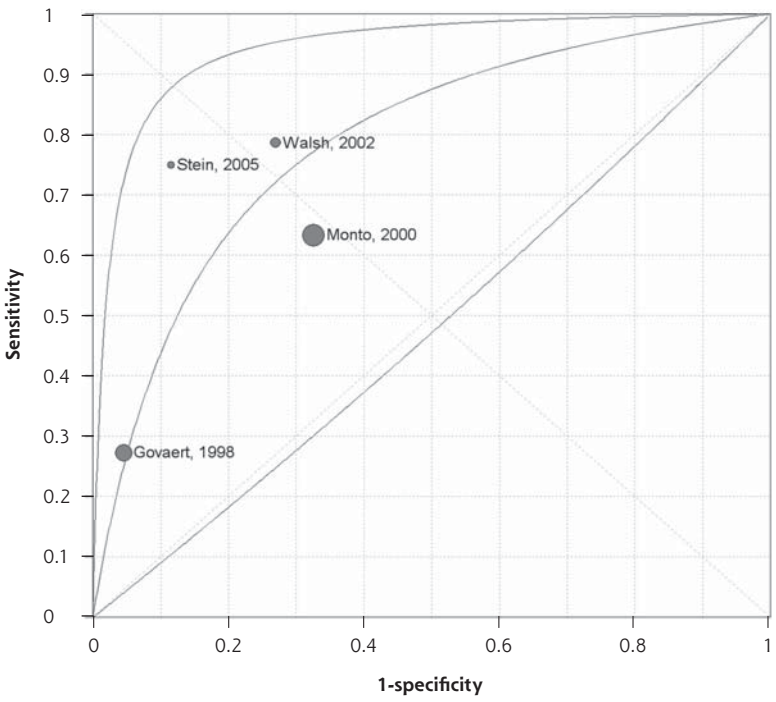
ROC = receiver operating characteristic.

Note: The curve graphs sensitivity vs 1−specificity for the 5 studies reporting data for this combination of symptoms. Lines are area under ROC curve and its 95% confidence interval. Size of circles is proportional to study size. The area under the ROC curve is 0.701, calculated as [SE(AUC) = 0.011], indicating moderate accuracy for distinguishing influenza from other infections.[8,11-13,22]

of the model in terms of sensitivity, specificity, predictive value, likelihood ratios, or area under the receiver operating characteristic curve.[13,21]

Several studies evaluated simple clinical heuristics such as the fever and cough rule, the fever, cough, and acute onset rule, and the cough, headache, and pharyngitis rule (Table 5). The fever and cough rule was evaluated in 5 studies, and the fever, cough, and acute onset rule was evaluated in 4 studies. The positive likelihood ratio for these rules ranged from 1.7 to 6.5, and the negative likelihood ratio ranged from 0.3 to 0.8. The ratio of positive to negative likelihood ratio (a measure of the ability to discriminate between diseased and nondiseased individuals) ranged from 3.6 to 21.7.[8,11-13,21,22] The area under the receiver operating characteristic curve was 0.70 for the fever and cough rule (Figure 1) and 0.79 for the fever, cough, and acute onset rule (Figure 2). The studies were too heterogeneous and the diagnostic threshold varied too extensively

Figure 2. Summary ROC curve for the fever, cough, and acute onset rule.



ROC = receiver operating characteristic.

Note: The curve graphs sensitivity vs 1−specificity for the 4 studies reporting data for this combination of symptoms. Lines are area under ROC curve and its 95% confidence interval. Size of circles is proportional to study size. The area under the ROC curve is 0.788, calculated as [SE(AUC) = 0.055], indicating moderate accuracy for distinguishing influenza from other infections.[8,11,13,21]

to estimate summary measures of accuracy for these simple heuristics, however. In general, surveillance studies and those with broader inclusion criteria had lower sensitivity and higher specificity,[8,13,22] whereas studies enrolling patients with influenza-like illness had higher sensitivity but lower specificity.[11,12,21] Only a single study evaluated a point score,[13] and that study had by far the lowest prevalence of influenza, because patients were enrolled and then reported any symptom occurring during flu season. No study prospectively proposed a point score or multivariate model, or used a split-sample or bootstrap procedure to evaluate such a model.

DISCUSSION

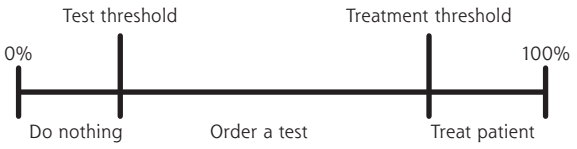
Although influenza is common and an important source of morbidity and mortality, studies of the diagnosis of this infection are largely small, use varied inclusion criteria and reference standards, and do not report their results in a way that would assist clinicians. In many cases, the inclusion criteria for the study (fever plus at least 1 other symptom) are also among the variables being evaluated for their accuracy, a potential source of bias. No study has prospectively evaluated a clinical score, CDR, or multivariate model.

Only simple clinical heuristics, such as the fever and cough rule, and the fever, cough, and acute onset rule, have been prospectively validated. Their sensitivity and specificity varied considerably, however, and it was not possible to calculate summary measures of accuracy for these rules. In part, this inability was due to varying selection criteria and use of the variables in question as some of the inclusion criteria for patients to be studied. It could be argued that it is inappropriate to combine data from a population-based study with a low prevalence of influenza[13] with data from 3 outpatient studies. It is interesting, however, that the positive and negative likelihood ratios of the studies of Govaert et al[13] and Stein et al[8] were almost identical despite their different populations. In addition, results of all 5 studies evaluating the fever and cough rule closely follow the same receiver operating characteristic curve, suggesting that they are measuring the same underlying construct. Because all 5 studies defined fever as a temperature of greater than 37.8°C or 38°C, it seems likely that the implicit definition of cough or how it was measured may have varied between studies.

The findings of our systematic review provide guidance for the design and conduct of future studies. For example, polymerase chain reaction should be used as the reference standard rather than culture because of its greater sensitivity for the detection of influenza.[10,28] Future studies should also ensure that they have an adequate sample size and include a broad range of patients with either respiratory tract infection or suspected influenza (without regard to whether they have fever, cough, or other symptoms).

Studies to date of influenza diagnosis have not prospectively validated multivariate models, an important next step in this area of research. In addition to multivariate models, researchers should validate point scores (based on multivariate models) that are simpler to use at the point of care. They should also explore alternate analytic methods, such as classification and regression trees and artificial neural networks. The latter have

Figure 3. Threshold model for diagnosis of influenza.



Note: Model indicates thresholds for ordering a test and for initiating therapy.

been widely used in biomedical research to develop classification and pattern recognition tools.[29] Originally designed to mimic the behavior of neurons, these networks are "trained" on one set of data and tested or validated on another. Mathematically, they are similar to a fully saturated multivariate model. Although prone to overfitting, the bootstrap procedure identifies the point at which error in the test set begins to rise because of overfitting.

When evaluating a patient with acute respiratory tract infection, a clinician must use information from the history and physical examination to decide among 3 courses of action: (1) rule out influenza and consider other diagnoses, (2) order a point-of-care test, or (3) treat empirically as influenza. These options are illustrated in the threshold diagram in Figure 3. Researchers have not adequately considered this clinical context. In addition to developing and validating a model, future researchers should identify the most useful test and treatment thresholds, either by surveying physicians to determine when they are comfortable ruling out or ruling in influenza, or by using quantitative analysis based on the harms and benefits of testing and treating.[30] Once established, these thresholds for testing and treating would be used to guide model development. These models for the diagnosis of influenza should identify at least 3 groups for whom different management strategies are indicated: low risk (do not pursue further testing or treatment for influenza), moderate risk (consider confirmatory testing), and high risk (give empiric therapy if the patient is seeking care within 48 hours of symptom onset).

What are clinicians to do? During influenza season, patients with fever and cough, especially if the onset was acute, have a high likelihood of influenza and do not require further testing unless complications such as pneumonia are suspected. For example, given a 33% pretest probability and using the primary care data from the study by Stein and colleagues,[8] the post-test probability of flu in such a patient is 76%. In the so-called shoulder season leading up to and following peak influenza season, a patient with fever, cough, and acute onset has a 42% likelihood of having influenza assuming a 10% pretest probability. Conversely, during shoulder season, a patient without this symptom triad has only a 3% likelihood of having influenza. These statistics can help clinicians and patients make informed decisions about care while CDRs undergo more rigorous evaluation.

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/cgi/content/full/9/1/69.

Key words: Influenza; clinical decision rule; systematic review; diagnosis; prediction model; decision model; clinical rule

Submitted April 6, 2010; submitted, revised, August 2, 2010; accepted September 1, 2010.

References

1. Jefferson T, Jones M, Doshi P, Del Mar C, Dooley L, Foxlee R. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. *Cochrane Database Syst Rev*. 2010;17(2):CD001265.
2. Ebell MH, White LL, Casault T. A systematic review of the history and physical examination to diagnose influenza. *J Am Board Fam Pract*. 2004;17(1):1-5.
3. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? *JAMA*. 2005;293(8):987-997.
4. Rothberg MB, Bellantonio S, Rose DN. Management of influenza in adults older than 65 years of age: cost-effetiveness of rapid testing and antiviral therapy. *Ann Intern Med*. 2003;139(5 Pt 1):321-329.
5. Rothberg MB, Fisher D, Kelly B, Rose DN. Management of influenza symptoms in healthy children: cost-effectiveness of rapid testing and antiviral therapy. *Arch Pediatr Adolesc Med*. 2005;159(11):1055-1062.
6. Ebell MH. Diagnosing and treating patients with suspected influenza. *Am Fam Physician*. 2005;72(9):1789-1792.
7. Senn N, Favrat B, D'Acremont V, Ruffieux C, Genton B. How critical is timing for the diagnosis of influenza in general practice? *Swiss Med Wkly*. 2005;135(41-42):614-617.
8. Stein J, Louie J, Flanders S, et al. Performance characteristics of clinical diagnosis, a clinical decision rule, and a rapid influenza test in the detection of influenza infection in a community sample of adults. *Ann Emerg Med*. 2005;46(5):412-419.
9. Hulson TD, Mold JW, Scheid D, et al. Diagnosing influenza: the value of clinical clues and laboratory tests. *J Fam Pract*. 2001;50(12):1051-1056.
10. van Elden LJ, van Essen GA, Boucher CA, et al. Clinical diagnosis of influenza virus infection: evaluation of diagnostic tools in general practice. *Br J Gen Pract*. 2001;51(469):630-634.
11. Monto AS, Gravenstein S, Elliott M, Colopy M, Schweinle J. Clinical signs and symptoms predicting influenza infection. *Arch Intern Med*. 2000;160(21):3243-3247.
12. Boivin G, Hardy I, Tellier G, Maziade J. Predicting influenza infections during epidemics with use of a clinical case definition. *Clin Infect Dis*. 2000;31(5):1166-1169.
13. Govaert TM, Dinant GJ, Aretz K, Knottnerus JA. The predictive value of influenza symptomatology in elderly people. *Fam Pract*. 1998;15(1):16-22.
14. Bender JM, Ampofo K, Gesteland P, et al. Development and validation of a risk score for predicting hospitalization in children with influenza virus infection. *Pediatr Emerg Care*. 2009;25(6):369-375.
15. Denoeud L, Turbelin C, Ansart S, Valleron AJ, Flahault A, Carrat F. Predicting pneumonia and influenza mortality from morbidity data. *PLoS One*. 2007;2(5):e464.

16. Talmor D, Jones AE, Rubinson L, Howell MD, Shapiro NI. Simple triage scoring system predicting death and the need for critical care resources for use during epidemics. *Crit Care Med*. 2007;35(5):1251-1256.

17. Challen K, Bright J, Bentley A, Walter D. 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*. 2007;7(1):33.

18. Hak EWF, Wei F, Nordin J, Mullooly J, Poblete S, Nichol KL. Development and validation of a clinical prediction rule for hospitalization due to pneumonia or influenza or death during influenza epidemics among community-dwelling elderly persons. *J Infect Dis*. 2004;189(3):450-458.

19. Zambon M, Hays J, Webster A, Newman R, Keene O. Diagnosis of influenza in the community: relationship of clinical diagnosis to confirmed virological, serologic, or molecular detection of influenza. *Arch Intern Med*. 2001;161(17):2116-2122.

20. Friedman MJ, Attia MW. Clinical predictors of influenza in children. *Arch Pediatr Adolesc Med*. 2004;158(4):391-394.

21. Walsh EE, Cox C, Falsey AR. Clinical features of influenza A virus infection in older hospitalized persons. *J Am Geriatr Soc*. 2002;50(9): 1498-1503.

22. van den Dool C, Hak E, Wallinga J, van Loon AM, Lammers JW, Bonten MJ. Symptoms of influenza virus infection in hospitalized patients. *Infect Control Hosp Epidemiol*. 2008;29(4):314-319.

23. Carrat F, Tachet A, Rouzioux C, Housset B, Valleron AJ. Evaluation of clinical case definitions of influenza: detailed investigation of patients during the 1995-1996 epidemic in France. *Clin Infect Dis*. 1999;28(2):283-290.

24. Ohmit SE, Monto AS. Symptomatic predictors of influenza virus positivity in children during the influenza season. *Clin Infect Dis*. 2006;43(5):564-568.

25. Ong A, Chen M, Lin L, et al. Improving the clinical diagnosis of influenza—a comparative analysis of new influenza A (H1N1) cases. *PLoS ONE*. 2009;4(12):e8453.

26. Babcock HM, Merz LR, Dubberke ER, Fraser VJ. Case-control study of clinical features of influenza in hospitalized patients. *Infect Control Hosp Epidemiol*. 2008;29(10):921-926.

27. Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess*. 2004;8(25):iii, 1-234.

28. Gharabaghi F, Tellier R, Cheung R, et al. Comparison of a commercial qualitative real-time RT-PCR kit with direct immunofluorescence assay (DFA) and cell culture for detection of influenza A and B in children. *J Clin Virol*. 2008;42(2):190-193.

29. Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. *Curr Opin Biotechnol*. 2002;13(1):72-76.

30. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302(20):1109-1117.