

This item is the archived peer-reviewed author-version of:

Validity of a clinical model to predict influenza in patients presenting with symptoms of lower respiratory tract infection in primary care

Reference:

van Vugt Saskia F., Broekhuizen Berna D.L., Zuithoff Nicolaas P.A., Coenen Samuel, Ieven Margaretha, Lammens Christine, Goossens Herman, et al.- Validity of a clinical model to predict influenza in patients presenting with symptoms of lower respiratory tract infection in primary care

Family practice - ISSN 0263-2136 - 32:4(2015), p. 408-414

Full text (Publishers DOI): <http://dx.doi.org/doi:10.1093/fampra/cmz039>

To cite this reference: <http://hdl.handle.net/10067/1266660151162165141>

Health Service Research

Validity of a clinical model to predict influenza in patients presenting with symptoms of lower respiratory tract infection in primary care

Saskia F van Vugt^{a,*}, Berna DL Broekhuizen^a, Nicolaas PA Zuihthoff^a, Gerrit A van Essen^a, Mark H Ebell^b, Samuel Coenen^{c,d}, Margareta Ieven^d, Christine Lammens^d, Herman Goossens^d, Chris C Butler^e, Kerenza Hood^f, Paul Little^g and Theo JM Verheij^a, on Behalf of the GRACE Consortium

^aUniversity Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, Utrecht, The Netherlands,

^bDepartment of Epidemiology and Biostatistics, University of Georgia, College of Public Health, Athens, GA, USA,

^cCentre for General Practice, University of Antwerp and ^dLaboratory of Medical Microbiology, Vaccine & Infectious Disease Institute (VAXINFECTIO), Antwerp, Belgium, ^eInstitute of Primary Care and Public Health, Cardiff University, School of Medicine, and ^fInstitute of Translation, Innovation, Methodology and Engagement, Cardiff, UK, and

^gUniversity of Southampton, FMedSci FRCGP Primary Care and Population Sciences Unit, Faculty of Medicine, Southampton, UK.

*Correspondence to Saskia F van Vugt, Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Stratenum 6.131, PO Box 85500, 3508 GA Utrecht, The Netherlands; E-mail s.f.vanvugt@umcutrecht.nl

Abstract

Background. Valid clinical predictors of influenza in patients presenting with lower respiratory tract infection (LRTI) symptoms would provide adequate patient information and reassurance.

Aim. Assessing the validity of an existing diagnostic model (Flu Score) to detect influenza in LRTI patients.

Design and Setting. A European diagnostic study recruited 1801 adult primary care patients with LRTI-like symptoms existing ≤ 7 days between October and April 2007–2010.

Method. History and physical examination findings were recorded and nasopharyngeal swabs taken. Polymerase chain reaction (PCR) for influenza A/B was performed as reference test. Diagnostic accuracy of the Flu Score (1× onset <48 hours + 2× myalgia + 1× chills or sweats + 2× fever and cough) was expressed as area under the curve (AUC), calibration slopes and likelihood ratios (LRs).

Results. A total of 273 patients (15%) had influenza on PCR. The AUC of the Flu Score during winter months was 0.66 [95% CI (95% confidence interval) 0.63–0.70]. During peak influenza season, both influenza prevalence (24%) and AUC were higher [0.71 (95% CI 0.66–0.76)], but calibration remained poor. The Flu Score assigned 64% of the patients as ‘low-risk’ (10% had influenza, LR – 0.6). About 12% were classified as ‘high risk’ of whom 32% had influenza (LR + 2.7). During peak influenza season, 60% and 14% of patients were classified as low and high risk, respectively, with influenza prevalences being 14% (LR – 0.5) and 50% (LR + 3.2).

Conclusion. The Flu-Score attributes a small subgroup of patients with a high influenza risk (prevalence 32%). However, clinical usefulness is limited because this group is small and the association between predicted and observed risks is poor. Considerable diagnostic imprecision remains when it comes to differentiating those with influenza on clinical grounds from the many

other causes of LRTI in primary care. New point of care tests are required that accurately, rapidly and cost effectively detect influenza in patients with respiratory tract symptoms in primary care.

Keywords: Cough, diagnostic accuracy, Europe, influenza, lower respiratory tract infection, primary health care.

Introduction

Lower respiratory tract infection (LRTI) is among the most common reasons for consultation in primary care, particularly during winter months when 5% of all visits are related to respiratory infections (1). If acute cough is accompanied by acute onset of fever, headache and muscle ache, this syndrome is called influenza-like illness (ILI) because this illness is assumed to be caused by an infection with an influenza virus. However, of all patients suffering from ILI, only 20–30% has true influenza, with a peak percentage of influenza virus positivity of 54% (1,2). Differentiating those infected with influenza from the large number of LRTI and ILI patients is important to provide adequate patient information and reassure patients about the expected self-limiting course, possible complications and appropriate preventive measures such as social distancing and hand washing, in case of early presentation (3,4).

A review assessing the performance of multivariable models and clinical decision rules to distinguish influenza from other causes of ILI found too heterogeneous studies to estimate summary measures of accuracy (5). Therefore, a recent study combined data from two similarly designed studies to create a larger dataset to develop a new diagnostic clinical model for influenza (Flu-Score: 1× onset <48 hours + 2× myalgia + 1× chills or sweats + 2× fever and cough) (6). After internal validation and comparison to previously proposed models, the Flu-Score was found to be the most accurate [e.g. highest reported likelihood ratios (LRs) and most useful (e.g. simple and easy to remember for point of care use)] diagnostic model for influenza (6). However, the prevalence and clinical presentation of infection with different Influenza subtypes varies between seasons. The diagnostic value of the Flu Score over several influenza seasons, when different influenza strains predominate, and during periods outside influenza outbreaks, is unknown. Moreover, not only in ILI but all LRTI patients a valid diagnosis is warranted in daily practice. Therefore, external validation of this score in an extended context of general practice is required before it is promoted for use in everyday care. We aimed to explore the validity of the Flu Score for identifying influenza infection among patients seeking health care for LRTI in a large European primary care study over three consecutive winters.

Methods

Design

This was a cross-sectional observational study which was part of the GRACE (Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe; www.grace-lrti.org) study.

Setting

The GRACE study collected data in 16 primary care research networks (PCNs) in 12 European countries. Participating general practitioners (GPs) recruited patients presenting with LRTI from October 2007 to April 2010.

Patients

Eligible patients were at least 18 years old, with any clinical presentation considered to be caused by LRTI by the responsible GP.

Symptom onset had to be ≤ 7 days, because the influenza detection rate is shown to be minimal after 1 week since symptom onset (7,8). All patients were consulting for the first time for this illness episode. Further inclusion criteria are similar to previous publications (9). Medical ethics committees of the participating centres approved the study.

Reference test (diagnostic outcome)

Nasopharyngeal swabs, taken by trained research staff within 24 hours after recruitment and before any antimicrobial treatment had started, were immediately placed on universal transport medium, stored on dry ice at -70°C and transported to the central microbiological laboratory of the University of Antwerp. The swabs were analyzed for Influenza A and B by reverse transcriptase polymerase chain reaction (RT-PCR) (2,4) after EssayMag extraction (10). Microbiologists who determined the results were blind to the clinical information. Influenza was considered present if the PCR was positive for influenza A or B.

Measurements and follow up

GPs recorded patients' symptoms, signs, comorbidities (e.g. diabetes, respiratory and cardiovascular disease) and influenza vaccination status at recruitment in a standardized Case Report Form (CRF).

Data analysis

Less than 1% of history items, <3% of physical examination items and 6% of nasopharyngeal swabs were missing (Table 1). Missing data rarely occur completely at random and therefore we used multiple imputation techniques to account for missing values, including on the diagnostic outcome influenza (11,12).

First, we calculated the prevalence of influenza and the clinical characteristics for the total study population, who had been recruited during winter months. In these patients, we explored the diagnostic value of the Flu Score to identify influenza. Because the baseline probability of influenza in the community is lower outside influenza outbreaks, clinical usefulness of the Flu Score is likely to vary between periods of low influenza prevalence compared to influenza peak season. We therefore repeated our analyses on the subgroup of patients who had been recruited during periods of documented influenza peak season. Data from the European Influenza Surveillance Network (EISN) (13–15) were used to determine weeks of influenza peak prevalence during our study for each PCN in Europe separately.

To assess the validity, first discrimination and calibration of the Flu Score were calculated (16). Discrimination is the ability to distinguish patients with influenza from patients without, and the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals (CIs) was calculated to quantify the discriminative ability of the Flu Score in our sample. Calibration is a measure for correspondence between predicted probabilities and the observed prevalence. This was assessed with both a calibration slope and the Hosmer and Lemeshow (HL) statistic (16).

Next, the Flu Score was used to calculate an individual influenza score for all patients based on the presence or absence of the signs

Table 1. Clinical characteristics of 1801 patients presenting with LRTI in primary care, during winter months (October–April) and during peak influenza season 2007–10 (validation cohort), compared to derivation cohort (6)

Diagnostic variable	Derivation cohort (6)		Validation cohort			Peak influenza season	
	Influenza season		Winter months (October–April)			Influenza A or B present	
	Total N = 459 (%)	Missing N (%)	Total N = 1801 (%) ^a	Influenza A or B present N = 273 (15%) ^b	Total N = 505 (%) ^a	Influenza A or B present N = 119 (24%) ^b	
Risk status							
High-risk medical condition ^b	n.a.	4 (0.2)	431 (24)	52 (19)	113 (22)	23 (19)	
Influenza vaccination last year	n.a.	1 (0.1)	378 (21)	30 (11)	90 (18)	12 (10)	
History (day 1)							
Age (mean, SD)	37 (14)	0 (0.0)	48 (16)	45 (15)	47 (16)	44 (15)	
Male gender	204 (44)	0 (0.0)	732 (41)	103 (38)	209 (41)	48 (40)	
Current smoker	37 (14) ^c	1 (0.1)	969 (54)	132 (48)	278 (55)	57 (48)	
Days ill prior to consultation (mean, SD)	6 (5) ^c	31 (1.7)	5 (2)	4 (2)	6 (2)	4 (2)	
Onset of symptoms ≤2 days	151 (33)	31 (1.7)	299 (17)	69 (25)	93 (18)	37 (31)	
Cough present	421 (92)	0 (0)	1800 (100)	272 (100)	504 (100)	118 (99)	
Phlegm present	315 (88) ^c	3 (0.2)	1417 (79)	194 (71)	387 (77)	82 (69)	
Runny nose present	348 (76)	2 (0.1)	1327 (74)	207 (76)	382 (76)	94 (79)	
Fever, chills or sweating present	324 (71)	3 (0.2)	749 (42)	181 (66)	218 (43)	83 (70)	
Chest pain present	92 (36) ^c	3 (0.2)	901 (50)	156 (57)	270 (54)	71 (60)	
Myalgia	335 (73)	2 (0.1)	1027 (57)	193 (71)	310 (61)	94 (79)	
Headache	359 (78)	1 (0.1)	1121 (62)	204 (75)	311 (62)	95 (80)	
General feeling unwell/fatigue	381 (83)	2 (0.1)	1479 (82)	244 (89)	415 (82)	104 (87)	
Interference with daily activities	n.a.	1 (0.1)	1243 (69)	224 (82)	356 (71)	95 (80)	
Physical examination (day 1)							
Any abnormal auscultation sound	n.a.	12 (0.7)	712 (40)	97 (36)	181 (36)	38 (32)	
Tachycardia (>100/minute)	n.a.	23 (1.3)	80 (4)	12 (4)	24 (5)	4 (3)	
Tachypnoea (>24/minute)	n.a.	45 (2.5)	32 (2)	1 (0)	7 (1)	0 (0)	
Nasopharyngeal swabs (day 1)							
Influenza A or B	157 (34)	116 (6.4)	273 (15)	273 (100)	119 (24)	119 (100)	

Values are numbers (percentages) unless mentioned otherwise. n.a., not available from the original study; SD, standard deviation.

^aCalculated as the mean of 10 imputations.

^bHeart or lung conditions, diabetes mellitus, chronic renal insufficiency, chronic Staphylococcal infections or immune-related diseases (35).

^cBased on one of the two derivation cohorts only (22).

and symptoms from the score (1× onset <48 hours + 2× myalgia + 1× chills or sweats + 2× fever and cough) to stratify them into a low-, intermediate- or high-risk group using the cut points proposed for the score (Table 1). Cutoff thresholds of 10% and 50% were suggested to be reasonable for, respectively, ruling out and ruling in influenza, based on a previous report on the opinion of 5 experienced primary care physicians, an informal email poll of 20 academic generalists and a previous decision threshold analysis (6,17). Patients with an estimated probability between 10% and 50% were defined as intermediate-risk patients, in whom influenza diagnosis remains inconclusive.

Finally, we determined positive and negative LR for the low risk group (intermediate and high versus low) and the high risk group (high versus low and intermediate) and compared these to the LR derived in the original study (6).

All analyses were repeated for a second Flu Score presented in the original study of Ebell *et al.* (6) that was more complex (Flu Score 2): 7× fever + 5× cough + 4× onset <48 hours + 8× myalgia.

Because treatment with antivirals should be taken within 48 hours of the onset of symptoms to have any benefit (18), we repeated all analyses among patients with onset of symptoms ≤2 days ($n = 299$).

As in most multicentre studies, clustering of patient data occurs to some extent within sites (in our case, the 16 PCNs), which could affect associations between diagnostic variables and outcomes (19). We accounted for such possible nonrandom differences within PCNs (clusters) using multilevel analyses (19).

Data were analysed using SPSS (Version 20.0 for Windows), R (Version 2.11.1) and SAS (Version 9.2) (20).

Results

Patient characteristics

A total of 1801 patients consulting their GP with LRTI symptoms existing 7 days or less were recruited during three winters from 2007 to 2010. PCR results were not available for 116 patients (6%), but these patients were similar on key characteristics to patients for whom PCR results were available (Supplementary Table 1). The mean age of all participants was 48 years (SD 16) and 41% were male. About 299 patients consulted their GP within 48 hours of symptom onset. About 75% of the participants with influenza presented after more than 2 days. Of the 378 patients (21%) who had

received an influenza vaccination in the previous year, 30 (8%) had influenza. Patient characteristics are displayed in Table 1.

Prevalence of influenza

Of the 1801 participants, 273 had influenza (15%). About 166 (61%) of these had influenza A, 86 (32%) influenza B and 8% remained undetermined. Influenza prevalence was 24% during the peak of the influenza season. In those 65 years and older (18% of the total study population), the overall proportion of influenza A/B infections was 10%. The prevalence of influenza A and B varied by winter (Fig. 1) and PCN, ranging from 2% in Spain in 2010 to 37% in Poland in 2008. In the 299 participants recruited within 2 days of illness duration, 23% had influenza. About 75% of patients with influenza presented later than 2 days of symptom onset.

Validity of the Flu Score

During the entire winter period (October–April), 64% of patients were classified as low risk for influenza according to the Flu Score and 12% as high risk. The proportion with influenza in the low-risk group was 10% (negative LR 0.6) compared to 32% in the high-risk group (positive LR 2.7). The remaining 25% of the patients were classified as intermediate risk, with an observed influenza presence of 21% (Table 2). The AUC of the Flu Score for detecting influenza A/B was 0.66 (95% CI 0.63–0.70). Calibration was poor (P value for HL test <0.001, and a calibration slope of 0.71). Recalibration to correct for difference in baseline prevalence between the original and the present study population did not improve calibration (P value HL test <0.001). LR for each risk group are shown in Table 2.

During peak influenza season discrimination was better (AUC 0.71, 95% CI 0.66–0.76, $P < 0.05$), but calibration remained poor (P value HL test <0.001). Almost 60% of patients were classified as low-risk for influenza, with a post-test probability of 14% for influenza (negative LR 0.5). The 14% classified as high risk had a post-test probability of 50% (positive LR 3.2, Table 2).

In patients presenting with an onset of symptoms ≤2 days, the AUC was 0.70 (95% CI 0.63–0.77, whole winter period) and 0.74 (95% CI 0.64–0.85, during influenza peak season, not shown). The risk classification is shown in Table 2.

The second diagnostic model developed by Ebell (Flu Score 2: 7× fever + 5× cough + 4× onset <48 hours + 8× myalgia), generated a

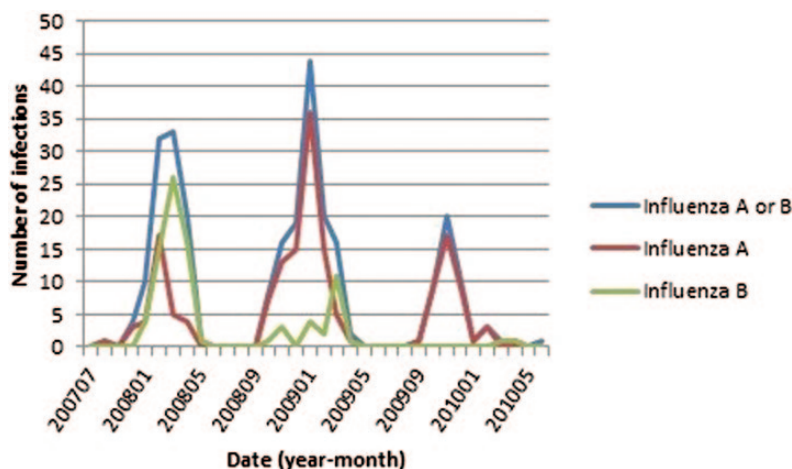


Figure 1. Influenza A/B virus detected by PCR in nasopharyngeal swabs in 1801 primary care patients presenting with LRTI over time (number of infections per week)

Table 2. Performance of the Flu Score^a for the diagnosis of influenza A/B in different settings and subgroups

	Validation study					AUC (95% CI)
	Points	Influenza/total (%)	LR – (1/LR–) ^b	LR+ ^b	% in group	
Derivation study (6) (n = 459)						Unknown
Low risk	0–2	12/149 (8)	0.2 (5.6)	1.7	33	
Intermediate risk	3	39/129 (30)	0.4 (2.4)	2.7	28	
High risk	4–6	106/181 (59)			39	
Winter months (n = 1801)						0.66 (0.63–0.70)
Low risk	0–2	111/1146 (10)	0.6 (1.7)	1.8	64	
Intermediate risk	3	95/447 (21)	0.8 (1.2)	2.7	25	
High risk	4–6	67/208 (32)			12	
Influenza peak season ^c (n = 505)						0.71 (0.66–0.76)
Low risk	0–2	41/301 (14)	0.5 (2.0)	2.0	60	
Intermediate risk	3	43/134 (32)	0.8 (1.3)	3.2	27	
High risk	4–6	35/70 (50)			14	
Onset of symptoms ≤2 days (n = 299)						0.70 (0.63–0.77)
Low risk	0–2	11/96 (12)	0.4 (2.3)	1.3	32	
Intermediate risk	3	15/82 (18)	0.6 (1.8)	1.8	27	
High risk	4–6	43/121 (36)			41	

High LRs [threshold of >3 to indicate clear diagnostic value (36)] in bold.

^aFlu Score = 1 × onset <48 hours + 2 × myalgia + 1 × chills/sweats + 2 × fever × cough. Note: Original derivation model: Risk of influenza A or B = 1/(1 + exp – (–2.808 + 0.721 × onset <48 hours + 1.260 × myalgia + 0.401 × chills/sweats + 1.361 × fever × cough)) (6). The LRs of the low-risk group are calculated from a two-by-two table with the sums of intermediate and high risk groups labelled as test positives and the values for the low risk group as test negatives. For calculation of the LRs of the high-risk group, the two-by-two table with the sums for low- and intermediate-risk groups labelled as test negatives and the values for high risk as test positives were used.

^cInfluenza season defined per PCN, according to data from the European Influenza Surveillance Network (EISN) (13–15).

similar risk classification, but more (42%) patients were classified as intermediate risk (Supplementary Table 2). The AUC was 0.63 (95% CI 0.59–0.66).

Discussion

Summary

In 1801 patients presenting to primary care with symptoms of LRTI during three consecutive winter periods (October–April), 15% had influenza on PCR testing of nasopharyngeal swab samples. The influenza prevalence was higher during peak influenza season (24%). A recently published diagnostic model, the Flu Score (onset <48 hours, cough, myalgia, fever, chills and sweats), assigned 14% of patients with a high risk of having influenza (LR + 3.2) during peak influenza season. Outside the peak influenza season, diagnostic accuracy of the score (e.g. ROC area, calibration slopes and LRs) was even smaller.

Strengths and limitations

To our knowledge, this is the first formal external validation study of the Flu Score. A strength of this study is the broad setting (16 PCNs in 12 countries and inclusion of all adults presenting with LRTI in primary care during three consecutive winters). External validation in this extended context of primary care is useful as this setting mimics daily practice.

The timing of our study could nevertheless affect the generalizability of our results. First, the virulence of the circulating influenza viruses determines both prevalence as well as morbidity of the infection. The 2009 pandemic influenza A (H1N1) virus, e.g. has been shown generally to cause mild disease (21). However, we recruited patients over three influenza seasons (2007–10), so some of the seasonal variation in influenza morbidity patterns that is related to different flu strains are accounted for in our analyses.

Subgroups were too small to analyze patients for each season separately.

Second, we used European surveillance data on increased influenza activity as a dichotomized variable in our data to define times of documented peak influenza season. Influenza shows geographic variation across Europe. Although surveillance data for each country were used, these numbers remain only a rough approximation for the local prevalence. Therefore, some misclassification of the defined peak influenza season might have occurred. However, by using LRs as an outcome measure we corrected for differences in influenza prevalence that might have occurred due to geographic variation.

One of the criteria for study inclusion was the presence of LRTI symptoms less than 7 days, because studies show positive PCR up to 7 days after onset of symptoms (7,8). In this study, we added an additional analysis in the subgroup of patients (17%) with an onset of symptoms ≤2 days because anti-influenza therapy is potentially useful only in these patients. Disadvantage of this subgroup analysis is the preselection of the population using one of the variables in the Flu Score, which is likely to have decreased the diagnostic value of the Flu Score. Moreover clinical utility of influenza diagnosis may extend beyond the therapeutic window of 48 hours for anti-influenza therapy.

Cough was a diagnostic item for the presence of influenza in the Flu Score that was present in almost all our study participants (1800 of 1801). Also the other symptoms of the Flu Score were very common in this population. This might have affected the performance of the Flu Score. However, the original study also found high proportions of the different items of the Flu Score (e.g. over 90% of patients had a cough), which underlines the importance of these symptoms in patients suspected of having influenza.

Comparison with existing literature

In this validation study, influenza prevalence both during the whole winter period (15%), as during influenza peak season (24%), was

lower when compared with the derivation study (34%). A possible explanation is the different clinical spectrum of included patients. Although the duration of symptoms prior to study inclusion in our study was similar (5 days) when compared with the derivation study (6 days), the derivation study included patients who fulfilled pre-specified clinical criteria for influenza (22) or RTI (23), while our study included all patients with symptoms and signs suggestive of LRTI. This could have meant that more patients were included in the derivation study who were more likely to have influenza at the outset. However, in case of any prevalence difference, this would have affected positive and negative values, but not the LRs we used in this study. Besides, other studies of LRTI patients found influenza prevalence rates within a similar range (e.g. 5%–30%) as in our study (24,25).

Validation of the Flu Score showed a different pattern of risk classification compared to the original study, with smaller differences in the observed influenza prevalences between the three risk groups. Especially the influenza prevalence in the high-risk group was lower compared to the original study (Table 2). External validation almost always shows a decreased performance of the model under study because of overfitting (when the ratio of tested variables to the number of patients experiencing events is small, or reproducibility is low) and underfitting (when important independent predictors of outcome are omitted from the system, due to historical period, geographic location, methodological approach or disease spectrum) (26). It is likely that the difference in performance of the model in two populations was also caused by the clinical differences between the two study populations, as mentioned earlier.

Previous studies all found diagnostic value for the clinical variables cough, fever, myalgia and acute onset for detecting influenza (27–32). However, performance of previously published diagnostic models that included these variables varied widely, with LRs ranging from 1.7 to 6.5 (5). One of the possible explanation for these differences could be the variation in the interpretation of cough, chills, sweats and fever both by doctors and patients. This confirms that reported diagnostic models with point or weighted scores for each item present, are prone to overfitting and poor performance in new sets of patients (33). Moreover, the point and weighted scores often result in complex arithmetic models that are not easy to recall and their use is known to be limited in clinical practice (34). We therefore did not develop a new diagnostic model for influenza, but rather aimed to enhance recommendations for clinical practice after assessing the performance of the previously reported most robust signs and symptom combination.

Interpretation of the results

The Flu Score can only be useful for daily practice if it classifies patients at influenza risks that are substantially lower or higher than the pre-test probability. The post-test probability for excluding influenza (10%) remained close to the 15% pre-test probability, rendering limited clinical value. During influenza peak season, the post-test probability for patients assigned to the high-risk group reached a 5-fold increase to 50% (LR + 3.2). According to a previous decision threshold analysis (17), adequate patient information, reassurance and, if presentation is early, appropriate preventive measures could be considered in this high-risk group without further testing. However, the low number of patients in this category limits the clinical usefulness of this finding. Also the poor calibration shows the limited value for use in daily practice, because the models' predicted probabilities do not agree with the observed influenza prevalence.

Conclusions

A small subgroup of patients with a high-influenza risk can be identified when using the Flu Score. However, clinical usefulness is limited because the association between predicted and observed risks is poor. Considerable diagnostic imprecision remains when it comes to differentiating those with influenza on clinical grounds from the many other causes of LRTI in primary care. New point-of-care tests are required that accurately, rapidly and cost effectively detect influenza in patients with respiratory tract symptoms in primary care.

Supplementary material

Supplementary material is available at *Family Practice* online.

Acknowledgements

We would like to thank the entire GRACE team for their diligence, expertise and enthusiasm. Finally, we are indebted to all of the patients who consented to be part of GRACE, without whom this study would not have been possible. The following are members of the GRACE consortium (www.grace-lrti.org): Saskia van Vugt (the Netherlands), Lidewij Broekhuizen (the Netherlands), Peter Zuithoff (the Netherlands), Ted van Essen (the Netherlands), Mark Ebell (USA), Chris Butler (UK), KerENZA Hood (UK), Samuel Coenen (Belgium), Herman Goossens (Belgium), Margareta Ieven (Belgium), Christine Lammens (Belgium), Jordi Almirall (Spain), Francesco Blasi (Italy), Slawomir Chlabicz (Poland), Mel Davies (UK), Maciek Godycki-Cwirko (Poland), Helena Hupkova† (Slovakia), Janko Kersnik† (Slovenia), Artur Mierzecki (Poland), Sigvard Mölstad (Sweden), Michael Moore (UK), Tom Schaberg (Germany), An De Sutter (Belgium), Antoni Torres (Spain), Pia Touboul (France), Paul Little (UK), and Theo Verheij (the Netherlands).

Declaration

Funding: This study was part of the GRACE project (www.grace-lrti.org), funded by 6th Framework Program of the European Commission (Reference: LSHM-CT-2005-518226). In Flanders (Belgium) this work was supported by the Research Foundation, Flanders (G-0274-08N).

Ethical approval: Medical ethics committees of the participating centres.

Conflict of interest: none.

References

1. Fowlkes A, Dasgupta S, Chao E *et al*. Estimating influenza incidence and rates of influenza-like illness in the outpatient setting. *Influenza Other Respir Viruses* 2013; 7: 694–700.
2. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. Does this patient have influenza? *JAMA* 2005; 293: 987–97.
3. Michiels B, Van Puyenbroeck K, Verhoeven V, Vermeire E, Coenen S. The value of neuraminidase inhibitors for the prevention and treatment of seasonal influenza: a systematic review of systematic reviews. *PLoS One* 2013; 8: e60348.
4. Talbot HK, Falsey AR. The diagnosis of viral respiratory disease in older adults. *Clin Infect Dis* 2010; 50: 747–51.
5. Ebell MH, Afonso A. A systematic review of clinical decision rules for the diagnosis of influenza. *Ann Fam Med* 2011; 9: 69–77.
6. Ebell MH, Afonso AM, Gonzales R *et al*. Development and validation of a clinical decision rule for the diagnosis of influenza. *J Am Board Fam Med* 2012; 25: 55–62.
7. Lau LL, Cowling BJ, Fang VJ *et al*. Viral shedding and clinical illness in naturally acquired influenza virus infections. *J Infect Dis* 2010; 201: 1509–16.
8. Carrat F, Vergu E, Ferguson NM *et al*. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. *Am J Epidemiol* 2008; 167: 775–85.

9. van Vugt SF, Broekhuizen BD, Lammens C *et al.* Use of serum C reactive protein and procalcitonin concentrations in addition to symptoms and signs to predict pneumonia in patients presenting to primary care with acute cough: diagnostic study. *BMJ* 2013; 346.
10. Loens K, van Loon AM, Coenjaerts F *et al.* Performance of different mono- and multiplex nucleic acid amplification tests on a multipathogen external quality assessment panel. *J Clin Microbiol* 2012; 50: 977–87.
11. Mackinnon A. The use and reporting of multiple imputation in medical research - a review. *J Intern Med* 2010; 268: 586–93.
12. Moons KG, Donders RA, Stijnen T, Harrell FE Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006; 59: 1092–101.
13. European Centre for Disease Prevention and Control. Annual Epidemiological Report on Communicable Diseases in Europe 2009. Stockholm, Sweden: ECDC, 2009.
14. European Centre for Disease Prevention and Control. Annual Epidemiological Report on Communicable Diseases in Europe 2010. Stockholm, Sweden: ECDC, 2010.
15. European Centre for Disease Prevention and Control. Annual Epidemiological Report 2011: Reporting on 2009 Surveillance Data and 2010 Epidemic Intelligence Data. Stockholm, Sweden: ECDC, 2011.
16. Steyerberg EW. *Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating.* New York: Springer Science Business, 2009.
17. Sintchenko V, Gilbert GL, Coiera E, Dwyer D. Treat or test first? Decision analysis of empirical antiviral treatment of influenza virus infection versus treatment based on rapid test results. *J Clin Virol* 2002; 25: 15–21.
18. Jefferson T, Jones M, Doshi P, Del Mar C. Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis. *BMJ* 2009; 339.
19. Urbach DR, Austin PC. Conventional models overestimate the statistical significance of volume-outcome associations, compared with multilevel models. *J Clin Epidemiol* 2005; 58: 391–400.
20. R Development Core Team. *A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2008.
21. Nicoll A, McKee M. Moderate pandemic, not many dead—learning the right lessons in Europe from the 2009 pandemic. *Eur J Public Health* 2010; 20: 486–8.
22. 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: 412–9.
23. 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: 614–7.
24. Macfarlane JT, Colville A, Guion A, Macfarlane RM, Rose DH. Prospective study of aetiology and outcome of adult lower-respiratory-tract infections in the community. *Lancet* 1993; 341: 511–4.
25. Woodhead M, Blasi F, Ewig S *et al.* Guidelines for the management of adult lower respiratory tract infections—full version. *Clin Microbiol Infect* 2011; 17 (suppl 6): E1–59.
26. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999; 130: 515–24.
27. 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: 630–4.
28. Govaert TM, Dinant GJ, Aretz K, Knottnerus JA. The predictive value of influenza symptomatology in elderly people. *Fam Pract* 1998; 15: 16–22.
29. 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: 283–90.
30. Monto AS, Gravenstein S, Elliott M, Colopy M, Schweinle J. Clinical signs and symptoms predicting influenza infection. *Arch Intern Med* 2000; 160: 3243–7.
31. 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: 1166–9.
32. Michiels B, Thomas I, Van Royen P, Coenen S. Clinical prediction rules combining signs, symptoms and epidemiological context to distinguish influenza from influenza-like illnesses in primary care: a cross sectional study. *BMC Fam Pract* 2011; 12: 4.
33. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000; 19: 453–73.
34. Redelmeier DA, Lustig AJ. Prognostic indices in clinical practice. *JAMA* 2001; 285: 3024–5.
35. Hak E, Buskens E, van Essen GA *et al.* Clinical effectiveness of influenza vaccination in persons younger than 65 years with high-risk medical conditions: the PRISMA study. *Arch Intern Med* 2005; 165: 274–80.
36. Sackett D, Richardson WS, Rosenberg W, Haynes RB. *Evidence Based Medicine: How to Practice and Teach EBM.* New York, NY: Churchill Livingstone, 1997.