BMJ Open

Derivation and validation of clinical prediction rules for diagnosis of infectious mononucleosis: a prospective cohort study

Muireann de Paor [1], Fiona Boland [1], Xinyan Cai,[2] Susan Smith,[3] Mark H Ebell [2], Eoin Mac Donncha,[4] Tom Fahey [1]

¹HRB Centre For Primary Care Research, Division of Population Health Sciences (PHS), Royal College of Surgeons Ireland, Dublin, Ireland
²Epidemiology and Biostatistics, University of Georgia, Athens, Georgia, USA
³Community Health and General Practice, Trinity College Dublin, Dublin, Ireland
⁴Student Health Unit, National University of Ireland Galway, Galway, Ireland

Correspondence to
Dr Muireann de Paor;
muireann.depaor@hse.ie

ABSTRACT

Objectives Infectious mononucleosis (IM) is a clinical syndrome that is characterised by lymphadenopathy, fever and sore throat. Although generally not considered a serious illness, IM can lead to significant loss of time from school or work due to profound fatigue, or the development of chronic illness. This study aimed to derive and externally validate clinical prediction rules (CPRs) for IM caused by Epstein-Barr virus (EBV).

Design Prospective cohort study.

Setting and participants 328 participants were recruited prospectively for the derivation cohort, from seven university-affiliated student health centres in Ireland. Participants were young adults (17–39 years old, mean age 20.6 years) with sore throat and one other additional symptom suggestive of IM. The validation cohort was a retrospective cohort of 1498 participants from a student health centre at the University of Georgia, USA.

Main outcome measures Regression analyses were used to develop four CPR models, internally validated in the derivation cohort. External validation was carried out in the geographically separate validation cohort.

Results In the derivation cohort, there were 328 participants, of whom 42 (12.8%) had a positive EBV serology test result. Of 1498 participants in the validation cohort, 243 (16.2%) had positive heterophile antibody tests for IM. Four alternative CPR models were developed and compared. There was moderate discrimination and good calibration for all models. The sparsest CPR included presence of enlarged/tender posterior cervical lymph nodes and presence of exudate on the pharynx. This model had moderate discrimination (area under the receiver operating characteristic curve (AUC): 0.70; 95% CI: 0.62–0.79) and good calibration. On external validation, this model demonstrated reasonable discrimination (AUC: 0.69; 95% CI: 0.67–0.72) and good calibration.

Conclusions The alternative CPRs proposed can provide quantitative probability estimates of IM. Used in conjunction with serological testing for atypical lymphocytosis and immunoglobulin testing for viral capsid antigen, CPRs can enhance diagnostic decision-making for IM in community settings.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A prospective cohort study with external validation in a geographically separate population.

⇒ Proposes alternative clinical prediction rules for a relatively common condition in primary care, infectious mononucleosis.

⇒ Only examined infectious mononucleosis where the underlying cause was Epstein-Barr virus; 5%–10% of cases may be due to other organisms such as cytomegalovirus, acute HIV infection and toxoplasmosis.

⇒ Derivation and validation cohorts differ in terms of the reference standard diagnostic test used; lower specificity of the heterophile antibody test in the validation cohort may have resulted in more false-negative results.

INTRODUCTION

The diagnosis of infectious mononucleosis (IM) is made on clinical grounds, based on characteristic symptoms such as fever, sore throat, lymphadenopathy and malaise. Less common features include a maculopapular rash, jaundice, splenomegaly and more rarely hepatomegaly.[1] Accurate diagnosis of IM in family practice is important. Correct diagnosis enables family physicians to provide patients with advice on avoidance of contact sports activity (to prevent the possibility of splenic rupture) and will help patients avoid unnecessary antibiotic treatment.[2] It also supports the provision of prognostic information about IM, with some evidence suggesting that approximately 12% of patients affected will experience persistent fatigue.[3]

Clinical prediction rules (CPRs) are tools that combine and weight the contribution of symptoms, clinical signs and available diagnostic tests, and in doing so stratify patients according to the probability of having a target disorder.[4] There are three main stages in the development of a CPR: derivation, validation

BMJ

(internal and external) and impact analysis.[5 6] External validation is needed for two primary reasons: first, associations between explanatory variables and outcomes may occur because of chance, and second, the explanatory variables for the CPR may be particular to the derivation population setting and may not be widely generalisable to a new setting.[5 6]

There are no widely used CPRs for IM. A 2021 systematic review of diagnostic accuracy of clinical signs, symptoms and haematological parameters for IM concluded that the presence of posterior cervical lymphadenopathy, inguinal or axillary lymphadenopathy, palatine petechiae, splenomegaly and atypical lymphocytosis with or without high lymphocyte count increased the likelihood of a diagnosis of IM.[7] The systematic review concluded that well-designed prospective studies are required to further examine the importance of clinical and laboratory-based findings and study the natural history of IM.

The aim of this study was to develop a CPR to aid in the clinical diagnosis of IM and externally validate in an independent population.

METHODS
Cohorts and procedures
The derivation cohort was recruited from seven university student health centres in Ireland. Recruitment took place from September 2017 until May 2019. The external validation cohort was recruited from a student health centre at the University of Georgia, USA, between September 2015 and January 2019.

For the derivation cohort, inclusion criteria for participants were those aged over 18 years (who were able to give informed consent) presenting with sore throat and at least one other of the following symptoms: malaise, fatigue, lymphadenopathy, fever, headache, symptom duration 7 days or less. Patients were excluded if they suffered from a medical condition or treatment associated with significant impaired immunity as determined by the recruiting physician; if they had health literacy or language difficulties. For the validation cohort, participants were included if their treating physician had a clinical suspicion of IM and ordered a diagnostic test.

The outcome measure is the diagnosis of acute IM caused by Epstein-Barr virus (EBV). In the derivation cohort, the reference standard test was positive EBV serology determined by a positive anti-VCA (viral capsid antigen) IgM result. In the validation cohort, positive EBV status was determined by a positive heterophile antibody test.[1] In general, EBV antibody testing is more sensitive than heterophile antibody testing, and ideally, the diagnostic tests in the derivation and validation populations would be the same; however, this was not possible in this study for pragmatic reasons.

Explanatory variables considered for inclusion in the derivation CPR model were based on the four most likely clinical variables (presence of enlarged posterior cervical lymph nodes, presence of enlarged inguinal or axillary lymph nodes, palatine petechiae, splenomegaly) based on a systematic review of signs and symptoms for the diagnosis of IM.[1 2] Further variables including tonsillar exudate and presence of fever, based on consensus discussion between clinical colleagues, were also included. The validation population had similar characteristics to the derivation cohort, but many of the variables collected in the derivation cohort were not present in the validation cohort. Hence, for the purposes of validation, only one of the four alternative derived CPR models proposed was subject to external validation.

Sample size for the derivation study was calculated using a ratio of 1:4 of cases to non-cases. The guidance for estimating the sample size required to derive a CPR advises a minimum of 10 participants with the outcome and 10 participants without the outcome for each explanatory variable used.[8] Using the six most likely variables, a sample size of 300 participants (with sore throat at presentation) was calculated. For the validation study, there is no agreed consensus on determining an adequate sample size in external validation studies; however, a minimum of 100 events and 100 non-events is recognised as an acceptable sample size when externally validating a CPR.[9]

From the literature, EBV IM accounts for approximately 8% of cases of sore throat. However, the validation population differed, in that it consisted of patients in whom IM was suspected and the treating clinician requested a diagnostic test. As previously mentioned, approximately one-third of tests performed for EBV IM have a positive result. Using this 1:2 ratio, for the purposes of calculating sample size, it was estimated that a sample size of approximately 300 would be required to yield 100 events of EBV IM.

Missing values for each factor were tabulated and multiple imputation considered. However, due to the relatively small number of missing values, the analysis included only participants with complete data.

Descriptive statistics, univariable and multivariable associations
Descriptive statistics (mean, standard deviation (SD), frequency, percentage) were generated for patients testing positive and negative for EBV. An assessment of the diagnostic effect of each explanatory variable was considered separately, with the results expressed as odds ratio (OR) and associated 95% confidence intervals (CIs). A threshold p value of ≤0.15 in the univariate analysis, or variables which were considered important following literature review and/or consensus, were then included in a multivariable logistic regression model.

Four multivariable models were developed as CPRs and transformed to point-based rules. For two of the models, points/scores were assigned by rounding the regression coefficients to the nearest integer. For the other two models, the regression coefficients were first multiplied by 10 and then rounded to the nearest integer. In model 1, all explanatory variables of prior clinical importance or associated with a threshold p value of ≤0.15 in

the univariable analysis were included in the model. In model 2, stepwise backward regression was used to optimise the prediction model through simplification. The same variables were included as in model 1, but a backward selection was used with a p value threshold of 0.05. This simplification improves the practicality of use for the physician, by reducing the number of components of the CPR that need inputting. Lastly, the impact of global clinical assessment by the examining physician has been shown to be important in relation to the diagnosis of pneumonia.[10] For this reason, the impact of the explanatory variable 'clinical impression' on overall CPR performance was assessed in two further models (models 3 and 4) by excluding the 'clinician impression' variable from model 1 and model 2, respectively.

Discrimination, calibration and external validation
There are two main ways to assess how well a clinical prediction rule model performs: discrimination and calibration. Discrimination is a model's ability to differentiate between individuals at higher risk and lower risk of having an outcome. Discrimination of the CPRs was quantified using the area under the receiver operating characteristic curve statistic (AUC) and 95% CI. An AUC of 0.5 represents chance, 0.7–0.9 represents moderate discrimination and 1.0 represents perfect discrimination.[11]

Calibration is the agreement between the predicted absolute risk and the observed risk for the outcome of interest. Calibration is measured by assessing the predicted and observed risk at different strata of risk. In this study, initial internal calibration was carried out by means of split sampling in the derivation cohort and assessed visually and using the Hosmer-Lemeshow test for goodness of fit, which examines whether the difference between the predicted and observed outcome can be explained by chance.[11] Subsequent, external calibration was then carried out using the US population of participants on CPR model 4 again using the Hosmer-Lemeshow test for goodness of fit.[11]

Likelihood ratios, post-test probabilities and further diagnostic testing
Likelihood ratios (LRs) and subsequent post-test probabilities for EBV were calculated for all the models, to assess the extent to which using the different CPR models changed the probability of a positive diagnosis of EBV using a Bayesian approach.[12 12]

Based on a previous systematic review, the impact of further diagnostic testing after application of CPRs was assessed incorporating the variable 'atypical lymphocytosis >10%'. An LR for atypical lymphocytosis was calculated from the univariable analysis. This LR was applied in series, following the CPR, adopting a Bayesian approach.[12 13] The previously calculated post-test probabilities for all four models now became pretest probabilities, and the LR for atypical lymphocytosis was applied to these to arrive at new post-test probabilities.[12 13]

Data were analysed using Stata software V.16 (College Station, Texas, USA) for statistical analysis.[14]

Patient and public involvement
None.

RESULTS
Descriptive statistics, univariable and multivariable associations
In the derivation cohort, there were 328 participants, of whom 42 (12.8%) had a positive EBV serology test result. Participants' ages ranged from 17 to 39 years old, with 225 (68.6%) female. From unadjusted logistic regression analyses (table 1), the following clinical variables were identified (p≤0.15) for inclusion in the multivariable analysis: presence of enlarged/tender posterior cervical lymph nodes, presence of exudate on tonsils, presence of splenomegaly, absence of shortness of breath, absence of feeling generally unwell, presence of exudate on pharynx, atypical lymphocytosis and 'clinician impression'. In the validation population of 2342 participants, 379 (16.2%) had positive heterophile antibody tests.

We fitted four alternative CPR models (models 1–4) by means of multivariable analysis with results expressed as estimated regression coefficients and 95% CIs (table 2). These regression coefficients were then converted to integer-based CPR scores/points (table 2).

Discrimination, calibration and external validation
The discrimination and calibration were investigated for the four CPR models (table 1, figures 1 and 2). In terms of discrimination, the AUC values ranged from 0.70 to 0.80 for all models, indicating moderate discrimination between those with and those without EBV (figure 1). The number of expected EBV cases (expected) was compared with the actual number of EBV cases (observed) and shown in figure 2. The Hosmer-Lemeshow goodness-of-fit test showed no evidence of a significant difference between the expected and the observed risk of EBV for all models, indicating good calibration (table 1).

A summary of symptoms and signs in the derivation and validation cohorts is presented in table 3.

It is evident from table 3 that there were clear differences between the participants with EBV and the overall population for some of the variables, in particular for the variables 'enlarged/tender posterior cervical lymph nodes', 'exudate on pharynx' and 'atypical lymphocytosis'. The EBV CPR score using model 4 was calculated for each individual in the validation cohort. Figure 3 shows the AUC which was 0.69 (95% CI: 0.67–0.72) and the Hosmer-Lemeshow goodness-of-fit test showed no evidence of a significant difference between the expected and the observed risk of EBV (χ^2=1.94; p=0.378), indicating good calibration (figure 4).

Table 1 Summary of variables and univariable associations between explanatory variables and EBV positivity in derivation cohort

Symptom/sign	EBV positive n=42 N (%)*	Total n=328 N (%)*	Univariable analysis OR (95% CI)	P value
Temperature >38†	2 (6.5)	11 (3.4)	1.85 (0.38–9.00)	0.44
Exudate on tonsils†	25 (61.0)	162 (49.4)	1.68 (0.86–3.27)	0.13
Splenomegaly†	2 (5.1)	3 (0.9)	14.81 (1.31–167.37)	0.03
Palatal petechiae†	11 (27.5)	92 (28.1)	0.95 (0.45–1.98)	0.88
Enlarged/tender posterior cervical lymph nodes†	22 (52.4)	76 (23.2)	4.71 (2.40–9.23)	<0.01
Enlarged/tender inguinal lymph nodes†	2 (5.1)	15 (4.6)	0.96 (0.21–4.42)	0.96
Enlarged/tender axillary lymph nodes†	2 (5.1)	11 (3.4)	1.51 (0.32–7.28)	1.43
Gender‡ (reference: male)				0.98
Male	15 (35.7)	119 (36.4)	0.99 (0.50–1.95)	
Female	26 (61.9)	208 (63.6)		
Cough	18 (43.9)	148 (45.1)	0.93 (0.48–1.79)	0.82
Nasal congestion	21 (51.2)	185 (56.4)	0.77 (0.40–1.48)	0.43
Shortness of breath (absence of shortness of breath for univariable analysis)	7 (17.1)	90 (27.4)	2.02 (0.86–4.73)	0.12
Feeling generally unwell (absence of feeling generally unwell for univariable analysis)	33 (80.5)	285 (86.9)	1.91 (0.81–4.49)	0.14
History of fever	28 (68.3)	198 (60.4)	1.41 (0.70–2.83)	0.34
Headache	27 (65.9)	192 (58.5)	1.39 (0.70–2.77)	0.35
Muscle aches	21 (51.2)	163 (49.7)	1.04 (0.54–2.00)	0.92
Sweats/chills	28 (68.3)	217 (66.2)	1.07 (0.53–2.16)	0.85
Fatigue	36 (87.8)	266 (81.1)	1.63 (0.61–4.35)	0.33
Not sleeping well	28 (68.3)	228 (69.5)	0.89 (0.44–1.81)	0.76
Dizziness	10 (25.0)	103 (31.4)	0.69 (0.32–1.46)	0.33
Poor appetite	22 (53.7)	196 (59.8)	0.73 (0.38–1.41)	0.35
Exudate on pharynx	18 (43.9)	77 (23.5)	3.00 (1.52–5.92)	<0.01
Enlarged anterior cervical lymph nodes	37 (88.1)	278 (84.8)	1.35 (0.50–3.63)	0.55
Tender anterior cervical lymph nodes	33 (78.6)	241 (73.5)	1.36 (0.62–2.97)	0.44
Hepatomegaly	0	4 (1.2)	Not applicable—only 4 positive (controls)	
Respiratory rate 0–11	2 (7.7)	6 (2.6)	Reference	
Respiratory rate 12–16	21 (80.8)	165 (71.4)	0.29 (0.50–1.69)	0.17
Respiratory rate 17–28	3 (11.5)	60 (26.0)	0.11 (0.01–0.82)	0.03
Atypical lymphocytosis >10%	26 (63.4)	29 (8.8)	162.34 (44.1–597.7)	<0.01
Clinician impression EBV positive	12 (28.6)	29 (8.8)	8.13 (3.43–19.28)	<0.01
	Mean (SD)	**Mean (SD)**		
Age	20.69 (3.94)	20.57 (2.78)	1.02 (0.91–1.13)	0.78
Pulse	82.82 (13.92)	81.13 (11.36)	1.01 (0.98–1.05)	0.45

*Calculations exclude those with missing data for each variable.
†Clinically important variables identified in the literature and from clinical consensus.
‡The gender variable from the gender-neutral person was omitted from analysis, as any associations found in relation to just one participant would not lead to accurate conclusions. However, all other data from this participant were used.
EBV, Epstein-Barr virus.

Table 2 Multivariable analysis of alternative clinical prediction rule models 1–4

Explanatory variable	Model 1* Regression coefficient (95% CI)	Points	Model 2† Regression coefficient (95% CI)	Points	Model 3‡ Regression coefficient (95% CI)	Points	Model 4§ Regression coefficient (95% CI)	Points
Presence of a temperature >38	1.14 (–0.68 to 2.97)	11			1.13 (–0.61 to 2.88)	11		
Presence of exudate on tonsils	0.28 (–0.83 to 1.38)	3			0.06 (–0.98 to 1.09)	1		
Presence of splenomegaly	1.98 (–1.57 to 5.53)	20			3.50 (–0.19 to 7.18)	35		
Presence of palatal petechiae	0.21 (–0.87 to 1.30)	2			0.16 (–0.81 to 1.13)	2		
Presence of enlarged/tender posterior cervical lymph nodes	1.73 (0.70 to 2.76)	17	4.55 (0.42 to 2.68)	5	1.60 (0.66 to 2.55)	16	1.56 (0.69 to 2.44)	2
Absence of enlarged/tender inguinal lymph nodes	0.40 (–2.18 to 2.97)	4			0.51 (–2.32 to 3.33)	5		
Presence of enlarged/tender axillary lymph nodes	0.35 (–2.21 to 2.92)	4			0.49 (–2.17 to 0.15)	5		
Absence of shortness of breath	1.38 (–0.17 to 2.92)	14			1.18 (–0.18 to 2.53)	12		
Feeling generally unwell	1.12 (–0.23 to 2.48)	11			0.82 (–0.47 to 2.09)	8		
Clinician impression	1.92 (0.47 to 3.37)	19	1.67 (0.71 to 2.62)	2				
Presence of exudate on pharynx	0.69 (–0.51 to 1.89)	7			0.93 (–0.16 to 2.03)	9	0.98 (0.08 to 1.87)	1
Hosmer-Lemeshow goodness of fit	χ^2=3.15; p=0.87		χ^2=5.23; p=0.07		χ^2=0.26; p=0.88		χ^2=0.69; p=0.71	
AUC (95% CI)	0.81 (0.75 to 0.85)		0.73 (0.63 to 0.82)		0.77 (0.67 to 0.87)		0.70 (0.62 to 0.79)	

*Explanatory variables had to be of prior clinical importance or be associated with a threshold p value of ≤0.15 in univariate analysis. Score/points assigned to each predictor by multiplying the regression coefficient by 10 and rounding to nearest integer.
†Explanatory variables had to be of prior clinical importance or be associated with a threshold p value of ≤0.15 in univariate analysis, and then a backward selection was used with a p value threshold of 0.05. Score/points assigned to each predictor by rounding the regression coefficient to nearest integer.
‡Explanatory variables had to be of prior clinical importance or be associated with a threshold p value of ≤0.15 in univariate analysis. Clinician impression was not used in this model. Score/points assigned to each predictor by multiplying the regression coefficient by 10 and rounding to nearest integer.
§Explanatory variables had to be of prior clinical importance or be associated with a threshold p value of ≤0.15 in univariate analysis, and then a backward selection was used with a p value threshold of 0.05. Clinician impression was not included in this model. Score/points assigned to each predictor by rounding the regression coefficient to nearest integer.
AUC, area under the receiver operating characteristic curve.

LRs, post-test probabilities and further diagnostic testing

Online supplemental tables 1 and 2 show the positive LRs and post-test probabilities in the derivation cohort for the four alternative CPR models (1–4), based on a pretest probability estimate of 8%. Online supplemental table 3 shows the LRs and post-test probabilities for positive EBV in validation cohort, based on a range of pretest (prior) probabilities. Online supplemental tables 4 and 5 show that inclusion of the positive diagnostic test result of 'atypical lymphocytosis >10%' with a positive LR of 60.03 has a substantial impact on all CPR models with post-test probabilities all greater than 80% for patients with the lowest level of pretest probability (again, based on the derivation cohort). Finally, online
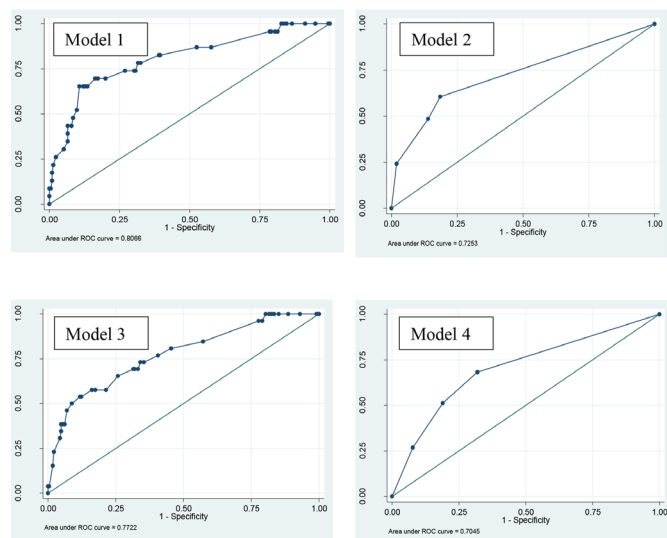
Figure 1 Receiver operating characteristic (ROC) curve for models 1–4; a graph of the sensitivity (y-axis) and the specificity (x-axis).

supplemental table 6 presents total scores/points and the number and percentage of EBV-positive and negative participants for each model.

DISCUSSION
Summary

We present four alternative multivariable CPRs that can be used to aid clinicians when assessing young adults with prolonged symptoms of fever, sore throat, lymphadenopathy or malaise where the possibility of IM arises as part of a differential diagnosis workup. The AUC values for the four models ranged from 0.70 to 0.80 indicating moderate discrimination between those with and without EBV. The more complex CPR models (1 and 3) provide slightly higher discrimination metrics; these CPRs can be integrated into computer-based clinical decision support
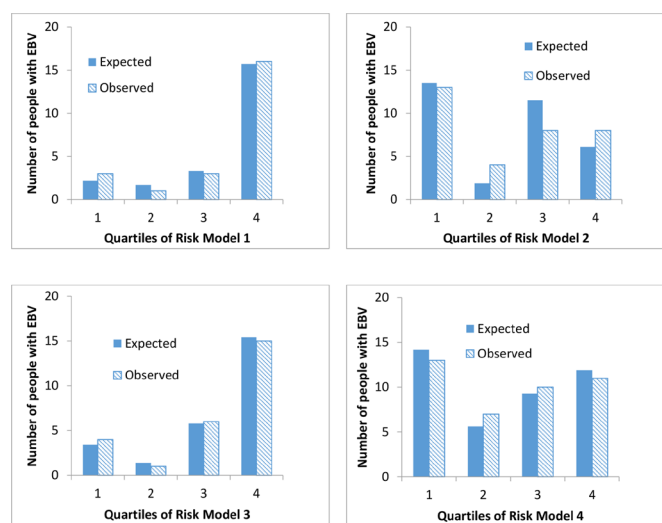
systems (CDSS) and used as diagnostic aids via electronic health record software though do require collection of more clinical data than might routinely be considered in clinical practice.[15] The briefer CPR models are also presented and provide diagnostic decision-making evidence by applying a minimal number of clinical variables converted as integer-based CPR scores.

Consistent with previous systemic reviews, incorporating atypical lymphocytosis (applied as a diagnostic test in series after initial CPR assessment) produces high post-test probabilities for EBV ranging from 94.8% to 99.0%, 'ruling in' EBV IM. Even applying atypical lymphocytosis to a low or zero CPR score produced a post-test probability of >80% (with relatively narrow CIs) for all four CPR models, confirming the high diagnostic value of this test.[2]

The simple, two-item (model 4) CPR had reasonable discrimination (AUC 0.69, 95% CI: 0.67–0.72) and good calibration when externally validated in a separate US population. In terms of clinical value, post-test probabilities were calculated from a range of plausible pretest probabilities (8%–16%). The results demonstrated that the post-test probability of EBV would be approximately 40% if a patient has a pretest probability of 16% and both CPR variables are present (online supplemental table 3). In this context, further diagnostic testing with a more specific blood test (positive anti-VCA IgM or IgG) would be required.[12] If one or both of the variables from the model 4 CPR are present, the patient may be categorised as intermediate risk of EBV IM diagnosis, and the clinician may be more inclined to perform a definitive test, whereas if neither variable is present, the patient may be categorised as low risk, allowing the clinician to adopt a 'watch-and-wait' approach. Ultimately, the test/treatment and test/reassurance thresholds will depend on several factors—availability and cost of diagnostic testing and patient's and clinician's willingness to accept diagnostic uncertainty without testing all patients who present with possible IM.

Strengths and limitations

Several limitations in relation to this study need to be considered. We only examined IM where the underlying cause was EBV and did not consider other known pathogens such as cytomegalovirus or herpes simplex virus, which may account for between 5% and 10% of IM cases.[1 2] The derivation and validation cohorts differ in terms of their inclusion criteria and in terms of the reference standard diagnostic test used though both cohorts targeted similar populations presenting with symptoms associated with IM. Broader inclusion criteria in the validation cohort meant that a higher proportion of participants tested positive for EBV. However, lower specificity of the heterophile antibody test in the validation cohort may have resulted in more false-negative results. Lastly, while we included 28 variables, we did not assess all known clinical variables that could be possibly associated with IM. This was on pragmatic grounds as participating physicians recruited, examined and reported on symptoms and

Figure 2 Internal validation showing observed and expected risk of EBV for models 1–4. EBV, Epstein-Barr virus.

Table 3 Summary of clinical variables and EBV positivity in derivation and validation cohorts

Symptom/sign	Derivation IM positive n=42 N (%)	Total n=328 N (%)	Validation IM positive n=379 N (%)	Total n=2342 N (%)
Temperature >38	2 (6.5)	11 (3.4)	11 (2.9)	153 (6.5)
Exudate on tonsils	25 (61.0)	162 (49.4)	131 (34.6)	451 (19.3)
Enlarged/tender posterior cervical lymph nodes	22 (52.4)	76 (23.2)	190 (50.1)	509 (21.7)
Cough	18 (43.9)	148 (45.1)	125 (33.0)	1012 (43.2)
Nasal congestion	21 (51.2)	185 (56.4)	163 (43.0)	1084 (46.3)
Headache	27 (65.9)	192 (58.5)	140 (36.9)	995 (42.5)
Muscle aches	21 (51.2)	163 (49.7)	83 (21.9)	657 (28.1)
Sweats/chills	28 (68.3)	217 (66.2)	159 (42.0)	964 (41.2)
Fatigue	36 (87.8)	266 (81.1)	195 (51.5)	1114 (47.6)
Exudate on pharynx	18 (43.9)	77 (23.5)	126 (33.2)	437 (18.7)
Enlarged/tender anterior cervical lymph nodes	38 (90.5)	284 (86.6)	312 (82.3)	1830 (78.1)
Atypical lymphocytosis >10%	26 (63.4)	29 (8.8)	280 (73.9)	334 (14.3)

EBV, Epstein-Barr virus; IM, infectious mononucleosis.

signs during real-time clinical practice without the help of a research assistant. Despite these limitations, there was broad similarity in terms of clinical features between the two cohorts (table 3) enabling external validation of the derived CPR.

Comparison with existing literature

The findings of this study add further evidence to previous systematic reviews about the signs and symptoms of IM caused by EBV in community settings.[2 7] Though there are currently no widely used CPRs for IM, many of the clinical variables in our four proposed CPR models have been identified as being of diagnostic value in previous literature.[2 7]

In addition, in the present study, we have been able to estimate the independent effects of symptoms and signs

by means of multivariable logistic regression. Previous systematic reviews, though useful, have only reported the diagnostic value of individual symptoms and signs based on univariable analysis.[2 7] Deriving and externally validating a CPR for IM provide clinicians with four alternative CPRs that can be integrated as part of a computer-based CDSS (models 1 and 3) or as a simpler point-scoring CPR (models 2 and 4).

The data collected may also be useful in the context of CPRs for sore throat caused by group A beta haemolytic *Streptococcus*, as diagnostic testing was performed for this on all participants also.

Implications for research and practice

The CPR would benefit from further validation studies in more diverse populations, particularly in terms of recruiting sufficient number of higher-risk patients who have most of the symptoms and signs of IM. Lastly, the validated CPRs should be subject to impact analysis by means of randomised controlled trials that assess the use of IM on patient outcome.

The CPRs developed provide additional diagnostic evidence about young adults who present to their family
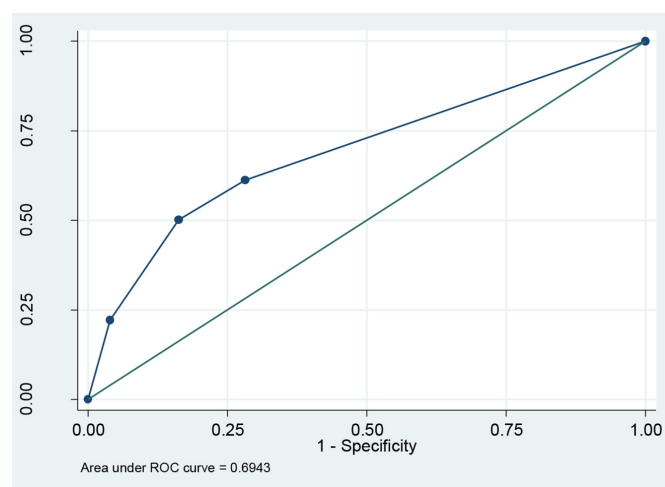
Area under ROC curve = 0.6943

Figure 3 External validation of model 4. ROC, receiver operating characteristic.
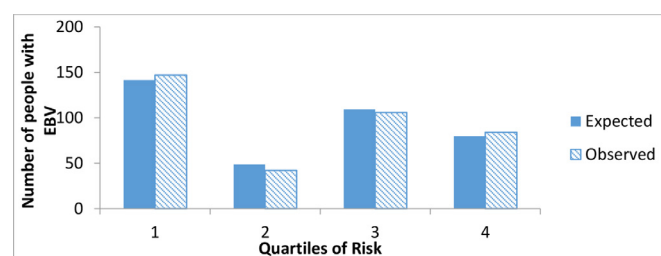
Figure 4 External validation of model 4—observed and expected risk of Epstein-Barr virus (EBV).

physician with symptoms suggestive of IM. Completing either the longer (models 1 and 3) or shorter (models 2 and 4) CPRs will provide clinicians with a means of estimating EBV infection and deciding on further diagnostic testing.[16] It seems reasonable that initial diagnostic testing should be performed to establish the presence of lymphocytosis and/or atypical lymphocytosis, as this is a readily available and low-cost test. More specific diagnostic testing based on positive anti-VCA immunoglobulin response (IgM or IgM) can be reserved for situations of remaining diagnostic uncertainty. Where the diagnostic threshold for confirming IM lies will depend on individual preferences and availability and cost of diagnostic tests. Although the CPR for diagnosis of IM caused by EBV demonstrates moderate rather than high discrimination, it may help clinicians stratify patients into low, intermediate and high-risk groups. Additional diagnostic testing in series by examining the presence of atypical lymphocytosis will aid in diagnostic risk stratification. In terms of clinical decision-making, application of a CPR is likely to be most useful in performing an important 'rule out' function. There are no predefined test/treat thresholds for IM—unlike some other conditions, for example, influenza, acute coronary syndrome and community-acquired pneumonia.[13 17] However, if neither CPR variable is present in a patient, especially if the patient's age is under 15 or over 25 years (at which point the pretest probability drops to approximately 2% or less), the clinician may be reasonably satisfied that the post-test probability is low enough for them to manage the patient expectantly, or consider an alternative diagnosis.[18] The LRs from having both variables in the CPR are too low (post-test probability at 23.2% for the age range 16–20 years, and lower at other ages) for a clinician to diagnose IM caused by EBV without ordering a triage (atypical lymphocytosis) or more definitive VCA serology test. Use of IM CPRs will provide greater explicit quantitative information about these implicit probability estimates.[12]

CONCLUSION

We have developed four simple CPRs to help family physicians identify which of their symptomatic patients have EBV-associated IM. External validation of a simple two-item CPR (model 4) shows good calibration. These CPRs should be subject to further external validation and impact analysis, but provide an evidence-based structure for assessment and further diagnostic testing of younger patients with suspected IM in community settings.

ORCID iDs
Muireann de Paor http://orcid.org/0000-0001-5387-931X
Fiona Boland http://orcid.org/0000-0003-3228-0046
Mark H Ebell http://orcid.org/0000-0003-3228-2877
Tom Fahey http://orcid.org/0000-0002-5896-5783

REFERENCES

1. Luzuriaga K, Sullivan JL. Infectious mononucleosis. *N Engl J Med* 2010;362:1993–2000.
2. Ebell MH, Call M, Shinholser J, *et al*. Does this patient have infectious mononucleosis?: the rational clinical examination systematic review. *JAMA* 2016;315:1502–9.
3. Buchwald DS, Rea TD, Katon WJ, *et al*. Acute infectious mononucleosis: characteristics of patients who report failure to recover. *Am J Med* 2000;109:531–7.
4. McGinn TG, Guyatt GH, Wyer PC, *et al*. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. evidence-based medicine Working group. *JAMA* 2000;284:79–84.
5. Galvin R, Joyce D, Downey E, *et al*. Development and validation of a clinical prediction rule to identify suspected breast cancer: a prospective cohort study. *BMC Cancer* 2014;14:743.
6. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006;144:201–9.
7. Cai X, Ebell MH, Haines L. Accuracy of signs, symptoms, and hematologic parameters for the diagnosis of infectious mononucleosis: a systematic review and meta-analysis. *J Am Board Fam Med* 2021;34:1141–56.
8. Moons KGM, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
9. Vergouwe Y, Steyerberg EW, Eijkemans MJC, *et al*. Substantial effective sample sizes were required for external validation

studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58:475–83.

10 Dale AP, Marchello C, Ebell MH. Clinical gestalt to diagnose pneumonia, sinusitis, and pharyngitis: a meta-analysis. *Br J Gen Pract* 2019;69:e444–53.

11 Alba AC, Agoritsas T, Walsh M, *et al*. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* 2017;318:1377–84.

12 Doust J. Using probabilistic reasoning. *BMJ* 2009;339:b3823.

13 Ebell MH, Locatelli I, Senn N. A novel approach to the determination of clinical decision thresholds. *Evid Based Med* 2015;20:41–7.

14 StataCorp. *Stata statistical software. release 16*. 2019.

15 Falk G, Fahey T. Clinical prediction rules. *BMJ* 2009;339:b2899.

16 Bossuyt PM, Irwig L, Craig J, *et al*. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ* 2006;332:1089–92.

17 Ebell MH, Locatelli I, Mueller Y, *et al*. Diagnosis and treatment of community-acquired pneumonia in patients with acute cough: a quantitative study of decision thresholds in primary care. *Br J Gen Pract* 2018;68:e765–74.

18 Del Mar C, Pincus D. Incidence patterns of respiratory illness in Queensland estimated from sentinel general practice. *Aust Fam Physician* 1995;24:625–9.