



Which tools best predict the incidence of work-related sensitisation and symptoms

E Suarthana, J-L Malo, D Heederik, H Ghezzi, J L'Archevêque and D Gautrin

Occup. Environ. Med. 2009;66;111-117; originally published online 18 Nov 2008;
doi:10.1136/oem.2008.041079

Updated information and services can be found at:

<http://oem.bmj.com/cgi/content/full/66/2/111>

These include:

Data supplement

"web only appendix"

<http://oem.bmj.com/cgi/content/full/oem.2008.041079/DC1>

References

This article cites 23 articles, 13 of which can be accessed free at:

<http://oem.bmj.com/cgi/content/full/66/2/111#BIBL>

1 online articles that cite this article can be accessed at:

<http://oem.bmj.com/cgi/content/full/66/2/111#otherarticles>

Rapid responses

You can respond to this article at:

<http://oem.bmj.com/cgi/eletter-submit/66/2/111>

**Email alerting
service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Occupational and Environmental Medicine* go to:

<http://journals.bmj.com/subscriptions/>

Which tools best predict the incidence of work-related sensitisation and symptoms

E Suarathana,^{1,2,3} J-L Malo,¹ D Heederik,² H Ghezzo,¹ J L'Archevêque,¹ D Gautrin¹

► Appendices are published online only at <http://oem.bmj.com/content/vol66/issue2>

¹ Department of Chest Medicine, Hôpital du Sacre-Coeur de Montreal, 5400 West Gouin Blvd, Montreal, Canada H4J 1C5; ² IRAS (Institute for Risk Assessment Sciences), Environmental Epidemiology Division, Utrecht University, Utrecht, The Netherlands; ³ Community Medicine Department, Faculty of Medicine, University of Indonesia. Jl. Pegangsaan Timur 16, 10320, Jakarta Pusat, Indonesia

Correspondence to: Dr Denyse Gautrin, Department of Chest Medicine, Hôpital du Sacre-Coeur de Montreal, 5400 West Gouin Blvd, Montreal, Canada H4J 1C5; d.gautrin@umontreal.ca

Accepted 1 August 2008
Published Online First
18 November 2008

ABSTRACT

Background/aim: This study used information from the questionnaire alone or in conjunction with clinical tests, such as skin-prick testing (SPT) and bronchial responsiveness (BR) testing at entry, to develop models for estimating the probability of the occurrence of specific IgE-sensitisation to and respiratory symptoms in contact with laboratory animal (LA) allergens after 32 months' training in an animal health technology programme.

Methods: Four multivariable logistic regression models were developed for each endpoint, consisting of: (1) questionnaire; (2) questionnaire and SPT; (3) questionnaire and BR testing; and (4) questionnaire, SPT and BR testing. The prognostic models were derived from a cohort of Canadian animal health technology apprentices. The models' internal validity and diagnostic accuracy were evaluated and compared.

Results: Symptoms indicative of asthma and allergic symptoms at baseline composed the final questionnaire model for the occurrence of occupational sensitisation and symptoms. Both questionnaire models showed a good discrimination (area under the receiver operating characteristics curve were 0.73 and 0.78, respectively) and calibration (Hosmer-Lemeshow test p value >0.10). Addition of SPT and/or BR testing increased the specificity of the questionnaire model for LA sensitisation, but not for symptoms at work. To facilitate their application in practice, the final questionnaire models were converted to easy-to-use scoring system.

Conclusions: Questionnaire is an easy tool that can give accurate prediction of the incidence of occupational sensitisation and symptoms.

The attributable risk from occupation for asthma in population-based studies ranges from 5% to 15%,¹⁻⁶ which makes work-related asthma an important public health concern. Several tools are used in the diagnosis of occupational asthma (OA), defined as a type of work-related asthma specifically caused by an agent present at the workplace.⁷ These include a questionnaire, skin-prick testing, assessment of non-specific bronchial responsiveness, examination of induced sputum, serial monitoring of peak expiratory flows and specific inhalation challenges.⁷ Where these approaches are used in clinical settings for diagnostic purposes, little is known about their prognostic value in predicting the future occurrence of occupational allergic diseases in open populations, either in general or at specific workplaces. If surveillance or even screening is to be applied in high-risk workplaces, which tools should be used when the worker is hired: a questionnaire alone or its combination with information on atopic status and bronchial responsiveness?

Recently, a multivariable prognostic model for sensitisation to laboratory animal (LA) allergens was proposed.⁸ The authors used information carried by simple questionnaire items independently or in combination with atopic status at baseline to estimate the probability of the occurrence of occupational sensitisation in three years of follow-up. Nevertheless, the predictive value of these tools with regard to the occurrence of respiratory symptoms at work has not yet been assessed.

Since 1991, we have conducted prospective studies in apprentices exposed to various occupational agents, including high-molecular-weight products—for example, animal-derived allergens, latex and cereals.⁹⁻¹² Our purpose was to investigate the natural history of occupational allergy and asthma. Students in an animal health technology programme showed a high incidence (8.9 per 100 person-years of sensitisation to LA allergens.¹⁰ This group also showed an incidence of probable OA of 2.7 per 100 person-years.¹¹

We used information from a questionnaire alone or in conjunction with clinical tests, such as skin-prick testing (SPT) and bronchial responsiveness (BR) testing at entry, to develop models for estimating the probability of the occurrence of specific IgE-sensitisation to and respiratory symptoms in contact with LA allergens after 32 months training in an animal health technology programme. For their practical application, the models were transformed into easy-to-use score charts. The use of such prognostic models among apprentices exposed to LA allergens may support coaching and job referral, decision-making in medical surveillance and enable early identification of individuals who need support through preventive measures.

METHODS

Study design and population

The prognostic model was developed in Canadian students in an animal health technology training programme from the cohort apprentice study.¹⁰ The apprentices answered questionnaires in a face-to-face interview and were subjected to SPT, lung function and bronchial non-specific provocation tests upon beginning their three-year apprenticeship programme. The tests were subsequently administered yearly until the end of their programme. Informed consent was obtained from each subject and the study was performed after approval from the Sacre-Coeur Hospital's ethics committee in accordance with Canadian ethical rules.

Original article

Table 1 Associations between the predictors and sensitisation to laboratory animal allergens from multivariable regression modelling

	Total (n = 314)	Not sensitised (n = 265)	Sensitised (n = 49)	OR (95% CI)
Age mean (SD) years	19.5 (9.4)	19.6 (3.1)	19.1 (2.7)	1.0 (0.9 to 1.1)
Male gender	43 (13.7)	43 (16.2)	0 (0.0)	—
Smoker	38 (12.1)	31 (11.7)	7 (14.3)	1.3 (0.5 to 3.0)
Symptoms indicative of asthma*	139 (44.3)	106 (40.0)	33 (67.3)	3.1 (1.6 to 5.9) **
Symptoms indicative of COPD†	30 (9.6)	25 (9.4)	5 (10.2)	1.1 (0.4 to 3.0)
Allergic symptoms‡	140 (44.6)	102 (38.5)	38 (77.6)	5.5 (2.7 to 11.3) **
Skin symptoms¶	60 (19.1)	48 (18.1)	12 (24.5)	1.5 (0.7 to 3.1) **
Respiratory symptoms at work§	35 (11.1)	22 (8.3)	13 (26.5)	4.0 (1.9 to 8.6) **
Positive SPT to common allergens††	119 (37.9)	85 (32.1)	34 (69.4)	4.8 (2.5 to 9.3) **
Measurable PC ₂₀ (≤ 32 mg/ml)	106 (33.8)	74 (27.9)	32 (65.3)	4.9 (2.5 to 9.3) **

Data presented as n (%) unless otherwise stated.

*Present if experienced at least one of the following: wheezing or whistling in the last 12 months; had an attack of shortness of breath in the last 12 months; asthma confirmed by physician; and/or start to cough induced by exercise, strenuous work, cold air, heavy smell, smoke or dust.

†Present if experienced at least one of the following: sometimes wake with a feeling of tightness in your chest first thing in the morning; awoken by cough at night in the last few months; bring up phlegm from chest first thing in the morning for at least 3 months each year; and/or had chronic bronchitis confirmed by doctor.

‡Present if experienced at least one of the following: develop eye or nasal or respiratory or skin symptoms when exposed to common allergen such as dust mite, animal hair or pollen; have an itchy runny nose or sneezing even when do not have a cold; and/or ever had "hay fever".

¶Present if experienced urticaria and/or eczema

§Present if experienced at least one of the following symptoms at work: get itchy, red and/or watery eyes; get runny or stuffy nose or sneezing; get itchy, redness or rash on your skin; feel shortness of breath; start to cough; and/or start to wheeze.

††Defined as two positive skin reactions to a set of 11 common inhalants.

**Univariable p value < 0.50.

95% CI, 95% confidence interval; COPD, chronic obstructive pulmonary diseases; OR, odds ratio; PC₂₀, level of the methacholine concentration that caused a decrease in FEV₁ of 20%; SPT, skin-prick test.

The number of subjects included in this analysis was different from the one in a previous publication.¹⁰ In the previous publication, subjects who showed a sensitisation to LA allergen at the initial visit were also included and considered at risk of developing a new sensitisation to another LA allergen. For the purpose of developing prognostic models for sensitisation to LA allergens, we included only apprentices who had a negative skin reaction to all LA allergens tested at the initial visit; rat urine, mouse urine, rabbit urine and/or rabbit hair. Prognostic models for symptoms at work were derived from apprentices who had neither rhinoconjunctivitis nor chest symptoms at work at the initial visit.

Potential predictors

We used questionnaire items, clinical information on atopic status—SPT results to common allergens—and BR to methacholine at baseline as potential prognostic predictors. The questionnaire was derived from the standardised questionnaire of the International Union Against Tuberculosis and Lung Disease (IUATLD) and was administered by a trained nurse, who was blinded to the outcome of SPT and bronchial challenge tests.^{13 14} We used principal components analysis (PCA) to reduce the number of questionnaire predictors to be included in the multivariate logistic regression analysis.¹⁵ PCA is often used to identify underlying variables, or factors, that explain the

Table 2 Distribution and strength of univariable association between the predictors at baseline and sensitisation to LA allergens at the 32nd month of the training programme

	Questionnaire only (model 1)		Questionnaire + SPT (model 2)		Questionnaire + BR (model 3)		Questionnaire + SPT + BR (model 4)	
	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)
Intercept	−2.86		−3.16		−3.44		−3.55	
Symptoms indicative for asthma	0.70	2.2 (1.1 to 4.3)	0.62	1.9 (1.0 to 3.9)	0.61	1.9 (1.0 to 3.8)	0.50	1.7 (0.8 to 3.5)
Allergic symptoms	1.38	4.6 (2.2 to 9.5)	1.10	3.2 (1.5 to 7.0)	1.39	4.3 (2.1 to 9.1)	1.07	3.1 (1.4 to 6.9)
Positive SPT to common allergens	—	—	0.99	2.8 (1.4 to 5.8)	—	—	0.84	2.4 (1.2 to 5.1)
Measurable PC ₂₀ (≤ 32 mg/ml)	—	—	—	—	1.32	4.0 (2.0 to 7.8)	1.21	3.6 (1.8 to 7.2)
ROC area (95% CI) after bootstrapping	0.731 (0.663 to 0.799)		0.754 (0.682 to 0.826)		0.777 (0.710 to 0.844)		0.785 (0.717 to 0.853)	
Hosmer-Lemeshow test (p value)	0.999		0.985		0.991		0.991	

ROC area SPT to common allergen 0.69 (95% CI 0.61 to 0.77).

ROC area measurable PC₂₀ 0.69 (95% CI 0.60 to 0.77).

From the questionnaire model, the individual probability of developing LA sensitisation can be estimated using the following formula:

$P(\text{sensitisation}) = 1/(1 + \exp(-(-2.86 + \text{symptoms indicative for asthma} \times 0.70 + \text{allergic symptoms} \times 1.38)))$.

Predictor is valued as 1 when present and 0 when absent.

β, regression coefficient after multiplication by correction factor from bootstrapping procedure; BR, bronchial responsiveness to methacholine; 95% CI, 95% confidence interval; OR, odds ratio; ROC, receiver operating characteristic; SPT, skin-prick testing.

pattern of correlations within a set of observed variables—for example, questionnaire items on symptoms. From the PCA, we identified clusters of correlated symptoms with Eigenvalues of 1 or more (that is, explaining more variance than a single symptom). Individual symptoms were only included in the clustered variable when factor loadings were higher than 0.4.

There were five clusters of correlated symptoms with Eigenvalues of 1 or more with a cumulative explained variability of 61.1%. The following clusters were identified and named post hoc as adequately as possible: “respiratory symptoms at work”, “symptoms indicative for asthma”, “allergic symptoms”, “symptoms indicative for chronic obstructive pulmonary diseases” and “skin symptoms”. For details regarding symptoms that were included in each cluster see appendix 1 online. A positive cluster was defined as the presence of at least one of the symptoms that composed the cluster.

Methacholine bronchial challenge tests were performed using guidelines slightly modified from those of the European Respiratory Society.¹⁶ The procedure for performing the bronchial challenge test was modified as described elsewhere to take into account the absence of an on-site physician.¹⁷ BR was considered measurable when the level of the methacholine concentration caused a decrease in forced expiratory volume in one second (FEV₁) of 20% (provocative concentration causing a 20% decrease in FEV₁, PC₂₀) of 32 mg/ml or lower—a measurable PC₂₀. We used a measurable PC₂₀ to define BR because this high cut-off value of PC₂₀ showed the strongest association with the endpoints compared to stricter definitions—for example, PC₂₀ 16 mg/ml or lower. SPT based atopy was defined as two positive skin reactions to a set of 11 common inhalants. A wheal diameter of 3 mm or more was regarded as a positive response, in the absence of any reaction to the diluent (glycerine, 50%) and in the presence of a positive reaction to histamine phosphate (1/200 mg/ml).¹⁰

Health endpoints

Incident cases of sensitisation to LA allergens were defined as individuals with a positive SPT response to rat urine, mouse urine, rabbit urine (Pharmacia Allergon AB, Angelholm, Sweden) and/or rabbit hair (Omega, Montréal, Canada) at the 32nd month visit, but no positive skin reactions to any of these LA allergens tested at the initial visit. A positive skin reaction was defined similarly as to common allergens. Incident cases of symptoms at work were defined as individuals with positive respiratory symptoms at work at the 32nd month visit, but no symptoms at the initial visit.

Data analysis

We developed the prognostic models for sensitisation to the LA allergens in 314 apprentices who had negative skin reactions to any of LA allergens at initial visit. The prognostic models for symptoms at work were developed in 296 apprentices with no work-related symptoms at baseline. We started the multi-variable analysis with the newly created clusters from the PCA as well as other personal characteristics with a univariable *p* value lower than 0.50. We used a backward stepwise selection procedure (using *p* < 0.157 for inclusion) to obtain a questionnaire model with the strongest predictors for sensitisation to LA allergens.¹⁸ Subsequently, results from the SPT to common allergens and the bronchial methacholine challenge test were then added to the model. Therefore, four multivariable logistic regression models were developed for each endpoint, consisting of: (1) questionnaire; (2) questionnaire and SPT; (3) questionnaire and BR testing; and (4) questionnaire, SPT and BR testing. The individual predicted probability of developing the outcome was calculated using all models. The model χ^2 of the second, third and fourth models were compared to the first model's χ^2 to see whether the addition of SPT, BR testing or both of them significantly improved the questionnaire model (the level of significance was *p* value < 0.05 using the likelihood ratio test).

The accuracy of the models was quantified using calibration and discrimination measures. Calibration, the agreement between the predicted probabilities and the observed frequencies of sensitisation, was assessed graphically and tested with the Hosmer–Lemeshow (H–L) test, where a *p* value of 0.10 and higher reflects good agreement.¹⁹ The discriminative ability was determined with the area under the receiver operating characteristic (ROC) curve. The ROC area can range from 0.5 (no discrimination) to 1.0 (perfect discrimination). Because the models were derived from the same subjects, the correlation between the models was taken into account when comparing the ROC area between models.²⁰ The internal validity of the models was assessed using the bootstrapping procedure. This procedure gives a correction factor for both the model's ROC area and for the regression coefficients of the predictors in the final model.²¹ The regression coefficients of the predictors in the final model were multiplied by this correction factor to prevent the model from producing optimistic predictions (that is, too high or too low estimations) when applied in future (new) workers/apprentices.

To facilitate the application of the final model for each outcome in practice, the corrected regression coefficients of the predictors in the final prognostic model were converted to easy-to-use numbers. To derive these numbers, the corrected

Table 3 Distribution and the strength of univariable association between the predictors at baseline and symptoms at work at the 32nd month of the training programme

	Total (n = 296)	No symptoms (n = 251)	Had symptoms at work (n = 45)	OR (95% CI)
Age mean (SD) years	18.7 (3.1)	18.6 (3.1)	19.1 (3.6)	1.1 (1.0 to 1.2)**
Male gender	44 (14.9)	37 (14.7)	7 (15.6)	1.1 (0.4 to 2.6)
Smoker	34 (11.5)	29 (11.6)	5 (11.1)	1.0 (0.4 to 2.6)
Symptoms indicative of asthma	125 (42.2)	96 (38.2)	29 (64.4)	2.9 (1.5 to 5.7)**
Symptoms indicative of COPD	27 (9.1)	18 (7.2)	9 (20.0)	3.2 (1.4 to 7.8)**
Allergic symptoms	118 (39.9)	81 (32.3)	37 (82.2)	9.7 (4.3 to 21.8)**
Skin symptoms	49 (16.6)	42 (16.7)	7 (15.6)	0.9 (0.4 to 2.2)
Positive SPT to common allergens	100 (33.8)	75 (29.9)	25 (55.6)	2.9 (1.5 to 5.6)**
Measurable PC ₂₀ (< 32 mg/ml)	92 (31.1)	70 (27.9)	22 (48.9)	2.5 (1.3 to 4.7)**

Data presented as n (%) unless otherwise stated.

**Univariable *p* value < 0.50.

95% CI, 95% confidence interval; COPD, chronic obstructive pulmonary diseases; OR, odds ratio; PC₂₀, level of the methacholine concentration that caused a decrease in FEV₁ of 20%; SPT, skin-prick test.

Original article

Table 4 Associations between the predictors and symptoms at work from multivariable regression modelling

	Questionnaire only (model 1)		Questionnaire + SPT (model 2)		Questionnaire + BR (model 3)		Questionnaire + SPT + BR (model 4)	
	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)
Intercept	-3.15		-3.32		-3.45		-3.40	
Symptoms indicative of asthma	0.62	2.0 (1.0 to 4.0)	0.60	1.9 (0.9 to 3.9)	0.63	1.9 (0.9 to 3.9)	0.55	1.8 (0.9 to 3.7)
Allergic symptoms	1.94	8.5 (3.7 to 19.2)	1.85	7.2 (3.1 to 16.8)	2.02	8.2 (3.6 to 18.6)	1.79	7.2 (3.1 to 16.8)
Positive SPT to common allergens	—	—	0.51	1.7 (0.8 to 3.5)	—	—	0.39	1.5 (0.7 to 3.2)
Measurable PC ₂₀ (≤ 32 mg/ml)	—	—	—	—	0.63	1.9 (1.0 to 3.9)	0.53	1.8 (0.8 to 3.7)
ROC area (95% CI) after bootstrapping	0.780 (0.713 to 0.847)		0.781 (0.713 to 0.849)		0.790 (0.721 to 0.859)		0.787 (0.719 to 0.855)	
Hosmer-Lemeshow test (p value)	0.999		0.999		0.991		0.993	

ROC area (95% CI) SPT to common allergen 0.65 (0.56 to 0.73).

ROC area (95% CI) Measurable PC₂₀ 0.61 (0.51 to 0.70).

From the questionnaire model, the individual probability of developing symptoms at work can be estimated using the following formula:

$P(\text{symptoms at work}) = 1 / (1 + \exp(-(-3.15 + \text{symptoms indicative for asthma} \times 0.62 + \text{allergic symptoms} \times 1.94)))$.

Predictor is valued as 1 when present and 0 when absent.

β , regression coefficient after multiplication by correction factor from bootstrapping procedure; BR, bronchial responsiveness to methacholine; 95% CI, 95% confidence interval; OR, odds ratio; ROC, receiver operating characteristic; SPT, skin-prick testing.

coefficients were divided by the smallest one and rounded to the nearest half integer. The discriminative accuracy of this scoring rule was again assessed. Finally, the sums of the scores were related to their corresponding probabilities of having the endpoint. A cut-off point of the sum scores was introduced to divide the apprentices into a group with a low versus high probability of having the endpoint. All analyses were performed using SPSS 15.0 for Windows and S-Plus 6 for Windows.

RESULTS

There were 49 of 314 (15.6%) and 45 of 296 (15.2%) apprentices who developed new sensitisation to LA allergens and symptoms at work, respectively, at the 32nd month of the training programme. Only 16 of 279 (5.7%) apprentices developed both new sensitisation to LA allergens and symptoms at work.

Predictors of incidence of sensitisation to LA allergens

At baseline (on average at 1.4 months after entry), 140 (44.6%) students had allergic symptoms. Based on univariable analysis results, this variable was the strongest predictor of the development of new sensitisation to LA allergens (table 1). SPT-based atopy or BR were both associated with the endpoints with comparable strength.

We started the multivariable questionnaire model with four clusters of symptoms from the PCA (respiratory symptoms at work, symptoms indicative for asthma, allergic symptoms and skin symptoms). After backward selection in the regression model, only symptoms indicative of asthma and allergic symptoms appeared as independent predictors of the development of LA sensitisation and, thus, remained in the questionnaire model (table 2, model 1).

All models showed a good internal validity: from the bootstrapping procedure, we obtained correction factors of

Table 5 Comparison of the accuracy of the selected cut-off points of the predicted probability of developing sensitisation to laboratory animal allergens produced by different models

	No of apprentices in the group (%) [*]	No of apprentices with sensitisation (n = 49) [†]		No of apprentices without sensitisation (n = 265) [‡]		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
		No (%)	No (%)	No (%)	No (%)				
Probability ≥ 0.10									
Model 1	195 (62.1)	44 (22.6)	151 (77.4)	89.8	43.0	22.6	95.8		
Model 2	176 (56.1)	41 (23.3)	135 (76.7)	83.7	49.1	23.3	94.2		
Model 3	187 (59.6)	44 (23.5)	143 (76.5)	89.8	46.0	23.5	96.1		
Model 4	159 (50.6)	41 (25.8)	118 (74.2)	83.7	55.5	25.8	94.8		
Probability ≥ 0.20									
Model 1	84 (26.8)	27 (32.1)	57 (67.9)	55.1	78.5	32.1	90.4		
Model 2	83 (26.4)	31 (37.3)	52 (62.7)	63.3	80.4	37.3	92.2		
Model 3	59 (18.8)	26 (44.1)	33 (55.9)	53.1	87.5	44.1	91.0		
Model 4	90 (28.7)	33 (36.7)	57 (63.3)	67.3	78.5	36.7	90.4		
Probability ≥ 0.30									
Model 1	84 (26.8)	27 (32.1)	57 (67.9)	55.1	78.5	32.1	90.4		
Model 2	56 (17.8)	24 (42.9)	32 (57.1)	49.0	87.9	42.9	90.3		
Model 3	59 (18.8)	26 (44.1)	33 (55.9)	53.1	87.5	44.1	91.0		
Model 4	48 (15.3)	24 (50.0)	24 (50.0)	49.0	90.9	50.0	90.6		

^{*}Proportion of all apprentices (n = 314).

[†]Proportion of apprentices with symptoms at work within the probability group.

[‡]Proportion of apprentices without symptoms at work within the probability group.

Model 1, questionnaire alone; model 2, questionnaire and SPT (skin-prick testing); model 3, questionnaire and BR (bronchial responsiveness to methacholine) testing; model 4, questionnaire, SPT and BR testing; NPV, negative predicted value; PPV, positive predictive value.

Table 6 Comparison of the accuracy of the selected cut-off points of the predicted probability of developing respiratory symptoms at work produced by different models

	No of apprentices in the group (%)*	No of apprentices with symptoms at work (n = 45)†	No of apprentices without symptoms at work (n = 251)‡	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
		No (%)	No (%)				
Probability ≥0.10							
Model 1	118 (39.9)	37 (31.4)	81 (68.6)	82.2	67.7	31.4	95.5
Model 2	133 (44.9)	38 (28.6)	95 (71.4)	84.4	62.2	28.6	95.7
Model 3	143 (48.3)	38 (26.6)	105 (73.4)	84.4	58.2	26.6	95.4
Model 4	124 (41.9)	37 (29.8)	87 (70.2)	82.2	65.3	29.8	95.3
Probability ≥0.20							
Model 1	118 (39.9)	37 (31.4)	81 (68.6)	82.2	67.7	31.4	95.5
Model 2	92 (31.1)	33 (35.9)	59 (64.1)	73.3	76.5	35.9	94.1
Model 3	86 (29.1)	32 (37.3)	54 (62.8)	71.1	78.5	37.3	93.8
Model 4	100 (33.8)	36 (36.0)	64 (64.0)	80.0	74.5	36.0	95.4
Probability ≥0.30							
Model 1	67 (22.6)	25 (37.3)	42 (62.7)	55.6	83.3	37.3	91.3
Model 2	67 (22.6)	25 (37.3)	42 (62.7)	55.6	83.3	37.3	91.3
Model 3	86 (29.1)	32 (37.3)	54 (62.8)	71.1	78.5	37.3	93.8
Model 4	57 (19.3)	23 (40.4)	34 (59.6)	51.1	86.5	40.4	90.8

*Proportion of all apprentices (n = 296).

†Proportion of apprentices with symptoms at work within the probability group.

‡Proportion of apprentices without symptoms at work within the probability group.

Model 1, questionnaire alone; model 2, questionnaire and SPT (skin-prick testing); model 3, questionnaire and BR (bronchial responsiveness to methacholine) testing; model 4, questionnaire, SPT and BR testing; NPV, negative predicted value; PPV, positive predictive value.

0.91 or higher. All of them also showed a good calibration (Hosmer-Lemeshow p value >0.10). Table 2 presents the corrected regression coefficients, intercepts and ROC areas, after multiplication by the correction factor from the bootstrapping procedure. The questionnaire (model 1) had a reasonable discrimination (ROC area was 0.731; 95% CI 0.663 to 0.799). The ROC area for SPT to common allergens or BR testing as a single test was equal: 0.69 (95% CI 0.60 to 0.77). The addition of SPT-based atopy to the questionnaire model (model 2) increased the ROC area to 0.754 (delta 0.023, 95% CI -0.014 to 0.062) and significantly increased the model χ^2 (deviance = 8.487, $df = 1$, $p = 0.004$). When BR testing was added to the questionnaire model (model 3), the ROC area increased by 0.046 (95% CI -0.004 to 0.098) and the model χ^2 significantly increased (deviance = 16.916, $df = 1$, $p < 0.001$). A complete combination of the questionnaire, SPT and BR testing (model 4) yielded a ROC area of 0.785 (delta 0.054; 95% CI -0.002 to 0.112) and significantly increased the model χ^2 (deviance = 22.505, $df = 2$, $p < 0.001$).

Predictors for incidence of symptoms at work

From the univariable analysis, allergic symptom was also the strongest predictor of the development of respiratory symptoms at work (table 3). Age and all symptom clusters except skin symptoms were included in the multivariable analysis. After backward selection in the regression model, only symptoms indicative of asthma and allergic symptoms remained in the final questionnaire model (table 4, model 1).

All models showed a good internal validity; correction factors from the bootstrapping procedure were 0.91 or higher. After multiplication by the correction factor, model 1 produced a good discrimination (corrected ROC area of 0.780 (95% CI 0.713 to 0.847)) and a good calibration (H-L test p value = 0.999). The ROC area of SPT to common allergens or BR testing alone was significantly lower than the questionnaire model (table 4, footnote). The addition of SPT to common allergens, BR testing, or both of them marginally increased the discriminative

ability of the questionnaire model (delta ROC area ranged between 0.001 and 0.010, 95% confidence interval -0.023 to 0.043). None of the additional tests significantly increased the model χ^2 of the questionnaire model (results not shown).

Application

To evaluate the clinical relevance of the prediction models, several probability thresholds were set to classify the apprentices into low and high probability group. The accuracy of the thresholds when applied in different models for predicting LA sensitisation (table 5) and respiratory symptoms at work (table 6) was examined.

For example, if we used the questionnaire model (model 1) in the prediction of LA sensitisation, there were 84 (26.8%) apprentices with a probability of 0.30 or higher. This threshold had a positive predictive value (PPV) of 32.1% (27/84) and a negative predictive value (NPV) of 90.4% (208/230). At higher probability thresholds, the specificity and PPV of the more complex models were higher than the questionnaire model alone. Table 6 illustrates that the questionnaire model for symptoms at work showed a comparable diagnostic accuracy to that from the other models across all range of probabilities.

To apply the models in daily practice, we transformed the regression questionnaire models for both endpoints into easy-to-use score charts (see appendices 2 and 3 online). As an example of how to use this chart, a student who had symptoms indicative of asthma would have a sum score of 1 (1+0). This corresponded to a 7% probability of developing sensitisation to LA allergens and a 4% probability of developing respiratory symptoms at work. If a sum score greater than 1 is chosen as the cut-off, then this student would be included in the high probability group.

DISCUSSION

We assessed the capacity of various tools—questionnaire, skin-prick testing and bronchial responsiveness testing—alone or in combination, to correctly predict the occurrence of specific

Original article

sensitisation to LA allergens and respiratory symptoms at work using data from a prospective study of apprentices.^{9 10 22} We found that the questionnaire models had a good accuracy in predicting incidence of LA sensitisation as well as symptoms at work. Additional information from skin-prick testing to common allergens and/or bronchial challenge to methacholine seemed to have limited predictive value in a relatively healthy population.

The questionnaire model with and without SPT and/or BR testing showed a good calibration, which meant that the predicted probabilities were in a good agreement with the observed frequency of the endpoint. It is also important to note that the questionnaire model had a comparable ability in discriminating sensitised from non-sensitised apprentices to that from SPT or BR testing alone. For symptoms at work, the discriminative ability of the questionnaire model was significantly higher than SPT as well as BR testing alone. The questionnaire model had a ROC area of 0.73, which meant that we could correctly assign a higher probability to a sensitised apprentice in 73% of pairs of apprentices in which one apprentice is sensitised to LA allergens and one is not.²³ This figure was equal to what was demonstrated by the questionnaire model for sensitisation to LA allergens that was developed in a population of Dutch workers (ROC area of 0.73; 95% CI 0.65 to 0.82). In the quoted study,⁸ the addition of total IgE and sensitisation to common animal allergens to the final questionnaire model substantially improved the ROC area to 0.81 (95% CI 0.73 to 0.89).

In our study, additional information from SPT and BR testing yielded marginal to moderate additional information on the prediction of the development of occupational allergy caused by exposure to LA allergens. This might be explained by the fact that the questionnaire includes items related to both the atopic history and bronchial responsiveness. Another possible explanation was that the models were derived from a relatively healthy population. Modelling in populations with disease manifestation—for example, in the clinical setting, may yield different models with different strengths of association between predictors and the endpoint. Therefore, it might be unnecessary to

confirm atopy and bronchial responsiveness with objective testing in a general population of healthy subjects; the questionnaire items representing satisfactory surrogates. However, depending upon the availability of resources, the healthcare and social system and legal context where the model will be applied, one may prefer to use a complete questionnaire-based model alone or a questionnaire in combination with additional test-based models.

As for the outcomes, we used a loose definition of respiratory symptom at work—developed any of nasoconjunctival or chest symptoms at work—due to power limitation. There were limited numbers of incident cases available for developing the models for a more complex outcome such as combination of symptoms and specific sensitisation to LA allergens. We did the PCA before modelling and after the models were developed, we did the bootstrapping procedure to check whether the models were reasonably valid. This procedure has been shown to be superior to split-sample or cross-validation methods.²⁴ It turned out that all models had a good internal validity, the correction factors for all models were 0.91 or above; the closer the correction factor is to 1, the less optimistic (provides too high or too low prediction). Nevertheless, an external validation of the models in a new population is necessary to confirm the performance of the models and their generalisability into other groups of animal health apprentices.^{25 26}

To the best of our knowledge, these are the first prognostic models for LA sensitisation and respiratory symptoms at work developed in an animal health apprentices setting. Such models enable the quantification of the probability of developing sensitisation to LA allergens or respiratory symptoms at work for individuals and, thus, they may support career counselling and health monitoring programmes among individuals with a high probability. At the individual level, this quantification will generate an awareness of relevant occupational allergic diseases. Apprentices with a high probability of developing sensitisation would be expected to be more vigilant—for example, use protective equipments to reduce the exposure to LA allergens. However, when applied in a working population, one should be aware of the potential negative implication of the model—for example, the predictive value is relatively low to be used in pre-employment screenings by employers or if the intervention was a change of job without signs of disease.

Another crucial issue in model application is the choice of the cut-off point above which workers will be assigned to a high probability group of developing sensitisation or respiratory symptoms at work. The choice must be based on an acceptable proportion of missed cases and of unnecessary referrals for rigorous medical tests for occupational allergies (that is, bronchial methacholine challenge test) in every worker. As shown in tables 5 and 6, a higher cut-off leads to fewer workers in the high-risk group; the specificity is higher but at the cost of lower sensitivity and vice versa. For example, when we apply the questionnaire model for sensitisation to LA allergens, we will save the expenses of unnecessary tests in almost 75% of the workers by not referring workers with a probability lower than 0.30. This cut-off point also has a high negative predictive value of 90% (208/230), which means that more than 90% of those who are not referred would indeed have a negative skin reactivity to LA allergens (if tested). However, 22 of all 49 cases will be missed. Given the nature of the disease this is not a major issue; our endpoints are not cancerous or life-threatening disease where misclassification would have dramatic impact on the prognosis. Nevertheless, if the aim is to detect as many cases as possible, a lower cut-off value might be of interest.

Main messages

- We developed the models for quantification of the probability of developing sensitisation to laboratory animal allergens or respiratory symptoms at work in an animal health apprentices setting.
- Information from questionnaire at entry can give an accurate prediction of the incidence of occupational sensitisation and symptoms owing to exposure to laboratory animal allergens at 32 months of the training programme. The added predictive value of skin-prick testing to common allergens and/or bronchial challenge to methacholine seemed to be limited in a relatively healthy population.

Policy implication

Application of the prognostic models for occurrence of occupational sensitisation and symptoms in an animal health apprentices setting may support career counselling and health monitoring programmes among individuals with a high probability.

In conclusion, we developed prognostic models to predict the occurrence of sensitisation to LA allergens and symptoms at work in animal health apprentices. The questionnaire model alone is an easy tool that can give an accurate prediction of the incidence of occupational sensitisation and symptoms. Application of prognostic models in surveillance will enable close monitoring of the identified high-risk group for preventive purposes. Yet, the predictive capacity of the models has to be validated in other populations exposed to laboratory animal allergens before they can be used with confidence. Our models might be applicable in a new apprentice setting with exposure to other high-molecular-weight allergens, but might not be generalised to settings with exposure to low-molecular-weight agents or irritant caused by a different underlying mechanism. Therefore, it would also be interesting to validate our models in populations exposed to other high molecular weight agents, as well as to develop similar models for low-molecular-weight agents or irritant causing occupational asthma in further study.

Acknowledgements: This study was supported by the Canadian Institutes of Health Research (grant MOP-53118), the Medical Research Council of Canada (grant MT-12256) and the Institut de recherche Robert-Sauvé en santé et en sécurité du travail (grant 099-164). ES is a research fellow with a grant from the Center for Asthma in the Workplace-Canadian Institutes of Health Research.

Competing interests: None.

Ethics approval: Informed consent was obtained from each subject and the study was performed after approval from the Sacre-Coeur Hospital's ethics committee in accordance with Canadian ethical rules.

REFERENCES

1. **Blanc P**, Toren K. How much asthma can be attributed to occupational factors? *Am J Med* 1999;**107**:580–7.
2. **Kogevinas M**, Anto JM, Sunyer J, *et al.* Occupational asthma in Europe and other industrialised areas: a population-based study. European Community Respiratory Health Survey Study Group. *Lancet* 1999;**353**:1750–4.
3. **Chan-Yeung M**, Malo JL, Tarlo SM, *et al.* Proceedings of the first Jack Pepys Occupational Asthma Symposium. *Am J Respir Crit Care Med* 2003;**167**:450–71.
4. **Kogevinas M**, Zock J, Jarvis D, *et al.* Exposure to substances in the workplace and new-onset asthma: an international prospective population-based study (ECRHS-II). *Lancet* 2007;**370**:336–41.
5. **Meyer JD**, Holt DL, Cherry NM, *et al.* SWORD '98: Surveillance of work-related and occupational respiratory disease in the UK. *Occup Med* 1999;**49**:485–9.
6. **Blanc PD**, Eisner MD, Balmes JR, *et al.* Exposure to vapors, gas, dust, or fumes: assessment by a single survey item compared to a detailed exposure battery and a job exposure matrix. *Am J Ind Med* 2005;**48**:110–7.
7. **Bernstein DI**, Campo P, Baur X. Clinical assessment and management of occupational asthma. In: Bernstein IL, Chan-Yeung M, Malo J, *et al.*, eds. *Asthma in the workplace*. 3rd ed. New York: Taylor & Francis, 2006:172.
8. **Meijer E**, Grobbee DE, Heederik D. A strategy for health surveillance in laboratory animal workers exposed to high molecular weight allergens. *Occup Environ Med* 2004;**61**:831–7.
9. **Gautrin D**, Infante-Rivard C, Dao T, *et al.* Specific IgE-dependent sensitization, atopy and bronchial hyperresponsiveness in apprentices starting exposure to protein-derived agents. *Am J Respir Crit Care Med* 1997;**155**:1841–7.
10. **Gautrin D**, Ghezzo H, Infante-Rivard C, *et al.* Incidence and determinants of IgE-mediated sensitization in apprentices: a prospective study. *Am J Respir Crit Care Med* 2000;**162**:1222–8.
11. **Gautrin D**, Ghezzo H, Infante-Rivard C, *et al.* Natural history of sensitization, symptoms and diseases in apprentices exposed to laboratory animals. *Eur Respir J* 2001;**17**:904–8.
12. **Rodier F**, Gautrin D, Ghezzo H, *et al.* Incidence of occupational rhinoconjunctivitis and risk factors in animal-health apprentices. *J Allergy Clin Immunol* 2003;**112**:1105–11.
13. **Burney PG**, Laitinen LA, Perdzet S, *et al.* Validity and repeatability of the IUATLD (1984) Bronchial Symptoms Questionnaire: an international comparison. *Eur Respir J* 1989;**2**:940–5.
14. **Burney P**, Luczynska C, Chinn S, *et al.* The European Community Respiratory Health Survey. *Eur Respir J* 1994;**7**:954–60.
15. **Bollen K**. *Structural equations with latent variables*. New York: John Wiley & Sons, 1989.
16. **Sterk PJ**, Fabbri LM, Quanjer PH, *et al.* Airway responsiveness. Standardized challenge testing with pharmacological, physical and sensitizing stimuli in adults. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;**16**:53–83.
17. **Troyanov S**, Malo JL, Cartier A, *et al.* Frequency and determinants of exaggerated bronchoconstriction during shortened methacholine challenge tests in epidemiological and clinical set-ups. *Eur Respir J* 2000;**16**:9–14.
18. **Harrell FE Jr**, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361–87.
19. **Hosmer D**, Lemeshow S. *Applied logistic regression*. New York: John Wiley and Sons, 1989.
20. **Hanley JA**, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
21. **Van Houwelingen JC**, Le Cessie S. Predictive value of statistical models. *Stat Med* 1990;**9**:1303–25.
22. **Gautrin D**, Infante-Rivard C, Ghezzo H, *et al.* Incidence and host determinants of probable occupational asthma in apprentices exposed to laboratory animals. *Am J Respir Crit Care Med* 2001;**163**:899–904.
23. **Hanley J**, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
24. **Peduzzi P**, Concato J, Kemper E, *et al.* A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;**49**:1373–9.
25. **Bleeker S**, Moll H, Steyerberg E, *et al.* External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;**56**:826–32.
26. **Justice A**, Covinsky K, Berlin J. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**:515–24.