

The Dutch diagnostic model for laboratory animal allergen sensitization was generalizable in Canadian apprentices

Eva Suarhana^{a,b,c}, Evert Meijer^b, Dick Heederik^b, Heberto Ghezso^a,
Jean-Luc Malo^a, Denyse Gautrin^{a,*}

^aDepartment of Chest Medicine, Hôpital du Sacré-Coeur de Montréal, Montreal, Canada

^bIRAS (Institute for Risk Assessment Sciences), Environmental Epidemiology Division, Utrecht University, Utrecht, The Netherlands

^cCommunity Medicine Department, Faculty of Medicine, University of Indonesia, Jl. Pegangsaan Timur, Jakarta Pusat, Indonesia

Accepted 3 June 2008

Abstract

Objective: To assess the transportability of an existing diagnostic questionnaire model for the sensitization to laboratory animal (LA) allergens.

Study Design and Setting: The model was externally validated in 414 Canadian animal health apprentices. Several approaches were used: (1) no adjustment; (2) recalibration of the intercept of the model; (3) re-estimation of the intercept and the regression coefficients of predictors; and (4) model revision, by excluding the existing predictor(s) and/or including new predictor(s). The bootstrapping procedure was done following the third and fourth methods. The calibration was assessed graphically and with the Hosmer–Lemeshow (HL) test. Discriminative properties were determined by the area under the receiver operating characteristic curve (ROC area).

Results: When applied without adjustment, the model's discriminative ability was adequate (ROC area was 0.74 vs. the original ROC area of 0.76); the calibration was poor (HL test $P < 0.001$). The other methods yielded models with good calibration ($P > 0.10$) and reasonable discrimination (ROC area ranged between 0.73 and 0.75). The refitted and revised model showed a good internal validity (correction factor from the bootstrapping procedure was more than 0.90).

Conclusion: Once updated, the diagnostic model is valid and can be applied with reasonable performance in an animal health apprentice setting. © 2009 Elsevier Inc. All rights reserved.

Keywords: Diagnostic model; High molecular weight allergens; Occupational sensitization; Questionnaire; Screening; Validity

1. Introduction

At baseline, in a cohort study of Canadian apprentices beginning animal health and veterinary medicine career programs, the skin reactivity to work-specific laboratory animal (LA) proteins was 13.8%. The study suggested that sensitization could ensue even if a very brief specific occupational exposure had occurred [1].

To demonstrate work-related sensitization, standardized allergen preparations are required, but these are costly and may not always be available in an occupational health practice. Therefore, Meijer et al. developed a diagnostic questionnaire model for specific sensitization to LA allergens

among Dutch laboratory workers exposed to rats, mice, and other rodents [2]. They demonstrated that a diagnostic model based on questionnaire items could be accurately used to predict the presence of workers at high or low risk of being sensitized without having to perform the more advanced reference test. Furthermore, they demonstrated that additional information from skin-prick test (SPT) responses to animal and non-animal common allergens improved the diagnostic performance. In a more recent study, it was shown that it is possible to develop a generic model for sensitization to occupational high molecular weight (HMW) allergens with some modifications for specific work environments [3]. The use of these diagnostic models can increase the efficiency of health surveillance, by allowing an occupational physician to predict for an individual worker the probability of being sensitized to LA allergens. On the basis of this prediction, a decision can be made to conduct additional specific tests to diagnose occupational allergies among workers with a high probability [4].

* Corresponding author. Department of Chest Medicine, Hôpital du Sacré-Coeur de Montréal, 5400 West Gouin Boulevard, Montreal H4J 1C5, Canada. Tel.: +514-3382222 ext. 2568; fax: 514-3383123
E-mail address: d.gautrin@umontreal.ca (D. Gautrin).

The utility of predictive models depends on how well they perform when applied to a population, which may be different from, but related to the individuals used to develop a model. In general, prediction models show a lower performance in populations other than where the model was derived [5]. For this reason, external validation of the model is necessary to address the accuracy of a model in different, but related, workers [5,6]. Canadian apprentices at 1.4 months after entry showed a different distribution of personal and exposure characteristics compared with Dutch workers. However, both Dutch workers and Canadian apprentices are exposed to the same occupational allergens, and thus, are at risk of developing the same occupational allergies [1,2].

Therefore, the objective of this study was to predict at an early phase the likelihood of sensitization to LA allergens in Canadian animal health technology apprentices, who had been exposed for about 1.4 months. We, thus, used the known predictors from the existing questionnaire model and externally validated this model in these trainees [5–7]. The second objective was to derive a diagnostic questionnaire model for the sensitization to LA allergens from the Canadian apprentices, and evaluate if whether or not the inclusion of predictors that were available in the Canadian setting could improve the performance of the existing model.

2. Methods

2.1. Populations

The existing diagnostic questionnaire model for the sensitization to LA allergens was derived from the first period of a cohort study investigating exposure–response relationships among 472 Dutch LA workers [2,8]. Questionnaire items, exposure determinants, IgE serology, SPT, and lung function tests were collected from all workers. The self-administered questionnaire was based on a Dutch version of an internationally accepted respiratory questionnaire [9]. The questionnaire included questions on age, gender, respiratory problems, personal and family history of allergic symptoms, and smoking history. Additional questions were asked about employment history, work duration with LAs, and allergic symptoms owing to working with LAs.

The apprentice study was conducted between 1993 and 1998, investigating the natural history of occupational asthma among Canadian apprentices exposed to high-molecular weight allergens [1]. At baseline, 417 Canadian animal health apprentices from four institutions offering a training program in animal health technology in Quebec, Canada, participated in the study. The apprentices answered questionnaires and were subjected to SPTs, lung function, and bronchial nonspecific provocation tests on beginning their 3-year apprenticeship program. The questionnaire was derived from the standardized questionnaire

of the International Union against Tuberculosis and Lung Disease (IUATLD) and was administered by a trained nurse [10].

The existing diagnostic models were externally validated in Canadian animal health apprentices at 1.4 months of their apprenticeship. An analysis on an item-by-item basis was done to identify questionnaire items from both studies, which were comparable at an acceptable level for this analysis. Informed consent was obtained from each subject, and both studies were performed according to Dutch and Canadian ethical rules.

2.2. Potential predictors

A history of asthma was defined as a positive answer to the following question; “Have you ever had an asthma attack in the last 12 months?” Symptoms of asthma were considered to be present if there were at least two relevant symptoms—wheezing, chest tightness, cough, and dyspnea. Allergic symptoms during work were considered present if the worker experienced respiratory symptoms (chest tightness, cough, or wheeze) or nasoconjunctival symptoms (runny nose or sneezing, running or itching eyes), and/or skin irritation when in contact with LAs. Workers were considered to have allergic symptoms during the past 12 months if they reported at least one eye, nasal, or respiratory problem when exposed to common allergens, such as, house dust, domestic animals, food, or pollen. Personal atopic history was defined as having a history of eczema, urticaria, or hay fever. Symptoms suggestive for bronchial hyperresponsiveness (BHR) were considered present if subjects experienced respiratory problems induced by exercise, strenuous work, very cold air, heavy smell, smoke, or dust.

2.3. Reference standard

SPT is useful to detect specific IgE responses to HMW allergens [11]. In both populations, six occupational allergens (rat urine, mouse urine, rat fur, mouse fur, guinea pig fur, and rabbit fur) and positive and negative controls were used for SPT. A wheal diameter of 3 mm or more was regarded as a positive response, after subtraction of any response to the negative control [8]. Sensitization to LA allergens was defined as a positive SPT response to any occupational allergens.

2.4. Data analysis

Of 417 eligible individuals, three (0.7%) apprentices with no SPT result were excluded, leaving $n = 414$ (99.3%) with complete data.

2.4.1. External validation of the existing diagnostics model

From the original Dutch logistic regression model, the individual probability of having a positive SPT response

to LA allergens can be estimated using the following formula [2]:

$$P(\text{sensitization}) = 1 / (1 + \exp(-(-1.82 + \text{history of asthma attacks} * 0.98 + \text{history of allergic symptoms} * 0.87 + \text{allergic symptoms during work} * 1.03 + \text{exposed to rats} \geq 20 \text{ hr/wk} * 0.79 + \text{male gender} * 0.46)))$$

Different statistical approaches have been introduced to externally validate a model in a new population [6]. We calculated individual probabilities using the equation from the existing model without any adjustment (no update, method 1). In the second approach, the same regression coefficients were used, but the intercepts were re-estimated in the validation population (recalibrate the intercept, method 2). In the third approach, the regression coefficients and the intercepts were re-estimated (refit the model, method 3). Finally, an evaluation whether exclusion of the existing predictor(s) and or inclusion of new predictor(s) from the Canadian setting could improve the performance of the Dutch model was done (model revision, method 4). Method 4 was carried out using the logistic regression analysis, where predictors from the Dutch model were frozen in the first block, and subsequently, predictors with univariable $P < 0.5$ from Canadian setting were added in the second block. Backward stepwise selection was applied with an inclusion criterion of $P < 0.15$ in the second block, and the overall χ^2 of the block was evaluated. The diagnostic accuracy of the diagnostic model was quantified using calibration and discrimination measures. The agreement between the predicted probabilities and the observed frequencies for sensitization (calibration) was evaluated graphically and with the Hosmer–Lemeshow (HL) test (where $P = 0.05$ and higher reflects good agreement) [12]. The model's ability to discriminate the sensitized from not-sensitized apprentices (discrimination) was determined with the area under the receiver operating characteristic curve (ROC area). The ROC area illustrates the relation between the false-positive rate (1-specificity) and the true-positive rate (sensitivity). The ROC area can range from 0.5 (no discrimination) to 1.0 (perfect discrimination) [13]. Methods 3 and 4 were followed by bootstrapping to assess the internal validity of the model [14]. Random bootstrap samples were drawn with replacement from the population consisting of all Canadian apprentices (100 replications). This bootstrapping procedure produced a corrected model's ROC area and a shrinkage factor. The regression coefficients of the predictors in the model were multiplied by this shrinkage factor to prevent the model from yielding optimistic predictions when applied in future (new) workers. Corrected ROC areas were compared for all methods.

2.5.2. Development and internal validation of the Canadian diagnostic model

The guidelines on the development of the prediction model were described by Harrell [15]. Briefly, questionnaire predictors with $P < 0.5$ in a univariable analysis were entered into a backward stepwise multivariable logistic regression procedure. The Akaike criterion ($P < 0.15$ for inclusion) was used to select a final questionnaire model with the strongest predictors for occupational sensitization. The diagnostic accuracy and internal validity of the final model were assessed as described earlier.

2.5.3. Model application

For screening purposes, a cutoff point of the predicted probabilities produced by different validation methods was considered to divide the population into apprentices with a low and a high sensitization probability. The sensitivity, specificity, positive and negative predictive values of the selected cutoff point were calculated. To facilitate the application of the diagnostic model in practice, the validated diagnostic model with the best diagnostic performance was converted to a nomogram [16]. Nomograms can be used to manually obtain predicted values from a regression model and are convenient tools in clinical practice. The nomogram has a reference line for reading scoring points (default range: 0–100) for each predictor in the model. Once the reader manually totals the points, the predicted probabilities can be read at the bottom.

All analyses were run using version 14.0 for Windows (Statistical Package for Social Sciences, Chicago, IL, USA) and S-Plus version 6.0 (Insightful Corp., Seattle, WA, USA) using the Hmisc and Design libraries [16,17].

3. Results

The prevalence of sensitization to LA allergens among the Dutch LA workers and Canadian animal health apprentices were 27.3% and 13.8%, respectively (Table 1). The Dutch workers were older, exposed longer to LA allergens, and most of them were males. Contrarily, most of the Canadian apprentices were female, and they had just started their vocational training in animal health technology. Twelve point that 5% of the Dutch worked 20 hr/wk or more with rats, whereas all Canadian apprentices were exposed less than 1 hr/wk to rats in their first months of apprenticeship. The Canadian apprentices somehow had higher percentages of allergic symptoms during the past 12 months, personal atopic history, and positive SPT response to non-animal common allergens, whereas the percentages of asthma attacks and symptoms suggestive of BHR were comparable with the Dutch workers. Interestingly, both populations showed comparable strength of associations (reflected by the odds ratio) between predictors from the earlier diagnostic model (gender, allergic symptoms during the past 12 months, allergic symptoms during work, and asthma attacks) and sensitization to LA allergen based on SPT (Table 2).

Table 1
Characteristics of samples in the Dutch and Canadian study populations

| Characteristics | Dutch workers (n = 472) | Canadian apprentices (n = 414) |
|--|----------------------------|--------------------------------------|
| Age (yr) ^a | 34.4 (9.4) | 18.3 (3.3) |
| Exposure duration (yr) ^a | 8.3 (8.9) | |
| Time exposed to rats (hr/wk) ^a | 8.5 (10.4) | 0.13 (0.15) |
| Work with rats >20 hr/wk | 59 (12.5) | 0 (0.0) |
| Male gender | 293 (62.1) | 57 (13.8) |
| Allergic symptoms during the past 12 months ^b | 131 (27.8) | 191 (41.6) |
| Allergic symptoms during work ^c | 152 (32.2) | 72 (17.4) |
| Asthma attacks | 66 (14.0) | 48 (11.6) |
| Symptoms of asthma ^d | 45 (9.5) | 75 (18.1) |
| Symptoms suggestive of BHR ^e | 192 (40.8) | 192 (46.4) |
| Personal atopic history ^f | 110 (23.3) | 147 (35.5) |
| Positive SPT to common animal allergens (to cat and or dog) | 121 (25.6) | 74 (17.9) |
| Positive SPT to common non-animal allergens (to house dust mite and or pollens) | 177 (37.5) | 205 (49.5) |
| <i>Outcome</i> | | |
| Positive SPT to LA allergens (to rat/mouse urine and/or to rabbit/guinea pig dander) | 129 (27.3) | 57 (13.8) |

Data presented as n (%) unless otherwise stated.

Abbreviations: BHR, bronchial hyperresponsiveness; SPT, skin-prick test.

^a Mean (standard deviation).

^b Experience at least one eye, nasal, or respiratory problem when exposed to common allergens, such as house dust, domestic animals, food, or pollen.

^c Experience chest tightness, running nose or sneezing, running or itching eyes, and/or itching skin on contact with laboratory animals.

^d Present if there were at least two relevant symptoms (wheezing, chest tightness, cough, and dyspnea).

^e Present if experience respiratory problems induced by exercise, strenuous work, very cold air, heavy smell, smoke, or dust.

^f Present in the case history of eczema, urticaria, or hay fever.

3.1. Diagnostic questionnaire model derived from the Canadian apprentices

We fitted a full model comprising age, gender, and allergic symptoms during the past 12 months, allergic symptoms during work, asthma attacks, symptoms of asthma, symptoms suggestive for BHR, and personal atopic history. The selected predictors in the final Canadian questionnaire model were allergic symptoms during the past 12 months, allergic symptoms during work, symptoms of asthma, and personal atopic history (Table 3, last column). The intercept and regression coefficients were multiplied by a shrinkage factor of 0.88 from the bootstrapping procedure, and the corrected ROC area was 0.74. Calibration of the model was good (HL test: $P = 0.298$).

3.2. External validation of the Dutch diagnostic questionnaire model

Recalibration of the model's intercept (method 2) gave a new intercept of -2.855 (Table 3, column 3). The re-estimated intercept and regression coefficients from method

3 presented in the fourth column had been multiplied by 0.95 (shrinkage factor from the bootstrapping procedure). Except for working hours with rats, which could not be re-estimated (all apprentices were exposed <20 hr/wk), the third method yielded comparable regression coefficients of the predictors as compared with values in the original model. In the model revision method, working hours with rats was excluded from the model, whereas symptoms of asthma and personal atopic history from the Canadian model were selected and added into the model (model χ^2 deviance = 11.8, df = 2). The internal validity of the revised model was good; the correction factor from the bootstrapping procedure was 0.94.

The ROC area produced by different validation approaches ranged between 0.73 and 0.75. Calibration plots in Fig. 1 demonstrate the graphical assessment of the agreement between the predicted probabilities and the observed frequencies. The calibration plot for the non-updated model obviously deviated from the ideal line (Fig. 1a) with a very significant HL test ($P < 0.001$). Recalibration of the intercept alone improved the calibration substantially (Fig. 1b); the HL test P value was 0.999, which reflects a satisfactory calibration. The calibration of the model produced by the third and fourth validation methods was good (the HL test $P > 0.1$).

3.3. Model application

To generate case-specific advice and support decision making by health professionals, several cutoff points of the predicted probability were selected, and their diagnostic accuracy was compared across the updated models (Table 4). A higher cutoff leads to a higher specificity and positive predictive value (PPV), but at the cost of lower sensitivity, and vice versa. A cutoff value of the predicted probability of 0.15 or higher was chosen as an example (prior probability, i.e., the prevalence of sensitization, was 13.8%). If we use the revised model, 135 (32.6%) apprentices will be classified in the high-probability group; 40 of 57 sensitized apprentices will be captured (70% sensitivity); and 262 of 279 apprentices in the low-probability group would have positive skin reactivity to LA allergens if tested (93.9% negative predictive value, NPV). Overall, the revised model yielded the best diagnostic properties across different cutoff points. Therefore, the revised model was converted to a nomogram (Fig. 2). The total scoring points corresponded to the predicted probabilities of being sensitized, which was presented at the bottom of the nomogram. As an example on the use of the nomogram, a female apprentice, with a history of atopy and asthma attacks, has a total number of points of 145 (0 + 0 + 100 + 45 + 0 + 0), which corresponds to a predicted probability of sensitization of 0.15.

4. Discussion

We aimed at assessing the transportability of an earlier obtained diagnostic questionnaire model for sensitization

Table 2
Strength of association between the predictors and sensitization to laboratory animal allergens in the Dutch and Canadian study populations

| Predictors | Dutch workers | | | Canadian apprentices | | |
|---|-------------------------|-----------------------------|--------------------------|------------------------|-----------------------------|--------------------------|
| | Sensitized (n = 129) | Not sensitized (n = 343) | OR (95% CI) ^a | Sensitized (n = 57) | Not-sensitized (n = 357) | OR (95% CI) ^a |
| Age (yr) ^b | 34.0 (9.1) | 34.6 (9.5) | 1.00 (1.00–1.01) | 18.8 (5.0) | 18.2 (2.9) | 1.05 (0.98–1.14) |
| Exposure duration (years) ^b | 8.4 (8.9) | 8.2 (9.0) | 1.00 (1.00–1.02) | | | |
| Time exposed to rats (hr/wk) ^b | 10.2 (11.5) | 7.9 (9.9) | 1.02 (1.00–1.04) | | | |
| Work with rats > 20 hr/wk | 23 (17.8) | 36 (10.5) | 1.9 (1.1–3.3) | | | |
| Male gender | 88 (68.2) | 205 (59.8) | 1.4 (0.9–2.2) | 10 (17.5) | 47 (13.2) | 1.4 (0.7–3.0) |
| Allergic symptoms during the past 12 months | 71 (55.9) | 60 (17.5) | 6.0 (3.8–9.4) | 44 (77.2) | 147 (41.2) | 4.8 (2.5–9.3) |
| Allergic symptoms during work | 79 (61.2) | 73 (21.3) | 5.8 (3.8–9.1) | 25 (43.9) | 47 (13.2) | 5.2 (2.8–9.5) |
| Asthma attacks | 38 (29.5) | 28 (8.2) | 4.7 (2.7–8.1) | 17 (29.8) | 31 (8.7) | 4.5 (2.3–8.8) |
| Symptoms of asthma | 18 (14.0) | 27 (7.9) | 1.9 (1.0–3.6) | 23 (40.4) | 52 (14.6) | 4.0 (2.2–7.3) |
| Symptoms suggestive of BHR | 62 (48.4) | 130 (37.9) | 1.5 (1.0–2.3) | 38 (66.7) | 154 (43.1) | 2.6 (1.5–4.8) |
| Personal atopic history | 25 (19.4) | 85 (24.8) | 0.7 (0.4–1.2) | 39 (68.4) | 108 (30.3) | 5.0 (2.7–9.1) |
| Positive SPT to common animal allergens | 77 (59.7) | 44 (12.8) | 10.1 (6.3–16.2) | 33 (57.9) | 41 (11.5) | 10.6 (5.7–19.7) |
| Positive SPT to common non-animal allergens | 89 (69.0) | 88 (25.7) | 6.4 (4.1–10.1) | 48 (84.2) | 157 (44.0) | 6.8 (3.2–14.3) |

Data presented as n (%) unless otherwise stated.

^a Odds ratio (95% confidence interval). ^b Data presented as mean (standard deviation).

to LA allergens derived from LA workers population to the animal health apprentices' setting. Our analyses suggest that with local adjustment, the diagnostic model is externally valid, and has a reasonable performance in predicting the presence or absence of sensitization to LA allergens in the Canadian animal health apprentices [5,7].

4.1. Issues on modeling and external validation

It is generally acknowledged that regression models used in diagnostic prediction research perform better on data from which the model is derived than on new data the same model is applied, especially in small data sets. Therefore, external validation is necessary to evaluate whether a model is applicable in another population. The calibration plot in Figure 1a clearly demonstrates that,

when the Dutch model is applied without any adjustments to a Canadian population, it systematically overestimates the sensitization probability [15]. The most likely explanation is the considerable difference in exposure characteristics between the two populations. The proxy variable for exposure from the Dutch model could not be evaluated properly, because it used a crude classification with a cut-point of 20-hr/wk exposure, whereas all the Canadian apprentices worked with rats less than 20 hr/wk during their apprenticeship. The ROC area produced by this method was 0.74, which was a little lower than the ROC area in the derivation set (0.76; 95% confidence interval [CI]: 0.71–0.82).

Figure 1b clearly illustrates how the recalibration of the intercept improves the existing model. The calibration line was closer to the ideal diagonal line and the HL test yielded

Table 3
The Dutch questionnaire model when validated in Canadian apprentices across different validation methods, compared with the Canadian questionnaire model

| | Validation of the Dutch model | | | | Canadian model |
|---|-------------------------------|---|-------------------------------|------------------------------|----------------|
| | No update (method 1) | Recalibrate the intercept (method 2) | Refit the model (method 3) | Model revision (method 4) | |
| | β^a | β^a | β^b | β^b | β^b |
| Intercept | -1.82 | -2.86 | -2.91 | -3.19 | -3.11 |
| Work with rats > 20 hr/wk | 0.79 | 0.79 | ^c | | |
| Male gender | 0.46 | 0.46 | 0.46 | 0.47 | |
| Allergic symptoms during the past 12 months | 0.87 | 0.87 | 1.0 | 0.62 | 0.56 |
| Allergic symptoms during work | 1.03 | 1.03 | 0.93 | 0.69 | 0.68 |
| Asthma attacks | 0.98 | 0.98 | 0.96 | 0.45 | |
| Symptoms of asthma | | | | 0.66 | 0.84 |
| Personal atopic history | | | | 1.01 | 0.96 |
| ROC area | 0.74 | 0.74 | 0.73 | 0.75 | 0.74 |
| Goodness of fit test (P value) | <0.001 | 0.999 | 0.999 | 0.602 | 0.298 |

Abbreviation: ROC, receiver operating characteristic curve.

^a Regression coefficient.

^b Regression coefficient after multiplication by correction factor from bootstrapping procedure.

^c Not applicable.

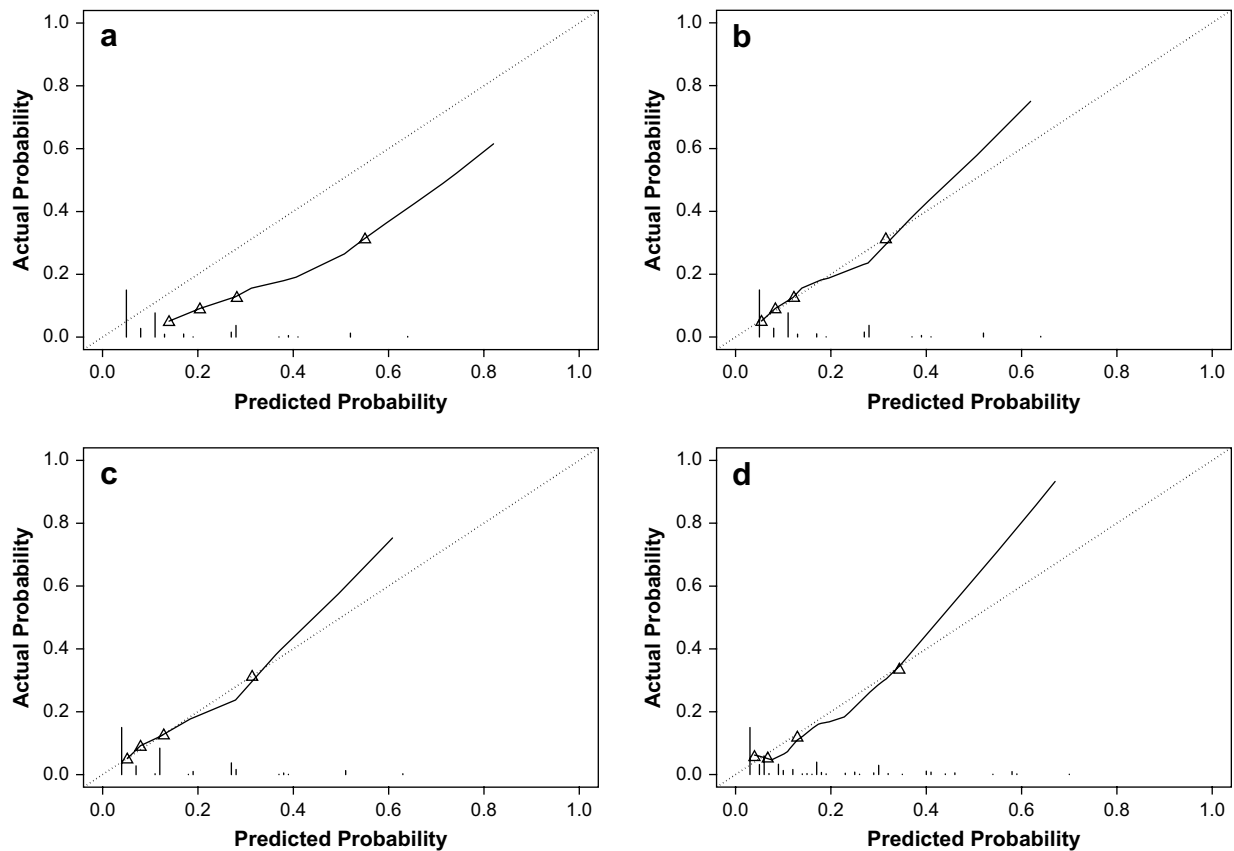


Fig. 1. Calibration plots of the Dutch questionnaire model when: applied without adjustment (a); the intercept was recalibrated (b); the model was refitted (c); and the model was revised; and (d) in Canadian apprentices. The solid line is a smoothed curve that represents a non-parametric estimate of the relation between the predicted probability and the observed sensitization rate. Ideally, this line fits the dotted line that represents perfect calibration. Triangles indicate the observed sensitization rate per equal-size quartiles of predicted probability. Distribution of the predicted probabilities is indicated with vertical lines at the bottom.

$P > 0.10$ for the recalibrated model, indicating no significant difference between the predicted probabilities and the observed frequencies of sensitization to LA allergens. Improvement of the performance of a model with an intercept adjustment indicated that there might be other important predictor(s) in the validation population not captured by the model. In accordance, two predictors were identified in the Canadian model, which were not captured in the earlier obtained model in the Dutch population. Symptoms of asthma and personal atopic history appeared to be strong predictors in the Canadian, but not in Dutch, population. Revision of the Dutch model by excluding the exposure variable and including the important predictors from the Canadian apprentice settings significantly improved the model χ^2 . However, the ROC area of the revised model was only slightly higher compared with the ROC area obtained by the simpler methods. This could be explained by changes in regression coefficients after the inclusion of the new predictors.

As expected, re-estimating the intercept and the regression coefficients of predictors yielded very similar estimations when compared with the original values. This was not surprising, because the predictors were defined in the same way for both populations, and the strength of

association between predictors in the model and the outcome were comparable. Re-estimation of the regression coefficient is recommended when the predictor is defined differently in the different populations.

The Canadian model was based on a relatively small number of cases (57 events for eight potential predictors in the full model; event per variable ratio: 7:1). When limited positive cases are available, statistical methods, such as bootstrapping procedure, should be used to check whether a developed model is reasonably valid or needs to be adjusted for potential optimism. This procedure has been shown to be superior over split-sample or cross-validation methods [18]. It turned out that the model had a reasonable internal validity; we obtained a correction factor of 0.88; the closer the correction factor is to 1, the lesser the optimism.

4.2. Clinical application

The model revision produced a valid model with the highest discriminative ability over the other approaches. The revised model had a corrected ROC area of 0.75, which meant that in 75% of all possible pairs of apprentices, in which one apprentice is sensitized and one is not, a higher predicted probability is assigned to the apprentice who is

Table 4
Comparison of diagnostic accuracy of the selected cutoff point of the predicted probability produced by different validation methods

| Cutoff point | Number of apprentices (%) ^a in the probability group | Number of apprentices with sensitization (<i>n</i> = 57) <i>n</i> (%) ^b | Number of apprentices without sensitization (<i>n</i> = 357) <i>n</i> (%) ^c | Sensitivity (%) | Specificity (%) | PPV ^d (%) | NPV ^e (%) |
|--------------------|---|---|---|-----------------|-----------------|----------------------|----------------------|
| Cutoff > 0.10 | | | | | | | |
| Recalibrated model | 204 (49.3) | 45 (22.1) | 159 (77.9) | 78.9 | 55.5 | 22.1 | 94.3 |
| Refitted model | 204 (49.3) | 45 (22.1) | 159 (77.9) | 78.9 | 55.5 | 22.1 | 94.3 |
| Revised model | 192 (46.4) | 44 (22.9) | 148 (77.1) | 77.2 | 58.5 | 22.9 | 94.1 |
| Cutoff > 0.15 | | | | | | | |
| Recalibrated model | 102 (24.6) | 32 (31.4) | 70 (68.6) | 56.1 | 80.4 | 31.4 | 92.0 |
| Refitted model | 102 (24.6) | 32 (31.4) | 70 (68.6) | 56.1 | 80.4 | 31.4 | 92.0 |
| Revised model | 135 (32.6) | 40 (29.6) | 95 (70.4) | 70.2 | 73.4 | 29.6 | 93.9 |
| Cutoff > 0.20 | | | | | | | |
| Recalibrated model | 90 (21.7) | 29 (32.2) | 61 (67.8) | 50.9 | 82.9 | 32.2 | 91.4 |
| Refitted model | 89 (21.5) | 29 (32.6) | 60 (67.4) | 50.9 | 83.2 | 32.6 | 91.4 |
| Revised model | 86 (20.8) | 32 (37.2) | 54 (62.8) | 56.1 | 84.9 | 32.6 | 92.4 |

^a Proportion of all apprentices (*n* = 414).

^b Proportion of apprentices with occupational sensitization within the sum score category.

^c Proportion of apprentices without occupational sensitization within the sum score category.

^d Positive predictive value.

^e Negative predictive value.

sensitized. This is a good achievement for a simple model that comprises only six questionnaire items. In their 2002 report, Meijer et al. evaluated the additional value of results from SPT to common animal and non-animal allergens. They obtained a ROC area of 0.76 for the model based on questionnaire information only and 0.86 for the model based on questionnaire information and SPTs. This indicates that additional SPT information involving reactivity to common allergens would lead to a 10% improved discrimination over information obtained from the questionnaires alone. However, following the clinical setting where a diagnosis is started with the anamnesis, an occupational physician can consider the use of a questionnaire-only model as a practical

and inexpensive tool in surveillance for individual prediction of sensitization to LA allergens [4].

If an occupational physician applies the revised diagnostic model and uses the predicted probability of 0.15 or higher as a cutoff to refer the apprentices for SPT to confirm occupational sensitization to LA allergens, one-third of the total apprentices will be referred for SPT, and 70% of all sensitized apprentices will be captured. However, with a low PPV (30%), 70% of the apprentices in the high-probability group will have a negative SPT. The low PPV can be explained by the relatively low incidence of sensitization in the population under study (although this prevalence is much higher than in general population). An occupational physician can also choose a higher cutoff, which offers a higher specificity and PPV, but at the cost of lower sensitivity.

The PPV for this test is comparable with that of prostate-specific antigen (PSA) screening for early detection of prostate cancer, which has a sensitivity of 46%, specificity of 91%, and PPV of 21% for PSA values between 4 and 10 ng/mL and 44% for PSA greater than 10 ng/mL [19]. Interestingly, although both screening tests shared similar diagnostic properties, they were applied in a very different context. In general, false positive findings will lead to unnecessary stress and health care burden, whereas false negative findings will create false sense of security. Nevertheless, the misclassification issue is more crucial when the outcome of interest is a cancer, such as prostate cancer, because false-positive cases will undergo intensive and often invasive diagnostic tests and treatment procedures, whereas false-negative cases may suffer from poor prognosis that could have been prevented by early detection. When the outcome is an (occupational) allergic disease, the consequence of the misclassification yielded by the screening test is less dramatic. In our case, false-positive cases will

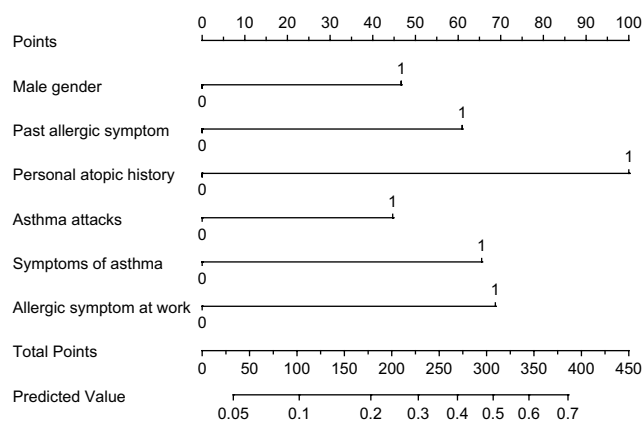


Fig. 2. Nomogram for manual calculation of the predicted probability of sensitization to laboratory animal allergen. Each predictor has a reference line for reading scoring points (default range: 0–100). Once the reader manually totals the points, the predicted probabilities can be read at the bottom. For example, a female apprentice, who has a history of atopy and asthma attack, has a total number of points of 145 (0 + 0 + 100 + 45 + 0 + 0), which corresponds to a predicted probability of wheat sensitization of 0.15.

undergo a simple serological test to confirm the presence of work-related sensitization, and investigation will end when they eventually have a negative serology result. Because occupational sensitization is not a disease, but rather a pre-condition, which is strongly associated with development of occupational respiratory symptoms, it is acceptable to speculate that false-negative cases could be captured in a future round of a periodic surveillance.

In conclusion, we externally validated an existing diagnostic questionnaire model for sensitization to LA allergens. After being updated, the model, which was derived from LA workers, demonstrated its transportability to the animal health apprentices setting. Addition of symptoms of asthma and personal atopic history to the model provided the best diagnostic accuracy, and thus, the revised model could be adopted in Canadian apprentices. The use of this diagnostic questionnaire-based model to predict the likelihood of sensitization to LA allergens can increase the efficiency of health surveillance programs in these apprentices.

Acknowledgments

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). One author (E.S.) is a research fellow with a grant from the Center for Asthma in the Workplace-Canadian Institutes of Health Research.

Conflicts of interest: none declared.

References

- [1] Gautrin D, Infante-Rivard C, Dao T, Magnan-Larose M, Desjardins D, Malo J. Specific IgE-dependent sensitization, atopy, and bronchial hyperresponsiveness in apprentices starting exposure to protein-derived agents. *Am J Respir Crit Care Med* 1997;155(6):1841–7.
- [2] Meijer E, Grobbee DE, Heederik D. Detection of workers sensitised to high molecular weight allergens: a diagnostic study in laboratory animal workers. *Occup Environ Med* 2002;59(3):189–95.
- [3] Suarthana E, Vergouwe Y, Nieuwenhuijsen M, Heederik D, Grobbee DE, Meijer E. Diagnostic model for sensitization in workers exposed to occupational high molecular weight allergens. *Am J Ind Med* 2005;48(3):168–74.
- [4] 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(10):831–7.
- [5] Bleeker SE, Moll HA, Steyerberg EW, Donders ART, Derksen-Lubsen G, Grobbee DE, et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;56:826–32.
- [6] Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;23(16):2567–86.
- [7] Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130(6):515–24.
- [8] Hollander A, Doekes G, Heederik D. Cat and dog allergy and total IgE as risk factors of laboratory animal allergy. *J Allergy Clin Immunol* 1996;98(3):545–54.
- [9] Medical Research Council Committee on the Aetiology of Chronic Bronchitis. Instruction for the use of questionnaire on respiratory symptoms. Dawlish, UK: Holtman, Ltd; 1966.
- [10] Burney PG, Laitinen LA, Perdrizet S, Huckauf H, Tattersfield AE, Chinn S, et al. Validity and repeatability of the IUATLD (1984) Bronchial Symptoms Questionnaire: an international comparison. *Eur Respir J* 1989;2(10):940–5.
- [11] Grammer L, Patterson R. Immunologic evaluation of occupational asthma. In: Bernstein IL, Chan-Yeung M, Malo J, Bernstein DI, editors. *Asthma in the workplace*. 2nd ed. New York: Marcel Dekker Inc.; 1999. p. 159–71.
- [12] Hosmer D, Lemeshow S. *Applied logistic regression*. New York: John Wiley and Sons, Inc; 1989.
- [13] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1):29–36.
- [14] Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med Nov* 1990;9(11):1303–25.
- [15] 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(4):361–87.
- [16] Harrell F. Design: S functions for biostatistical/epidemiologic modeling, testing, estimation, validation, graphics, prediction, and typesetting. Available at: <http://hesweb1.med.virginia.edu/biostat/s/Design.html>; 2002.
- [17] Harrell F. The Hmisc library 2002.
- [18] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49(12):1373–9.
- [19] Gann PH, Hennekens CH, Stampfer MJ. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. *JAMA* 1995;273(4):289–94.