

# **Predicting Occupational Lung Diseases**

**Eva Suarhana**

©2008 Eva Suarhana  
Predicting Occupational Lung Diseases  
Thesis, Utrecht University  
ISBN 9789039347867

Printed by : Drukkerij Ridderprint, Ridderkerk  
Cover illustration : Eva Suarhana and Lisanne van der Voort  
(Multimedia Centre of Veterinary Medicine)  
Layout : H. Otter, Multimedia Centre of Veterinary Medicine

# **Predicting Occupational Lung Diseases**

*Voorspellen van Beroepsgebonden Longziekten*  
(met een samenvatting in het Nederlands)

## **Proefschrift**

ter verkrijging van de graad van doctor aan de Universiteit Utrecht  
op gezag van de rector magnificus, prof.dr. J.C.Stoof,  
ingevolge het besluit van het college voor promoties  
in het openbaar te verdedigen  
op vrijdag 6 juni 2008 des middags te 12.45 uur

door

**Eva Suarthana**

geboren op 24 April 1977  
te Jakarta, Indonesia

**Promotoren:** Prof. dr. ir. D. Heederik  
Prof. dr. D.E. Grobbee

**Co-promotor:** Dr. E. Meijer

The research reported in this thesis was financially supported by the Institute for Risk Assessment Sciences (IRAS), Utrecht University; Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht; Productschap GZP, the Hague; and Aladdin programme for occupational health research, The Netherlands Ministry Social Affairs and Employment.

"If you don't go after what you want, you'll never have it.  
If you don't ask, the answer is always no.  
If you don't step forward, you're always in the same place."

**Nora Roberts**

**Untuk Bapak, Mama, Santhy, Dan Sherley**



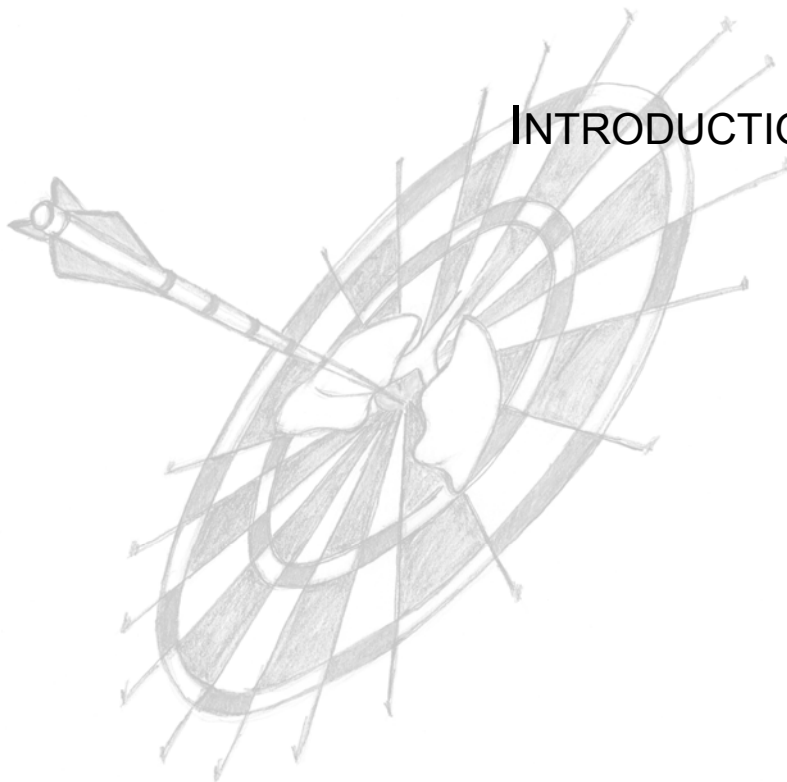
# Contents

1	Introduction	1
2	Diagnostic Models	7
2.1	A diagnostic model for sensitization to wheat allergens	9
2.2	A diagnostic model for sensitization to occupational high molecular weight allergens	27
2.3	A diagnostic model for ruling out pneumoconiosis among construction workers.	39
3	Prognostic Models	53
4	Model Validation	71
5	Model Application	87
6	General Discussion	103
	Summary	109
	Nederlanse Samenvatting	113
	Rangkuman dalam Bahasa Indonesia	117
	Affiliation of Contributors	121
	List of Publications	123
	Acknowledgements	125
	Curriculum Vitae	127



# CHAPTER 1

INTRODUCTION





# INTRODUCTION

## Occupational Lung Diseases

The association between occupational exposures to dust and fumes and lung diseases in different industries (e.g. mining, agriculture, construction industry) has been established a long time ago. In the beginning of the 20<sup>th</sup> century, pneumoconiosis was the largest contributor to the burden of occupational respiratory diseases. Pneumoconiosis mortality, especially silicosis mortality, has been constantly declining with the introduction of exposure control technology and reduction of the workforce in mining industries and foundries after the world war II.(1) Unexpectedly, in the last decade, silicosis seemed back from being absent in the construction industry, as a result of the introduction of portable hand tools that led to high dust exposure.(2)

In the late twentieth century, occupational asthma (OA) has become the most common work-related respiratory disorder.(3, 4) Surveillance of work-related and occupational respiratory diseases (SWORD) in the UK between 1988 and 1998 shows that OA had the highest incidence rate of work related respiratory diseases reported by chest and occupational physicians.(3) Blanc and Toren reported a population attributable risk (PAR) of occupational associated asthma of almost 10% on the basis of general population surveys.(5) A similar figure was reported by researchers from the European Community Respiratory Health Survey (PAR 5-10%).(6) An annual incidence of 2-20 occupational asthma cases per 100,000 workers has been estimated using occupational disease registry information.(7) A more recent study highlights that OA remains the major respiratory problem related to the workplace. From a European open population multi-center survey, the annual incidence of occupational asthma was estimated as 25-30 cases per 100,000 workers. No distinction could be made between specific occupational asthma and work-aggravated asthma.(8)

## Prevention of Occupational Diseases

In general, there are two distinctly different approaches for the prevention of work-related diseases. The first is primary prevention through exposure reduction and exposure control. This is the occupational hygiene approach, which makes use of health based exposure standards for chemical and biological agents in the air and translates these into exposure reduction strategies. Unfortunately this approach is not always possible or difficult to carry out. In such case, secondary prevention approach is required: early detection of occupational diseases through regular medical surveillance.(9-11)

As an illustration, the major cause of OA is the exposure to aeroallergens in the workplace.(11, 12) Although individuals with high exposure to occupational allergens are more likely to have developed sensitization and respiratory symptoms(13, 14), individuals with intermittent or occasional exposure may also be affected(15). A recent exposure-response study suggests that sensitization risk does occur even at very low exposure levels and some residual excess risk, even at the lowest possible exposure can not be excluded.(16) Therefore, parallel to exposure reduction and exposure control, surveillance should routinely be conducted to early detect the occupational diseases.(9-11)

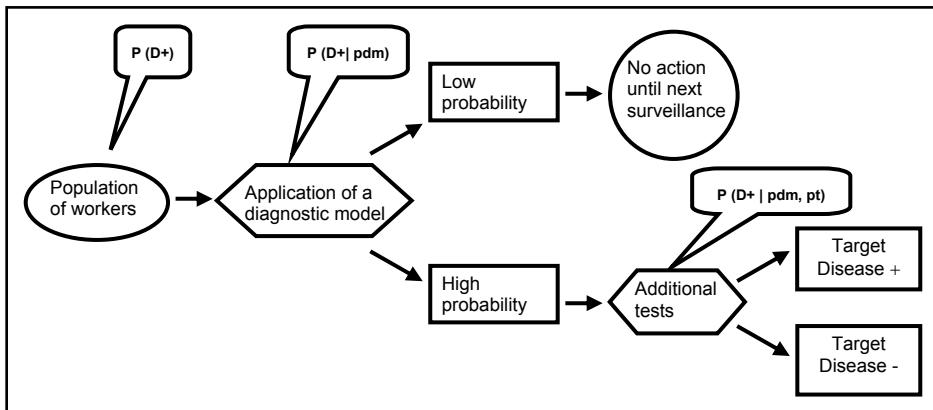
## Evidence Based Occupational Health Practice

The new paradigm in occupational health, known as evidence based occupational health practice, expects the occupational health practitioner to use the results of appropriate studies that have evaluated cause-effect relationships and the efficacy of prevention strategies. These guidelines stress the importance of interventions focused at population

level rather than at the individual level and provides suggestions for policy-making decision.(17)

The British Occupational Health Research Foundation recently produced comprehensive guidelines for prevention, identification, and management of occupational asthma based on almost 500 original studies.(11) Having all available information collected and rated for quality, the question is how practical these strategies are to be implemented in occupational health practice.(18) Methods commonly used in surveillance programs to identify cases of occupational asthma are respiratory questionnaires, spirometry, and skin-prick testing or identification of serum specific IgE to occupational sensitizers.(9) Nevertheless, no attempt has been made to quantify the prior- and post-test probability of the presence of OA given a test result. For example, after completion of a set of questionnaires, can an occupational physician accurately quantify the probability of OA? Is the probability sufficiently high enough to decide that further medical tests are needed to confirm the presence of OA? Or, is the probability low enough so that withholding advanced tests will not harm this worker?

Prediction studies may answer to the above questions by developing models in which personal and work related characteristics are used to estimate the individual probability of the presence or occurrence of an outcome (e.g. disease of interest or its related condition). With a standard and objective quantification, an occupational physician can identify workers with a low probability of having a disease and exclude them from further medical investigations (Figure 1). Such scientific approach may greatly assist the occupational physician in formalizing the decision-making process, as has been proposed by Meijer *et al* in surveillance of laboratory animal workers.(19) It would also increase the efficiency of the surveillance, as it decreases the number of unnecessary tests.



**Figure 1** Surveillance for respiratory allergy among an allergen exposed population.  $P(D+)$  is the prior probability of the target disease in the population.  $P(D+ | pdm)$  is the predicted probability of the target disease given the result of the application of the diagnostic model. Workers will be stratified based on their disease probability; additional test will only be performed in workers with a high probability.  $P(D+ | pdm, pt)$  is the predicted probability of the target disease given the result of the application of the diagnostic model and additional test.

The motive of secondary prevention approach is to early detect a disease in order to allow early intervention and management in the hope to reduce complications from the disease. Nevertheless, like other medical diagnostic tests, a diagnostic model may produce misclassified outcomes. It may classify a diseased subject as non-diseased and thus, create a false sense of security. On the other hand, it may assign a healthy subject into the diseased group and thus, leads to unnecessary stress and intervention. Rational decision-

making in a surveillance setting is heavily dependent on improvement of the clinical outcome as a result of early diagnoses; the burden of disability from the clinical outcome; and adequacy of the cost, accuracy, and acceptability of the surveillance test.(20) Therefore, assuring the diagnostic performance of the model and choosing a favourable cut-off point for the decision making process is important before implementing such prediction model.

### **Aims and Outline of This Thesis**

A clinical diagnosis requires a synthesis of all available test results to estimate the probability of the presence of a disease. Test results can be obtained for example from questionnaire, physical examination, laboratory test, or imaging. A diagnosis rarely arises from the presence of one test result and predicting the likelihood of a disease is complicated by the fact that many test results generate more or less identical information. Therefore, it is important to evaluate the independency and additional predictive value of a test given the presence of earlier information. Multivariable regression analysis offers a solution for this. This technique allows that weights are given to each independent predictor in the estimation equation. The ability of a test to discriminate individuals with and without the outcome is evaluated using the area under the receiver operating characteristic curve. The contribution of a test as a single test or in combination with others test in the prediction of a disease is assessed. This approach leads to so called prediction models, which are decision making tools for clinicians and provide estimation of the probability of an outcome at present (diagnosis) or in future (prognosis). When applied in occupational health practice, diagnostic models enable occupational physician to deal with uncertainties in diagnosing workers at risk of having occupational diseases. The main goal is to optimize risk estimation at low costs. Prognostic models may initiate counseling and interventions and are thus useful for identification of specific groups at risk. In this thesis, the concepts of diagnostic and prognostic research have been applied to occupational lung diseases. Optimization and identification of specific risk groups is also explored.

Chapter 2 describes several diagnostic models for occupational lung diseases. In Chapter 2.1 a screening tool for sensitization to wheat allergens was developed and validated in Dutch bakery workers. Chapter 2.2 illustrates that it is possible to develop one model for sensitization to high molecular weight allergens in Dutch bakery workers (exposed to wheat and or fungal alpha amylase allergens) and laboratory animal workers (exposed to rat and or mouse urinary allergens). The model was validated in British laboratory animal workers. In Chapter 2.3 a diagnostic model for pneumoconiosis was developed in Dutch construction workers who were exposed to silica dust. All diagnostic models were transformed into easy-to-use scoring system to facilitate their use in practice.

Chapter 3 describes the development of prognostic models for occurrence of occupational sensitization and respiratory symptoms in Canadian apprentices in animal health technology. The prognostic value of questionnaire, skin-prick test, and bronchial challenge to methacholine test, as a single test or in combination with others, was assessed.

Chapter 4 focuses on assessing the generalizability of a prediction model. A series of statistical approaches were used to externally validate a questionnaire model for sensitization to LA allergens -which was developed in Dutch workers- in Canadian animal health technology apprentices setting. A new model was eventually developed in Canadian apprentices and compared to the original Dutch model. Model revision was done to evaluate if inclusion of new predictors from the Canadian setting could improve the performance of the original model.

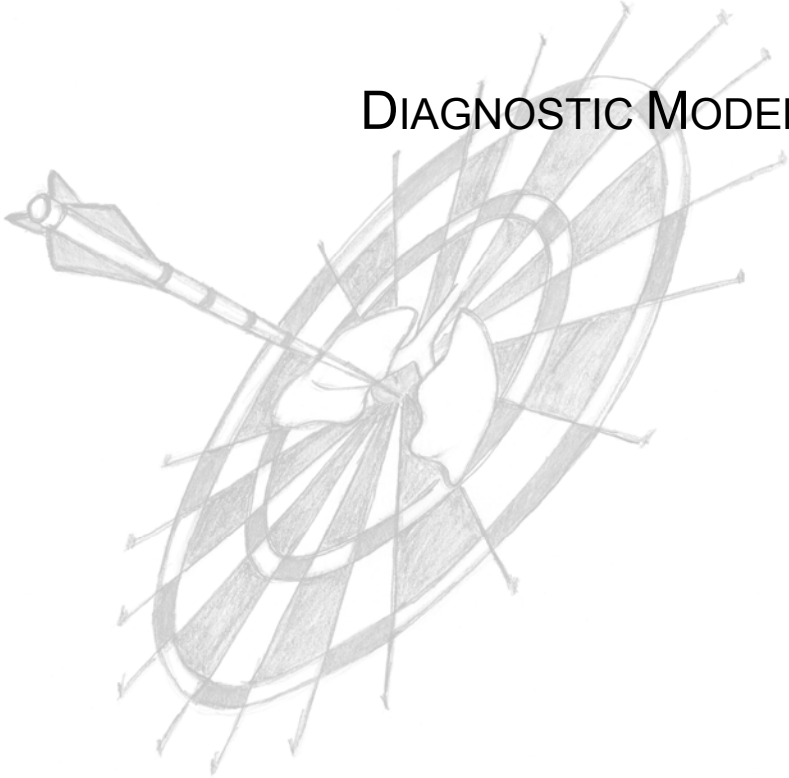
Chapter 5 describes the application of a diagnostic model for occupational sensitization in ongoing surveillance for respiratory diseases in the baking and flour-producing industries in the Netherlands. The diagnostic model was applied in 5,325 Dutch bakery and flour and enzyme exposed workers who participated in surveillance. This chapter illustrates how a diagnostic model may improve the efficiency of such surveillance.

**REFERENCES**

1. Heederik D, Attfield M. Characterization of Dust Exposure for the Study of Chronic Occupational Lung Disease: A Comparison of Different Exposure Assessment Strategies. In; 2000. p. 982-990.
2. Tjoe Nij E, Burdorf A, Parker J, Attfield M, van Duivenbooden C, Heederik D. Radiographic abnormalities among construction workers exposed to quartz containing dust. *Occup Environ Med* 2003;60(6):410-7.
3. Meyer JD, Holt DL, Cherry NM, McDonald JC. SWORD '98: Surveillance of work-related and occupational respiratory disease in the UK. *Occup Med* 1999;49(8):485-489.
4. Balmes J, Becklake M, Blanc P, Henneberger P, Kreiss K, Mapp C, et al. American Thoracic Society Statement: Occupational contribution to the burden of airway disease. *Am J Respir Crit Care Med* 2003;167(5):787-97.
5. Blanc PD, Toren K. How much adult asthma can be attributed to occupational factors? *Am J Med* 1999;107(6):580-7.
6. Kogevinas M, Anto JM, Sunyer J, Tobias A, Kromhout H, Burney P. Occupational asthma in Europe and other industrialised areas: a population-based study. European Community Respiratory Health Survey Study Group. *Lancet* 1999;353(9166):1750-4.
7. Heederik D, Douwes J, Wouters I, Doekes G. Epidemiology of occupational respiratory diseases and risk factors. *Eur Respir Mon* 2000;15:429-47.
8. Kogevinas M, Zock JP, Jarvis D, Kromhout H, Lillienberg L, Plana E, et al. Exposure to substances in the workplace and new-onset asthma: an international prospective population-based study (ECRHS-II). *Lancet* 2007;370(9584):336-41.
9. Cullinan P, Tarlo S, Nemery B. The prevention of occupational asthma. *Eur Respir J* 2003;22(5):853-60.
10. Tarlo SM, Liss GM. Prevention of occupational asthma--practical implications for occupational physicians. *Occup Med (Lond)* 2005;55(8):588-94.
11. Nicholson PJ, Cullinan P, Taylor AJ, Burge PS, Boyle C. Evidence based guidelines for the prevention, identification, and management of occupational asthma. *Occup Environ Med* 2005;62(5):290-9.
12. Sastre J, Vandenplas O, Park H-S. Pathogenesis of occupational asthma. *Eur Respir J* 2003;22(2):364-373.
13. Houba R, Heederik D, Doekes G. Wheat sensitization and work-related symptoms in the baking industry are preventable. An epidemiologic study. *Am J Respir Crit Care Med* 1998;158(5 Pt 1):1499-503.
14. Hollander A, Heederik D, Doekes G. Respiratory allergy to rats: exposure-response relationships in laboratory animal workers. *Am J Respir Crit Care Med* 1997;155(2):562-7.
15. Gautrin D, Ghezzi H, Infante-Rivard C, Malo JL. Incidence and determinants of IgE-mediated sensitization in apprentices. A prospective study. *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1222-8.
16. Heederik D, Houba R. An exploratory quantitative risk assessment for high molecular weight sensitizers: wheat flour. In; 2001. p. 175-185.
17. Franco G. The future of occupational health practice: reconciling customer expectation and evidence-based practice. *Occup Med (Lond)* 2001;51(8):482-4.
18. Tarlo SM, Liss GM. Evidence based guidelines for the prevention, identification, and management of occupational asthma. *Occup Environ Med* 2005;62(5):288-9.
19. 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.
20. Sackett D, Haynes R, Guyatt G, Tugwell P. *Clinical Epidemiology: A basic science for clinical medicine*. 2nd ed. Boston: Little, Brown and Company; 1991.

# CHAPTER 2

## DIAGNOSTIC MODELS





# CHAPTER 2.1

## A DIAGNOSTIC MODEL FOR SENSITIZATION TO WHEAT ALLERGENS

The contents of this chapter are based on  
E Suarthana, Y.Vergouwe, KGM Moons, D.E. Grobbee, J de Monchy, D Heederik, E Meijer  
Development and validation of a diagnostic model for detection of  
sensitization to wheat allergens: *Prediction study in bakery workers*  
*Submitted in a revised version*



# A DIAGNOSTIC MODEL FOR SENSITIZATION TO WHEAT ALLERGENS

---

EVA SUARTHANA, YVONNE VERGOUWE, KARL MOONS, DIEDERICK E. GROBBEE,  
JAN DE MONCHY, DICK HEEDERIK, EVERT MEIJER

---

## ABSTRACT

**Background:** Occupational asthma (OA) has been shown to be the major respiratory problem related to exposure to high-molecular weight allergens, including wheat flour and fungal  $\alpha$ -amylase allergens in bakeries and the flour producing industry. Effective and timely detection of OA precursor may lead to a decrease in OA prevalence. A diagnostic study was aimed at developing and validating a questionnaire model for detection of sensitization to wheat allergens in bakery workers.

**Methods:** Multivariable logistic regression modelling was done in 867 bakery workers (development set). The discrimination and internal validity of the model were assessed. External validation of the original model was done in 390 independent bakery workers (validation set). Update methods were applied when necessary. Both the calibration and discrimination of the updated model were assessed. The final diagnostic model was transformed into a scoring rule.

**Results:** The questionnaire model with six predictors showed a good discrimination (area under the receiver operating curve (ROC) 0.76; 95% confidence interval 0.71 to 0.81), and internal validity (shrinkage factor of 0.89; corrected ROC area 0.75). Application of the original model in the validation set gave a reasonable discrimination (ROC area 0.69, 95% confidence interval 0.62 to 0.75), but poor calibration ( $U$  statistic=0.046). Recalibration of the intercept improved the model's calibration ( $U$  statistic=0.0004). Re-estimation of the intercept and all regression coefficients did not improve the calibration ( $U$  statistic of -0.004) and led to similar discrimination (ROC area 0.71). The model with re-estimated intercept was chosen as the final model.

**Conclusion:** The diagnostic model for sensitization to wheat allergens has good diagnostic accuracy. External validation and updating in a sample of independent bakery workers yielded a model that is ready to be used in a wider population of bakery workers.

**KEYWORDS:** accuracy, diagnostic model, occupational allergic disease, validity

## INTRODUCTION

In the past few decades, occupational asthma (OA) has been shown to be the major respiratory problem related to exposure to high-molecular weight allergens, including wheat flour and fungal  $\alpha$ -amylase allergens in bakery and flour producing industry.(1-3) Although highly exposed individuals are more likely to have serious complaints and disability(4), workers with intermittent or occasional exposures may also be affected(5). Recent studies suggest an elevated risk of sensitization even at very low exposure levels.(6) Therefore, strategies to enhance early detection should be considered.

Guidelines for early detection of OA have been published and usually include pre-placement medical evaluation and testing, periodic follow-up questionnaires, interval medical examinations, and selective use of additional tests such as skin-prick testing and spirometry.(3, 7) Nevertheless, none of these guidelines specify how early detection should be carried out and none of them use prediction model in which personal and work related characteristics are used in combination to estimate individual's probability of the presence or occurrence of OA. Because the true diagnosis of OA can only be made in a clinical setting, precursors of the disease can be detected first. By effectively and timely detecting this precursor, progress can be made in decreasing OA prevalence.

For the detection of sensitization to occupational high molecular weight allergens, a general diagnostic model was developed earlier and comprised questions as well as laboratory results (total IgE, and IgE to common allergens).(8) However, to have a simple and evidence based decision tool at low costs to be implemented in a national surveillance program to early detect bakers with allergic diseases, a different model had to be developed. To stay close to the diagnostic sequence in practice, a diagnostic model consisting of only questionnaire items is therefore attractive.

We aimed to develop a diagnostic questionnaire model for sensitization to wheat allergens in bakery workers. Then, we determined whether this diagnostic model performed adequately in new workers(9, 10) and whether the accuracy of the diagnostic model could be further improved (updated) (11, 12). The final diagnostic model was then simplified into easy-to-use scoring rule to enhance its use in practice. Such tool will enable occupational physicians to objectively estimate risk of having occupational allergic diseases for every individual at low costs and thus, improve the efficiency of surveillance program.

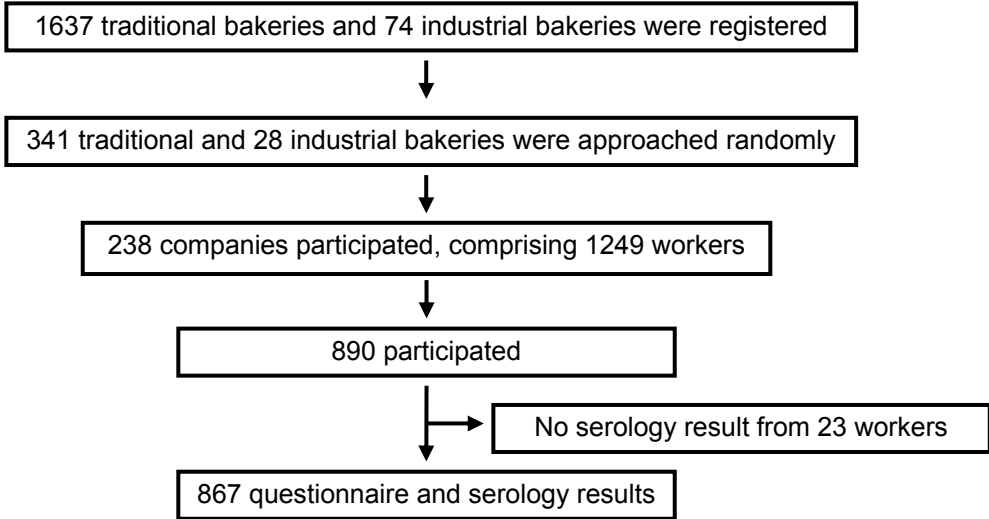
## METHODS

### Study subjects and design

We used data from a study performed in the framework of a National Occupational Respiratory Allergy Surveillance among bakery workers in the Netherlands to develop the diagnostic model (development set), including 660 traditional and 230 industrial workers (Figure 1).(13) We used a second data set from a cross sectional study in four sectors of baking and flour producing industries to validate and update the diagnostic model (validation set). This study included 153 workers from traditional bakeries, 85 from industrial bakeries, 83 from the milling industry, and 69 from the baking product manufacturers.(14) We considered the later three groups as non-traditional workers. Informed consent was obtained from all participating workers in both studies.

In both studies, workers were asked to complete a self-administered respiratory questionnaire which contained items from internationally standardized questionnaires. (15, 16) The questionnaire comprised items on employment (job, tasks) data, history of lower and upper respiratory symptoms, allergic symptoms due to common allergens, symptoms

suggesting bronchial hyperresponsiveness, work-related upper and lower respiratory symptoms, skin symptoms, absenteeism, medication use, changes in tasks or jobs, and smoking habits.



**Figure 1** Recruitment of participants in the development set

### Diagnostic outcome (reference standard)

In both studies, sensitization to wheat allergens was chosen as the outcome. In the development set, 23 workers without serology result were excluded from the analysis, leaving  $n=867$ . Specific IgE antibodies were measured with a modified enzyme immunoassay (EIA).(17) Sensitization to wheat allergens was considered present if the optical density (OD) of 492 exceeded the OD +0.10 of the reagent blank (no serum control). In the validation set, specific IgE antibodies were measured with a commercial immunoassay (Pharmacia CAP system, Pharmacia Diagnostics, Sweden). IgE sensitization class II (concentration 0.7 kU/L or greater) to wheat allergens was defined as positive IgE serology. The EIA for wheat allergens has been compared earlier to the CAP assay. The overall agreement was satisfactory, and good at higher titres.(17)

### Diagnostic predictors

Ideally there should be at least 10 events per candidate predictor (10 to 1 rule).(18) In the development set there were 108 sensitized workers whereas the questionnaire consisted of 75 items on symptoms. Therefore, we used principal components analysis (PCA) to reduce the number of candidate predictors.(19) Only components with an Eigenvalue greater than 1 (i.e. explaining more variance than a single predictor) were extracted in the PCA. Based upon these results, new clustered predictors (symptoms) were defined. Individual predictors were only included in the clusters when factor loadings were higher than 0.6.

The PCA resulted in 14 clusters of correlated symptoms with Eigenvalue of one or greater. The following clusters were identified and named (*post hoc*) as adequately as possible: “asthma symptoms in the last 12 months”, “shortness of breath and wheeze”, “cough and phlegm in the last 12 months”, “upper respiratory symptoms”, “allergic symptoms in contact with HDM and/or plants”, “allergic symptoms in contact with pets”, “allergic symptoms due to certain food”, “symptoms suggestive of BHR Induced by smoke”,

“symptoms suggestive of BHR induced by change of temperature”, “skin symptoms”, “during work upper respiratory symptoms”, “during work lower respiratory symptoms”, “after work upper respiratory symptoms”, and “after work lower respiratory symptoms”. For details regarding included symptoms in each cluster, see Appendix. A positive cluster was defined by the presence of at least one of the symptoms that composed the cluster.

To further reduce the number of clusters, we combined clusters with a Pearson correlation coefficient of 0.7 or higher, because they contain mainly the same information. From this step, “during work upper respiratory symptoms” and “after work upper respiratory symptoms” were combined and renamed as “work-related upper respiratory symptoms”. Similarly, “during work lower respiratory symptoms” and “after work lower respiratory symptoms” were combined and renamed as “work-related lower respiratory symptoms”.

Based on the literature, age, gender, and smoking habits were not included in the multivariable analysis because they are not associated with specific sensitization to wheat allergens.(2) (20) Therefore, we started the full multivariable logistic regression model with 13 predictors: 12 new cluster variables from the PCA and type of bakery (industrial or traditional).

## **Data analysis**

### *Model development*

A diagnostic model for the presence of sensitization to wheat allergens was derived from the development set. Backward stepwise selection of the candidate predictors was applied using the likelihood ratio test with a p-value according to Akaike’s Information Criterion (p-value < 0.157).(21, 22)

The model’s ability to discriminate sensitized from not-sensitized workers (discrimination) was assessed by the ROC area. The ROC area can range from 0.5 (no discrimination) to 1.0 (perfect discrimination).(23) A bootstrapping procedure was done to assess the internal validity of the model (S-Plus 6 for windows, Insightful Corp).(24) Hundred bootstrap samples were drawn from the development set. In each bootstrap sample, the modelling process was repeated. This resulted in 100 models that were applied to the original development set. This procedure produced a corrected model’s ROC area and a shrinkage factor. The regression coefficients of the predictors in the model were then multiplied by this shrinkage factor to prevent overfitting of the regression coefficients and optimism of the model when applied to new subjects.(12, 24, 25)

### *Methods to validate and update the diagnostic model*

We then evaluated the model’s external validity by applying the regression formula from the original developed model (i.e. no adjustment) to the validation set. We compared the predicted probabilities with the observed outcomes and used measures of calibration and discrimination. The agreement between the predicted probabilities and the observed frequencies for sensitization (calibration) was evaluated graphically.(26) A line described by an intercept of 0 and a slope of 1 indicates perfect calibration. We also used unreliability index  $U$  to quantify the miscalibration. A  $U$ -index close to 0 indicates that the intercept is not statistically significant different from 0 and the slope not different from 1.(27) The discrimination was assessed with ROC area. If the model did not perform well in the validation set, we updated the model to improve the validity (11, 12) First, only the intercept of the original model was adjusted. This method would not influence the model’s discrimination since the relative ranking of the sensitized workers versus not sensitized was not altered. Secondly, the intercept and the regression coefficients of the original model were adjusted to evaluate if more advanced updating methods could yield further improvement. All newly estimated regression coefficients (method 2) were shrunk to prevent overfitting.

### *Model application*

To facilitate the application of the final model in practice, the model was converted to an easy-to-use scoring rule. To derive the scores, the regression coefficients of the predictors were divided by the smallest one, and rounded to the nearest half integer. This scoring rule was then applied in the combined data of the development and validation set. For each worker, the sum score was calculated. Categories of these scores were related to the corresponding probabilities of being sensitized to wheat allergens.

### *Missing data*

Deleting subjects with a missing value does not only lead to a loss of statistical power, but often also to biased results. Therefore, imputing missing values is generally preferred to complete case analysis.(28, 29) Missing data were single imputed with values obtained from regression equations using SPSS version 15.0 (SPSS, Inc., Chicago, IL, USA). In the development set there were 286 (32.1%) subjects with 1 or more missing questionnaire items, of which 135 had 1 missing item. In the validation set there were 144 (36.9%) subjects with at least 1 missing questionnaire items, of which 100 had 1 missing item.

## **RESULTS**

The distribution of the candidate predictors in the development and validation set are shown in table 1. Most (74.4%) workers in the development set were traditional bakers, whereas this was 39.2% in the validation set. The prevalence of wheat sensitization was 12.5% and 19.5% in the development and validation set, respectively. Nasoconjunctival symptoms, asthma symptoms, shortness of breath and wheeze, work-related respiratory symptoms, both upper and lower, were strongly associated with wheat sensitization (table 2).

### *Model development*

The predictors that remained in the model after backward stepwise selection are presented in table 3. The model showed good discrimination (ROC area 0.77, 95% confidence interval 0.72 to 0.82). The bootstrapping procedure yielded an optimism corrected ROC area of 0.75 (0.71 to 0.81).

### *Model validation and update*

The discrimination of the original model in the validation set was reasonable (ROC area of 0.69, 95% confidence interval 0.62 to 0.75), but the calibration poor ( $U$  statistic=0.046). The calibration plot shows that the predicted probabilities were all systematically lower than the observed frequency (fig. 2a). Recalibration of the intercept (to -3.05) of the original model in the validation set (table 4) improved the calibration ( $U$  statistic=0.0004, figure 2b). Re-estimation of the intercept and all regression coefficients did not improve the calibration ( $U$  statistic of -0.004, figure 2c) and led to similar discrimination (ROC area 0.71). Therefore, we used the model with a new intercept of as our final model.

**Table 1** Distribution of personal characteristics and potential predictors in the development and validation set

	Development set n=867	Validation set n=390
Age years *	40 (17 to 79)	37 (17 to 63)
Working years *	12 (1 to 54)	10 (1 to 45)
Traditional bakers	645 (74.4)	153 (39.2)
Male gender	820 (94.6)	380 (97.4)
Ever smoker	509 (58.7)	220 (56.4)
Nasoconjunctival symptoms in the last 12 months †	560 (64.6)	274 (70.3)
Asthma symptoms in the last 12 months ‡	32 (3.7)	36 (9.2)
Shortness of breath and wheeze §	261 (30.1)	136 (34.9)
Cough and phlegm in the last 12 months ¶	138 (15.9)	not available
Allergic symptoms in contact with HDM and/or plants #	226 (26.1)	93 (23.8)
Allergic symptoms in contact with pets **	95 (11.0)	38 (9.7)
Allergic symptoms after ingestion particular food ††	103 (11.9)	9 (2.3)
Symptoms suggestive of BHR induced by dust and/or smoke ††	253 (29.2)	104 (26.7)
Symptoms suggestive of BHR induced by change in temperature §§	145 (16.7)	66 (16.9)
Skin symptoms ¶¶	357 (41.2)	171 (43.8)
Work-related upper respiratory symptoms ###	200 (23.1)	159 (40.8)
Work-related lower respiratory symptoms ***	104 (12.0)	88 (22.6)
<i>Outcome</i>		
IgE sensitization to wheat allergens	108 (12.5)	76 (19.5)

Data presented as n (%) unless otherwise stated.

\* Median (range)

† Present if experienced allergic symptoms including hay-fever; blocked/runny nose or sneezing; awoken due to nasal symptoms; and/or itchy/ watery/ red eyes in the last 12 months

‡ Present if experience at least one of the following in the past the last 12 months: asthma attack, awoken due asthma attack, asthma attack induced by exercise.

§ Present if experience at least one of the following: respiratory problem, shortness of breath after walking with normal pace, or wheezing

¶ Present if experienced chronic cough (3 months) with or without phlegm

# Present if experienced upper and/or lower respiratory and/or skin symptoms in contact with HDM or plants

\*\* Present if experienced upper and/or lower respiratory and/or skin symptoms in contact with pets

†† Present if experienced upper and/or lower respiratory and/or skin symptoms after ingestion of certain food

‡‡ Present if experienced respiratory problem due dust, smoke, or tobacco

§§ Present if experienced respiratory problem due to change in temperature

¶¶ Present if experienced dry skin, itchy skin, and/or eczema in the last 12 months

### Present if experienced itchy or blocked nose and/or itchy eyes during or after work, which improve when away from work.

\*\*\* Present if experienced asthma, wheezing, shortness of breath, and/or chest tightness during or after work, which improve when away from work.

**Table 2** Univariable association between the potential predictors and sensitization to wheat allergens in the development set

	Not-sensitized (n=759)	Sensitized (n=108)	OR (95% CI)
Traditional bakers	556 (73.3)	89 (82.4)	1.7 (1.0 to 2.8)
Nasoconjunctival symptoms in the last 12 months	465 (61.3)	95 (88.0)	4.6 (2.5 to 8.4)
Asthma symptoms in the last 12 months	18 (2.4)	14 (13.0)	6.1 (3.0 to 12.7)
Shortness of breath and wheeze	193 (25.4)	68 (63.0)	5.0 (3.3 to 7.6)
Cough and phlegm in the last 12 months	114 (15.0)	24 (22.2)	1.6 (1.0 to 2.7)
Allergic symptoms in contact with HDM and/or plants	177 (23.3)	49 (45.4)	2.7 (1.8 to 4.1)
Allergic symptoms in contact with pets	74 (9.7)	21 (19.4)	2.2 (1.3 to 3.8)
Allergic symptoms after ingestion particular food	81 (10.7)	22 (20.4)	2.1 (1.3 to 3.6)
Symptoms suggestive of BHR induced by dust and/or smoke	206 (27.1)	47 (43.5)	2.1 (1.4 to 3.1)
Symptoms suggestive of BHR induced by change in temperature	112 (14.8)	33 (30.6)	2.5 (1.6 to 4.0)
Skin symptoms	306 (40.3)	51 (47.2)	1.3 (0.9 to 2.0)
Work-related upper respiratory symptoms	144 (19.0)	56 (51.9)	4.6 (3.0 to 7.0)
Work-related lower respiratory symptoms	64 (8.4)	40 (37.0)	6.4 (4.0 to 10.2)

Data presented as n (%) unless otherwise stated.

OR: odds ratio, CI: confidence interval, HDM: house dust mite, BHR: bronchial hyperresponsiveness

**Table 3** Multivariable associations between the predictors and sensitization to wheat allergens

	$\beta^*$	OR (95% CI)
Intercept	-3.66	
Traditional baker	0.67	2.2 (1.2 to 3.9)
Nasoconjunctival symptoms in the last 12 months	0.72	2.3 (1.2 to 4.5)
Asthma symptoms in the last 12 months	0.63	2.0 (0.9 to 4.4)
Shortness of breath and wheeze	0.61	2.3 (1.3 to 3.8)
Work-related upper respiratory symptoms	0.47	1.7 (0.9 to 3.1)
Work-related lower respiratory symptoms	0.61	2.2 (1.1 to 4.4)
<i>U</i> -statistic	0.003	
ROC area, 95% CI	0.75 (0.71 to 0.81)	

\* Regression coefficient after multiplication by the shrinkage factor (0.89) from bootstrapping procedure

OR: odds ratio, CI: confidence interval, ROC: receiver operating area, CI: confidence interval

The predicted probability of wheat sensitization can be calculated using the following formula:  $P(\text{sensitization}) = 1 / (1 + \exp(-(-3.66 + \text{traditional baker} * 0.67 + \text{nasoconjunctival symptoms in the last 12 months} * 0.72 + \text{asthma symptoms in the last 12 months} * 0.63 + \text{shortness of breath and wheeze} * 0.61 + \text{work-related upper respiratory symptoms} * 0.47 + \text{work-related lower respiratory symptoms} * 0.61)))$ . Predictor is valued as 1 when present and 0 when absent.

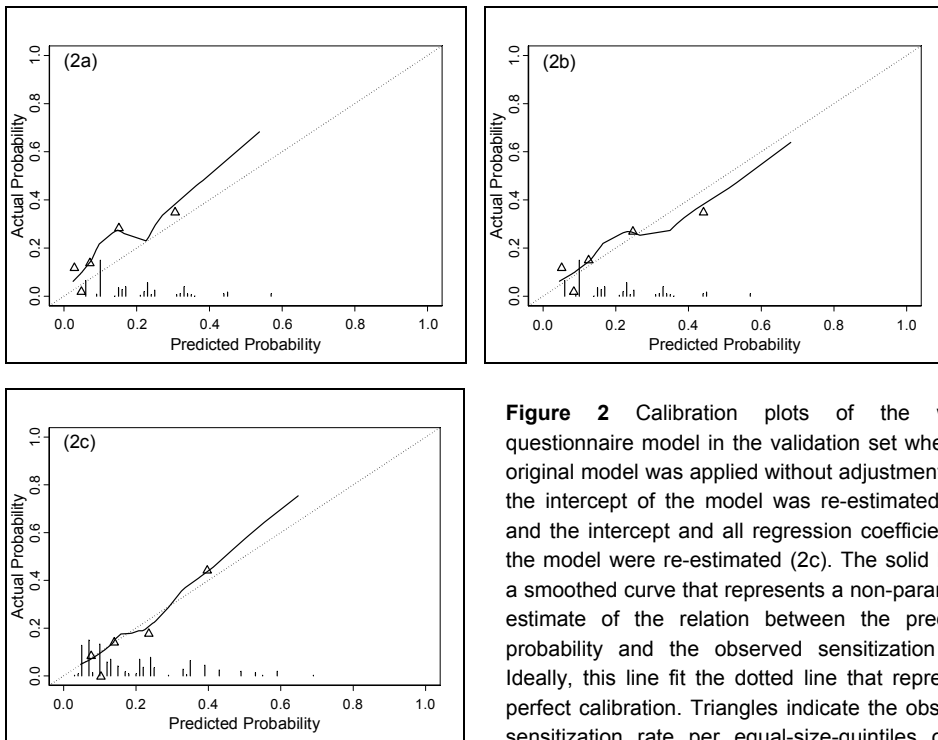
**Table 4** The updated model for wheat sensitization in the validation set and their accuracy

	Recalibration of the intercept (method 1)	Recalibration of the intercept and regression coefficients (method 2)
	$\beta^*$	$\beta^\dagger$
Intercept	-3.05	-2.61
Traditional baker	0.67	0.60
Nasoconjunctival symptoms in the last 12 months	0.72	0.22
Asthma symptoms in the last 12 months	0.63	0.74
Shortness of breath and wheeze	0.61	0.55
Work-related upper respiratory symptoms	0.47	1.11
Work-related lower respiratory symptoms	0.61	-0.39
<i>U</i> -statistic	0.0004	-0.004
ROC area, 95% CI	0.69 (0.63 to 0.76)	0.71 (0.64 to 0.77)

\* Regression coefficient

† Regression coefficient after multiplication by shrinkage factor (0.90) from bootstrapping procedure

ROC: receiver operating area, CI: confidence interval



**Figure 2** Calibration plots of the wheat questionnaire model in the validation set when the original model was applied without adjustment (2a), the intercept of the model was re-estimated (2b), and the intercept and all regression coefficients of the model were re-estimated (2c). 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 fit the dotted line that represents perfect calibration. Triangles indicate the observed sensitization rate per equal-size-quintiles of the predicted probability. Distribution of the predicted probabilities is indicated with vertical lines at the bottom.

**Model application in the combined data**

To enhance its application in practice, the final regression model was simplified into a scoring rule (figure 3). The discriminative ability of the scoring rule in the combined data was slightly higher than the discrimination of the final regression model in the validation set (ROC area 0.73, 95% confidence interval 0.69 to 0.77). The corresponding predicted probabilities of wheat sensitization for each sum scores was calculated (figure 3, lower part). The percentage of the sensitized workers rose with an increasing sum score. As an example of how to use this chart, a traditional baker with no symptom except hay-fever had a sum score of 2 (1+1+0+0+0). There were 331 bakers with a sum score of 2, of which 40 (12%) were sensitized. The mean predicted probability of wheat sensitization of this group was 16%.

Variable in the model	Value	Score					
Type of bakery	If traditional baker	1					
Nasoconjunctival symptoms in the last 12 months	If in the last 12 months experienced at least one of the following: <ul style="list-style-type: none"> <li>allergic symptoms including hay-fever</li> <li>blocked/runny nose or sneezing</li> <li>awaken due to nasal symptoms</li> <li>itchy/ watery/ red eyes</li> </ul>	1					
Asthma symptoms in the last 12 months	If in the last 12 months experienced at least one of the following: <ul style="list-style-type: none"> <li>asthma attack</li> <li>awaken due asthma attack</li> <li>asthma attack induced by exercise</li> </ul>	1					
Shortness of breath and wheeze	If experienced at least one of the following: <ul style="list-style-type: none"> <li>respiratory problem</li> <li>shortness of breath after walking in normal pace</li> <li>wheezing in the last 12 months</li> </ul>	1					
Work-related upper respiratory symptoms	If experienced at least one of the following symptom during or after work: <ul style="list-style-type: none"> <li>itchy nose</li> <li>blocked nose</li> <li>itchy eyes</li> </ul> Symptom should improve when away from work	1					
Work-related lower respiratory symptoms	If experienced at least one of the following symptom during or after work: <ul style="list-style-type: none"> <li>asthma</li> <li>wheezing</li> <li>shortness of breath</li> <li>chest tightness</li> </ul> Symptom should improve when away from work	1					
		Sum score ...					
Sum score	0	1	2	3	4	5	6
n	121	374	331	220	123	68	20
Observed wheat sensitization (%)	5	5	12	19	24	52	65
Mean predicted probability of wheat sensitization (%)	5	9	16	24	37	53	68

**Figure 3** Scoring rule for sensitization to wheat allergens and the corresponding predicted probability

We then categorized the workers into three groups: low probability (sum score 0-1), intermediate probability (sum score 2 -3) and high probability (sum score  $\geq 4$ ). A clear association could be observed between these probability groups and the prevalence of sensitization and symptoms (table 5). Only 3 to 8% of workers in the low probability group had allergic symptoms, 10% used medication to improve respiratory complaints, and 0.2% changed job due to allergic symptoms. These figures were very much lower compared to workers in the intermediate or high probability group.

**Table 5** General and health characteristics across low, intermediate and high probability groups

	Low probability (sum score 0-1) n=495	Intermediate probability (sum score 2-3) n=551	High probability (sum score $\geq 4$ ) n=211
Age years *	41 (18 to 79)	37 (17 to 79)	37 (17 to 63)
Work duration years *	13 (1 to 54)	9.6 (1 to 50)	11 (1 to 44)
Male	477 (96.4)	519 (94.2)	204 (96.7)
Allergic symptoms in contact with HDM and/or plants	37 (8)	171 (31)	111 (53)
Allergic symptoms in contact with pets	13 (3)	60 (11)	60 (28)
Allergic symptoms after ingestion particular food	20 (4)	54 (10)	38 (18)
Symptoms suggestive of BHR induced by dust and/or smoke	80 (16)	159 (29)	118 (56)
Symptoms suggestive of BHR induced by change in temperature	21 (4)	85 (15)	105 (50)
Use of medication to improve respiratory complaints in the last 12 months (e.g. inhalants)	48 (10)	173 (31)	116 (55)
Change of function or task due to respiratory symptoms	1 (0.2)	3 (0.5)	17 (8)
IgE sensitization to wheat allergens	25 (5)	82 (15)	77 (37)
Predicted probability of wheat sensitization †	0.08 (0.001)	0.19 (0.002)	0.45 (0.007)

Data presented as n (%) unless otherwise stated.

HDM, house dust mite; BHR, bronchial hyperresponsiveness

\* Median (range)

† Mean (standard error)

## DISCUSSION

We developed and validated a simple diagnostic model for the detection of sensitization to wheat allergens in bakery workers. The original model showed an adequate discrimination in the validation set, but the calibration was poor. Simple recalibration of the intercept improved the model's calibration. In its form as an easy-to-use scoring rule, the model enables manual calculation of the individual probability of being sensitized to wheat allergens. The model comprised of six questionnaire items: type of bakery, nasoconjunctival symptoms in the last 12 months, asthma problems in the last 12 months, shortness of breath and wheeze, during work upper and lower respiratory symptoms. Our findings support the idea that simple questionnaires can reasonably predict the presence of wheat sensitization.

We started by reducing 75 questionnaire items on symptoms to 12 candidate predictors. Modelling was based on 108 cases and 13 candidate predictors (ratio 8 to 1). We used bootstrapping to check whether the model produced optimistic (i.e. too high or too low) estimates in a new population. The model had a good internal validity with a shrinkage factor of 0.89.

It is generally acknowledged that differences between the population where a model is derived and where a model is applied may influence its performance and therefore, external validation is recommended to determine whether a model is applicable in new populations.(12, 25) The calibration plot in Figure 2a demonstrates that when the diagnostic model was applied without any adjustment in the validation set, it systematically produced

too low predicted probabilities (21), although the model could reasonably discriminate between the sensitized and not-sensitized workers (ROC area 0.69). Recalibration of the intercept alone already improved the calibration of the original model in the validation set. The more advanced update method did not improve the calibration (U statistic from method 1 vs. 2 were 0.0004 vs. -0.004, respectively), whereas the model's discrimination was comparable. Some of the re-estimated regression coefficients obviously deviated from the original values. Work-related upper respiratory symptoms showed a three times higher coefficient, whereas work-related lower respiratory symptoms showed a negative coefficient (which is not considered plausible). One possible explanation was that different associations were found in the validation set (results not shown). Work-related lower respiratory symptom was somehow not a strong predictor (OR 1.8, 95%CI 1.0 to 3.1) in the validation set. Therefore, the model with the re-estimated intercept was chosen as the final model and transformed to a scoring rule to facilitate its application in practice.

Sensitization to IgE to wheat allergens is closely associated with the development of respiratory diseases among bakery workers.(30) The diagnosis of whatever allergic disease (especially bronchial asthma) can only be made at an individual level in a clinical setting. However, at a population level, precursors of the disease of interest can be used for the detection of workers with an elevated risk of the outcome of interest. Subsequent diagnostic investigations can be limited to these high risk workers, leaving a considerable amount of workers in which no further medical investigations are needed. Therefore, instead of using a clinical definition such as OA, we used class II positive IgE serology to wheat allergens as our reference standard.

For decision making purpose, the choice of the cut-off point is an important issue. It must be based on the balance between the acceptable proportion of missed cases and the reduction of unnecessary tests. In general, a higher cut-off point leads to fewer subjects in the high probability group; the specificity is higher but at the cost of lower sensitivity. The context in where the model is going to be applied also determines the choice of the cut-off point. In the case of (occupational) allergic disease, false positive cases will 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.

As an illustration of how the model can formalize decision making and improve the efficiency of an early detection program, we stratified the workers into three probability groups. In workers with a high probability of wheat sensitization, advanced medical investigations are needed and may be best performed in an occupational respiratory health clinic. Workers with an intermediate probability can be evaluated further by occupational physicians. No further medical investigations are advised for workers with a low probability; they are enrolled in the next round of detection program. If a sum score of 4 or higher was proposed as a high cut-off point, 211 (16.8%) workers would be advised for additional clinical investigations. Only 77 (36.5%) of them eventually are sensitized to wheat allergens. Nonetheless, this group showed very high proportions of allergic and bronchial hyperresponsiveness symptoms, on average two and ten times higher than what was found in the intermediate and low probability group, respectively. This group also showed much higher rates of medication use and change in function due to respiratory or allergic symptoms than the other groups. In the intermediate group, the figures were lower than the high probability group, but definitely higher than the low probability group. Although we would miss 25 sensitized workers in the low probability group, but we correctly withhold serological tests in 470 workers. These findings suggested that the choice of the score cut-off points was reasonable.

In conclusion, we developed an easy to use and accurate diagnostic model for sensitization to wheat allergens to be used in future bakers. The model uses simple questionnaire items, which are commonly available in occupational health practice. With this approach the efficiency of health monitoring programs in baking industry may be increased.

External validation and updating in a sample of independent bakery workers yielded a model that is ready to be used in a wider population of bakery workers.

## REFERENCES

1. Brant A. Baker's asthma. *Curr Opin Allergy Clin Immunol* 2007;7(2):152-5.
2. Houba R, Doekes G, Heederik D. Occupational respiratory allergy in bakery workers: a review of the literature. *Am J Ind Med* 1998;34(6):529-46.
3. Nicholson PJ, Cullinan P, Taylor AJ, Burge PS, Boyle C. Evidence based guidelines for the prevention, identification, and management of occupational asthma. *Occup Environ Med* 2005;62(5):290-9.
4. Houba R, Heederik D, Doekes G. Wheat sensitization and work-related symptoms in the baking industry are preventable. An epidemiologic study. *Am J Respir Crit Care Med* 1998;158(5 Pt 1):1499-503.
5. Gautrin D, Ghezzi H, Infante-Rivard C, Malo JL. Incidence and determinants of IgE-mediated sensitization in apprentices. A prospective study. *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1222-8.
6. Heederik D, Houba R. An exploratory quantitative risk assessment for high molecular weight sensitizers: wheat flour. *Ann Occup Hyg* 2001; 45(3):175-185.
7. Tarlo SM, Liss GM. Prevention of occupational asthma--practical implications for occupational physicians. *Occup Med (Lond)* 2005;55(8):588-94.
8. Suarathana 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.
9. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19(4):453-73.
10. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130(6):515-24.
11. Janssen K, Moons K, Kalkman C, Grobbee D, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008;61(1):76-86.
12. 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.
13. Jacobs J, Meijster T, Meijer E, Suarathana E, Heederik D. Wheat allergen exposure and the prevalence of work-related sensitization and allergy. Accepted for publication in *Allergy, the European Journal of Allergy and Clinical Immunology* 2008.
14. Oostenbrink JH TJ, Tempels Z, Heide S, Steketee HA, Kerkhof M, Monchy JGR. Aard en omvang van beroepsgebonden klachten bij werknemers in bakkerijen, meelfabrieken en grondstoffenindustrie. Groningen: Academisch Ziekenhuis Groningen; 2002 24 Dec 2002.
15. 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.
16. van der Lende R, Orié NG. The MRC-ECCS questionnaire on respiratory symptoms (use in epidemiology). *Scand J Respir Dis* 1972;53(4):218-26.
17. Doekes G, Douwes J, Wouters I, de Wind S, Houba R, Hollander A. Enzyme immunoassays for total and allergen specific IgE in population studies. *Occup Environ Med* 1996;53(1):63-70.
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. Bollen K. *Structural equations with latent variables*. New York: John Wiley & Sons; 1989.
20. Nieuwenhuijsen MJ, Putcha V, Gordon S, Heederik D, Venables KM, Cullinan P, et al. Exposure-response relations among laboratory animal workers exposed to rats. *Occup Environ Med* 2003;60(2):104-108.
21. 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.
22. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;21(1):45-56.

23. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1):29-36.
24. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med* Nov 1990;9(11):1303-25.
25. Toll D, Janssen K, Vergouwe Y, Moons K. Validation, updating and impact of clinical prediction rules: a review. *Journal of Clinical Epidemiology* 2007; Accepted for publication.
26. Miller M, Langefeld C, Tierney W. Validation of probabilistic predictions. *Med Decis Making* 1993;13:49-58.
27. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;3(2):143-52.
28. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59(10):1087-91.
29. van der Heijden GJMG, T. Donders AR, Stijnen T, Moons KGM. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. *Journal of Clinical Epidemiology* 2006;59(10):1102-1109.
30. Chan-Yeung M MJ. Natural history of occupational asthma. In: Bernstein D C-YM, Malo JL, Bernstein DI, editor. *Asthma in the Workplace*. 2 ed. New York: Marcel Dekker; 1999. p. 129-144.

## APPENDIX

Clusters of correlating symptoms and percentage of explained variance as determined by principal component analysis (factor loadings after orthogonal varimax rotation between brackets)

A principal components analysis (PCA) was used to reduce the number of questionnaire predictors to be included in the multivariable logistic regression analysis. From the PCA, we identified clusters of correlated symptoms with Eigenvalue of one or greater (ref. 19). The identified clusters were named *post hoc* as adequately as possible. A positive cluster was defined as the presence of at least one of the symptoms that composed the cluster. For symptoms during and after work, we used additional condition if the symptoms improved when away from work.

---

### 1. Respiratory symptoms (cumulative explained variance 65.1%)

---

1.1. Asthma symptoms in the last 12 months  
(37.7%)

Have you had asthma attacks?  
(0.88)

Have you ever awoken due to asthma attack?  
(0.84)

Have you had asthma attack induced by  
exercise? (0.88)

1.2. Shortness of breath and wheeze  
(15.9%)

Have you ever had respiratory problems?  
(0.80)

Have you had shortness of breath  
after walking in normal pace? (0.60)

Have you ever had wheeze in the chest  
in the last 12 months? (0.77)

1.3. Cough and phlegm in the last 12 months  
(10.9%)

Have you had chronic cough (3 months)?  
(0.83)

Have you chronic cough with phlegm  
(3 months)? (0.81)

---

### 2. Upper respiratory symptoms (cumulative explained variance 50.5%)

---

Have you ever had allergic symptoms, including hay-fever? (0.71)

Have you had blocked/runny nose or sneezing in 12 months? (0.71)

Have you ever awoken duet o nasal symptoms? (0.67)

Have you had itchy/watery/red eyes in 12 months? (0.75)

---

### 3. Skin symptoms (cumulative explained variance 59.6%)

---

Have you ever had dry skin in the last 12 months? (0.70)

Have you ever had itchy skin in the last 12 months? (0.78)

Have you ever had eczema in elbow, knee, or neck? (0.81)

Have you ever had eczema in the last 12 months? (0.79)

---

---

**4. Allergic symptoms (cumulative explained variance 59.0%)**

---

**4.1. In contact with HDM and/or plants (35.2%)**

Respiratory symptoms in contact with house dust mites? (0.72)  
Respiratory symptoms in contact with plants or pollens? (0.73)  
Skin symptoms in contact with house dust mites? (0.68)  
Skin symptoms in contact with plants or pollens? (0.69)

**4.2. In contact with pets (12.6%)**

Respiratory symptoms in contact with pets? (0.87)  
Skin symptoms in contact with pets? (0.89)

**4.3. To certain food (11.2%)**

Respiratory symptoms after ingestion certain food? (0.78)  
Skin symptoms after ingestion certain food? (0.83)

---

---

**5. Symptoms suggestive of BHR (cumulative explained variance 51.5%)**

---

**5.1. Induced by smoke (35.9%)**

By dust (0.75)  
By smoke (0.89)  
By tobacco smoke (0.83)

**5.2. Induced by change of temperature (15.6%)**

From warm to cold (0.70)  
From cold to warm (0.64)  
By fog (0.640)  
By freezing cold (0.74)

---

---

**6. During work symptoms (cumulative explained variance 51.3%)**

---

**6.1. Upper respiratory symptoms (38.3%)**

Itchy eyes (0.65)  
Runny nose (0.83)  
Blocked nose (0.81)

**6.2. Lower respiratory symptoms (13.0%)**

Asthma attacks (0.70)  
Wheezing (0.77)  
Shortness of breath (0.74)  
Chest tightness (0.74)

---

---

**7. After work symptoms (cumulative explained variance 52.5%)**

---

**7.1. Lower respiratory symptoms (39.5%)**

Asthma attack (0.69)  
Wheezing (0.78)  
Shortness of breath (0.74)  
Chest tightness (0.70)

**7.2. Upper respiratory symptoms (13.0%)**

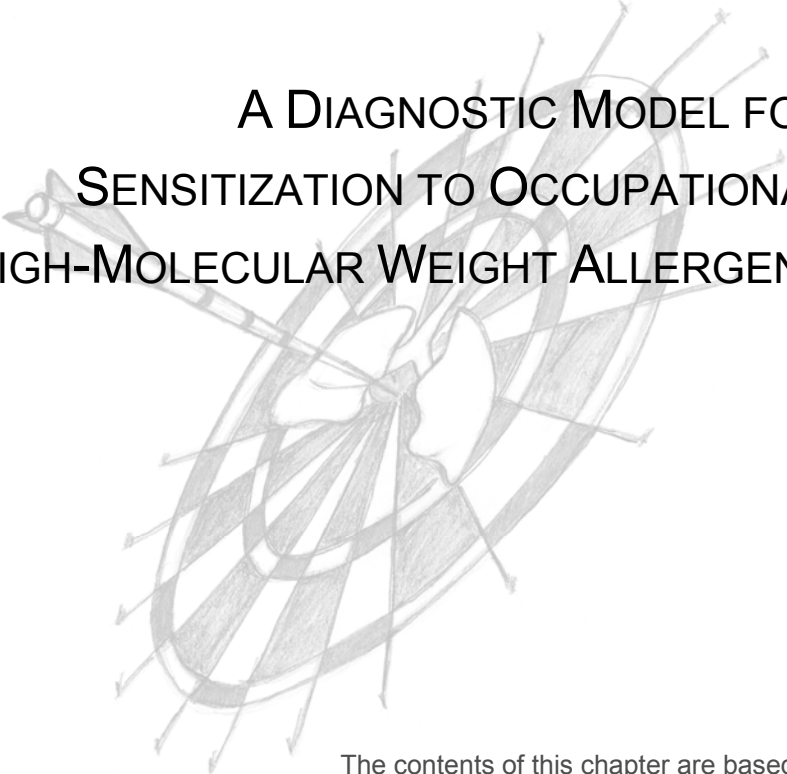
Itchy eyes (0.69)  
Runny nose (0.84)  
Blocked nose (0.81)

---



# CHAPTER 2.2

## A DIAGNOSTIC MODEL FOR SENSITIZATION TO OCCUPATIONAL HIGH-MOLECULAR WEIGHT ALLERGENS



The contents of this chapter are based on  
E. Suarhana, Y. Vergouwe, M. Nieuwenhuijsen,  
D. Heederik, D.E. Grobbee, E. Meijer  
Diagnostic model for sensitization in workers exposed to  
occupational high molecular weight allergens  
*Am J Ind Med* 2005;48(3):168-74



# A DIAGNOSTIC MODEL FOR SENSITIZATION TO OCCUPATIONAL HIGH-MOLECULAR WEIGHT ALLERGENS

---

EVA SUARTHANA, YVONNE VERGOUWE, MARK NIEUWENHUIJSEN,  
DICK HEEDERIK, DIEDERICK E. GROBBEE, EVERT MEIJER

---

## ABSTRACT

**Background:** Occupational allergy has great impact in workers exposed to high molecular weight (HMW) allergens. The present study aimed to develop and validate a generic diagnostic model for sensitization to HMW allergens, defined as positive IgE.

**Methods:** The model was developed in pooled data from Dutch laboratory animal (LA) workers and bakers using logistic regression analysis. Validity was assessed internally by bootstrapping procedure, and externally in British LA workers.

**Results:** The model included working hours/week, work-related symptoms, total IgE, and IgE to common allergen. Significant interactions between the type of work and the predictors resulted in different scores for LA workers and bakers. Internal and external validation showed that the model was satisfactory calibrated and discriminated workers at high and low risk of being sensitized.

**Conclusions:** It is possible to develop a generic model for sensitization to occupational HMW allergens. However, the weighing of predictors differs across specific work environments.

**KEYWORDS:** Diagnostic research, occupational allergy, bakers, laboratory animal workers, validity.

## INTRODUCTION

Occupational allergy has long been recognized as an important problem in workers exposed to high molecular weight (HMW) proteins that may cause allergic sensitization through IgE-mediated mechanisms. This particularly concerns rat urinary aeroallergen (RUA) and mouse urinary aeroallergen (MUA) exposure in laboratory animal workers, and wheat flour and fungal  $\alpha$ -amylase exposure in bakers (1-4).

The prevalence of occupational allergy has been reported to vary between 11-44 % in laboratory animal workers and 5-30 % in bakers (1, 5). Work related allergic symptoms, e.g. asthma, rhinitis, conjunctivitis, and/or eczema are frequently severe enough with social and economic consequences at individual level. A survey in the UK among patients diagnosed with asthma, showed that persons with occupational asthma had greater difficulty in finding new job, greater loss of income, and were less likely to be currently employed than those with asthma that was unrelated to work (1).

Exposure response studies show that, at the moment, sensitization can still occur when exposure is at the very lowest technologically achievable level (1). This stresses the importance of early detection of allergic sensitization in individual workers. However, standardized allergen preparations to demonstrate work-related sensitization can be costly and may not always be available. Therefore, a simple and cheap diagnostic strategy, e.g. a diagnostic decision model, is needed to identify workers at high risk of being sensitized to occupational HMW allergens.

We recently developed a strategy for health surveillance in laboratory animal workers exposed to RUA and MUA (6) from 351 Dutch laboratory animal workers participating in ongoing cohort study. However, it would be an advantage for occupational physicians if they have a generic diagnostic model for sensitization to HMW allergens, which they can apply in workers exposed to a broader spectrum HMW allergen. Therefore the current study aimed to develop and validate a single diagnostic model for sensitization to a broader spectrum occupational HMW allergens from questionnaires and laboratory tests commonly used in occupational health practice. We therefore developed a diagnostic model in pooled data set of Dutch laboratory animal workers and bakers and validated the model in British laboratory animal workers.

In contrast to the former study, we use positive IgE test (class I) to HMW allergen to diagnose sensitisation. This test was used rather than skin prick test (SPT), since within the European Union it is not allowed to prepare crude extracts of working place allergens to be used for SPT.

## METHODS

### Study Design and Population

The study population was drawn from laboratory animal facilities in the Netherlands and United Kingdom and bakeries in the Netherlands. Data of the Dutch bakers were derived from 21 Dutch bakeries(7). Data of Dutch laboratory animal workers were collected from a cohort of 4 universities, 3 commercial or industrial laboratories, and one laboratory school participated (5). Laboratory animal facilities in the United Kingdom involved 3 institutions (2 commercial and 1 academic) specialized in small animal research(8). The participation rates were 75%, 77% and 88%, which resulted in 427 Dutch bakers, 579 Dutch laboratory animal workers, and 357 British laboratory animal workers respectively.

The number of observations can differ from previously published analyses because only workers from whom questionnaires and specific IgE results were available were included.

The total study population contained 343 Dutch bakers, 568 Dutch laboratory animal workers and 352 British laboratory animal workers.

### **Questionnaire**

Workers completed a short self-administered respiratory questionnaire, which was extended with questions on personal history of allergic symptoms to common allergens, symptoms suggesting airways hyper-responsiveness (AHR), work-related allergic symptoms, smoking habits and job histories(5, 7).

A history of allergy ('Are you or have you been allergic to one or more agents?') was considered present if they reported history of chest tightness, running nose or sneezing, running or itching eyes, and itching skin to common allergens such as house dust, domestic animals, food or pollen. Persons were considered to have work-related allergic symptoms if they give positive answer to question 'Do you have symptoms of allergy (e.g. chest tightness, running nose or sneezing, running or itching eyes, and itching skin) during working hours, after contact with certain agent (e.g. flour in the bakery, rat in the laboratory) at work?'. Current smokers were defined as those who smoked more than one cigarette a day for at least one year. Data about AHR and personal history of respiratory symptoms from Dutch and British data sets could not be used because they were collected using different questionnaires.

### **Serologic Tests**

Specific IgE antibodies were measured in the blood serum. Specific IgE antibodies to wheat flour were measured with a commercial immunoassay (AlaSTAT; DPC, Apeldoorn, The Netherlands). Specific IgE antibodies to  $\alpha$ -amylase were measured with modified enzyme immunoassay (EIA) (9). Specific IgE antibodies to rat urinary aeroallergen (RUA) and mouse urinary allergen (MUA) were analyzed with CAP-RAST (10). Sera of class 1 or higher ( $> 0.35$  kU/L) were considered positive. Specific IgE antibodies to common allergens (house dust mites, grass pollen, birch pollen, cat fur, and dog fur) were measured with assay developed at the University of Wageningen, the Netherlands (9). An optical density (OD) of 492 exceeding the OD +0.05 of the reagent blank (no serum control) was considered as a positive reaction. Total IgE was measured with a sandwich EIA, a level  $\geq 100$  kU/L was defined positive (9).

### **Outcome Definition**

Sensitization to HMW allergens was defined as class I positive IgE serology to wheat flour or  $\alpha$ -amylase allergen in bakers, and RUA or MUA in laboratory animal workers.

### **Data Analysis**

A diagnostic model for HMW allergens sensitization was developed in the pooled data of Dutch bakers and laboratory animal workers ( $n=911$ , the development set). The univariable and multivariable associations between workers characteristics and presence of sensitization were modeled with logistic regression analysis. The regression coefficients are equal to the antilog of the odds ratios (OR).

We examined whether continuous variables could be transformed into simple categorizations. Cut-off values for the categories were chosen based on fitted restricted cubic spline (RCS) function (11). Model  $X^2$  of the models with the continuous and categorized variables were compared to see whether the categorizations were reasonable.

We first considered a full model including potential diagnostic characteristics from the questionnaire and serologic test with univariable p values below 0.5. Secondly, we used a backward stepwise procedure ( $p < 0.157$ ) to select a subset of the strongest predictors of

sensitization (12, 13). We also assessed whether the type of work (bakers or laboratory animal workers) modified the association of the selected predictors and sensitization using the interaction term. We performed an overall test to study the statistical significance of the interactions at  $p=0.3$ , conform current statistical guidelines(14).

Calibration of the model, the agreement between predicted probabilities of being sensitized and observed frequencies of sensitized workers, was assessed with the Hosmer–Lemeshow goodness-of-fit test. The discriminative ability was determined with the area under the receiver operating characteristic (ROC) curve. The ROC curve shows the relation between the false positive rate (1-specificity) against the true positive rate (sensitivity). The area under the curve (AUC) can range from 0.5 (no discrimination) to 1.0 (perfect discrimination) and reflects the probability that in all possible pairs of workers in which one worker is sensitized and one is not, a higher predicted probability is assigned to the worker who is sensitized(15).

We used bootstrapping to assess the internal validity of the diagnostic model. This procedure provides an estimate of the optimism-corrected ROC area. The corrected estimate is the value that can be expected in future patients. We also obtained a shrinkage factor from bootstrap procedure. The regression coefficients were multiplied by this shrinkage factor to prevent that predictions are too optimistic for new patients, i.e. that low predictions are too low and high predictions too high(16). External validity was studied in the British laboratory animal workers (validation set)(17, 18).

To facilitate application in practice, the regression coefficients of the diagnostic model were converted to a score chart. Then the sums of the scores were related to their corresponding probabilities. To derive the scores, the shrunk regression coefficients were divided by the smallest coefficient, and rounded to the nearest half integer. The sum scores were used to divide the population into a group of workers with a low and high risk of sensitization. Sensitivity and specificity of each score combination were plotted as ROC curves.

All statistical analyses were performed with SPSS 10.0 for Windows (Statistical Products and Service Solution, Inc, Chicago) and S-plus version 4.5 software (Mathsoft, Inc., Seattle, WA).

## RESULTS

### Worker Characteristics

The general characteristics and univariable associations with sensitization to HMW allergens of workers in the development set are shown in Table 1. Of 911 Dutch workers 650 (71%) were male and 146 (16%) were sensitized to occupational HMW allergens. The mean age was 35 (SD 9.8) years and the average of total working hours/week was 41 (SD 7.7).

### Model Development

The continuous variable total working hours/week was transformed into a dichotomous variable. The RCS function showed a clear change in risk of sensitization at 38 hours/week. The model  $X^2$  of the RCS and dichotomization at 38 hours/week yielded comparable results;  $X^2= 14.7$  for the continuous variable, and  $X^2= 11.8$  for dichotomous variable. Working more than 38 hours/week significantly increased the risk of sensitization by almost two times (OR=1.9) compared to working less than 38 hours (Table 1).

**Table 1** General characteristics and univariate association for sensitization to occupational HMW allergens in pooled data (the development set)

	Distribution n (%)	Univariable OR (95% CI)	p-value
<i>Questionnaire</i>			
Age †	35 (9.7)	0.99 (0.98 to 1.01)	0.508
Male gender	650 (71)	0.6 (0.4 to 0.9)	0.015
Current smoker	306 (34)	1.1 (0.7 to 1.6)	0.709
Personal history of atopic symptoms	301 (33)	1.9 (1.3 to 2.8)	< 0.001
Work related symptoms	294 (32)	5.5 (3.8 to 8.0)	< 0.001
Working > 38 hours/week	610 (67)	1.9 (1.3 to 2.7)	0.001
Previous exposure to HMW allergens	638 (70)	1.3 (0.9 to 1.9)	0.177
Type of work (laboratory animal worker)	568 (52)	1.0 (0.7 to 1.4)	0.996
<i>IgE results</i>			
Positive IgE to common allergens	287 (32)	4.4 (3.0 to 6.4)	< 0.001
Total IgE > 100 kU/l	199 (22)	3.4 (2.4 to 4.9)	< 0.001
<i>Outcome</i>			
Sensitized to HMW allergens	146 (16)		

†Mean (standard deviation)

OR: odds ratio; CI: confidence interval

We started the selection process with seven predictors, i.e. gender, personal history of allergic symptoms, work-related allergic symptoms, total working hours/week, previous exposure to HMW allergens, IgE to common allergen, and total IgE level. In agreement with previous publications age, and smoking habits were not associated with sensitization (univariable p-value > 0.5) and hence not considered in the backward selection (8, 10, 19, 20). Four predictors remained in the model after backward selection were work-related allergic symptoms, total working hours/week, IgE to common allergen, and total IgE level.

Testing the interaction between the first four predictors and type of work gave a significant result ( $X^2=11.6$ ,  $df=4$ ,  $p=0.021$ ). Therefore all interaction terms and type of work were included in the final model (Table 2).

The interaction terms show that the strength for predictors for laboratory animal workers and bakers are different. For example, the interaction of type of work with work related symptoms (OR=3.2) indicates that the strength of work related symptom is 3.2 times higher in laboratory animal workers than in bakers. Hence, the OR for work related symptoms is 6.7 (2.1\*3.2) in laboratory animal workers and 2.1 in bakers.

### Model Performance

The calibration of the diagnostic model is given in Figure 1. Observed frequencies and predicted probabilities of sensitization were in agreement for all predicted probabilities categories. The Hosmer-Lemeshow goodness-of-fit test showed good agreement between the predicted and observed probabilities of sensitization ( $p=0.2$ ).

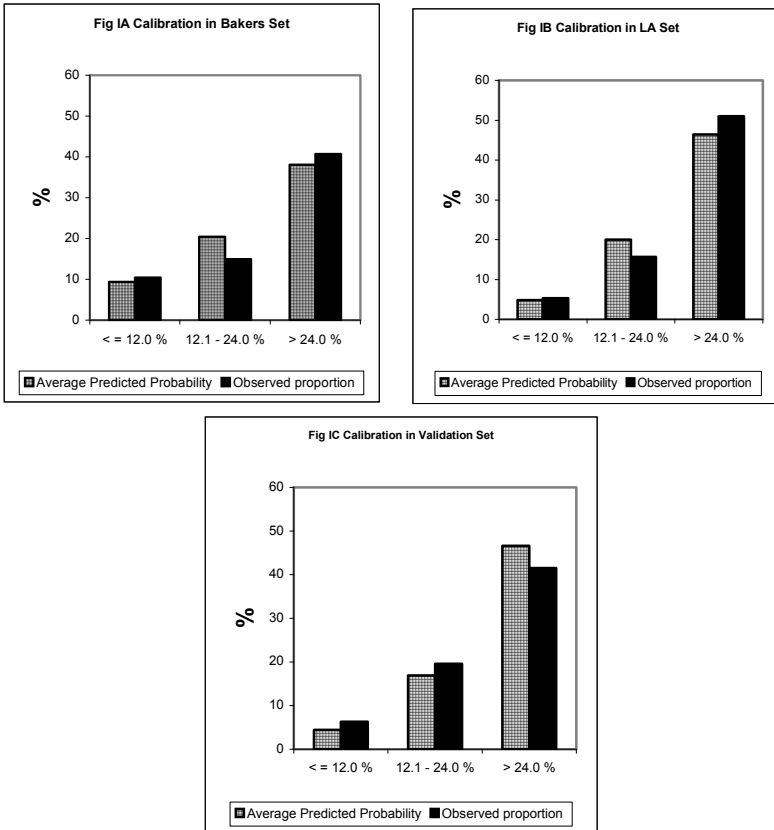
The discriminative ability of the diagnostic model in the development set was good (AUC 0.77; 95% CI 0.73 to 0.82). The AUC estimated in laboratory animal workers separately was 0.80 (95% CI 0.75 to 0.86) whereas in bakers it was 0.70 (95% CI 0.62 to 0.78). The model showed higher sensitivity with similar specificity in the laboratory animal workers, as compared to bakers, at all cut off points (Figure 2).

**Table 2** Strength of the predictors for sensitization to occupational HMW allergens in the model as estimated in pooled data (the development set)

	$\beta$ (SE)	OR (95% CI)
Work related symptoms	0.73 (0.33)	2.1 (1.1 to 3.9)
Working > 38 hours/week	0.87 (0.32)	2.4 (1.3 to 4.5)
Positive IgE to common allergens	0.82 (0.34)	2.3 (1.2 to 4.4)
Total IgE level > 100 kU/L	0.59 (0.35)	1.8 (0.9 to 3.6)
Type of work <sup>‡</sup>	-0.68 (0.50)	0.6 (0.2 to 1.3)
Work related symptoms * type of work	1.17 (0.43)	3.2 (1.4 to 7.4)
Working > 38 hours/week * type of work	-0.33 (0.48)	0.7 (0.3 to 1.8)
Positive IgE to common allergens * type of work	-0.43 (0.44)	0.6 (0.3 to 1.5)
Total IgE > 100 kU/l * type of work	0.67 (0.45)	1.9 (0.8 to 4.8)

$\beta$ : Regression coefficient ; SE: standard error; OR: odds ratio; CI: confidence interval

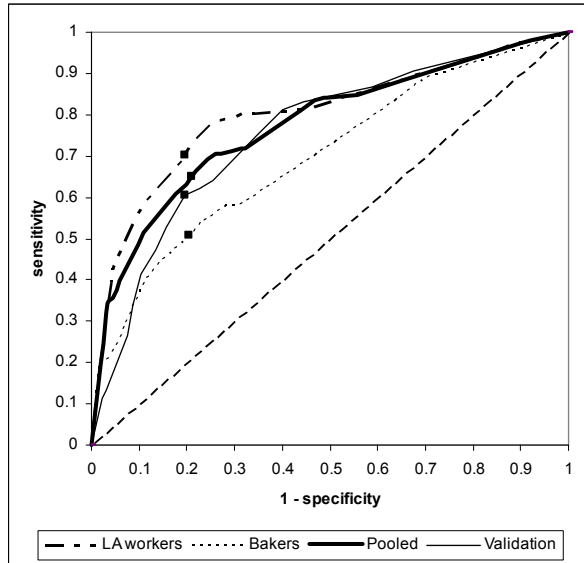
<sup>‡</sup>Type of work 1= laboratory animal workers, 0=bakers.



**Figure 1** Agreement between the observed proportion and the average predicted probability of sensitization to occupational allergens in Dutch bakers and laboratory animal workers (the development set) and British laboratory animal workers (validation set).

### Internal Validity

The bootstrapping procedure yielded an optimism-corrected AUC of 0.76 and a shrinkage factor of 0.95. This result showed that the optimism was small (AUC difference was 0.01, and the shrinkage factor of 0.95 was close to 1.0).



**Figure 2** Discriminative ability of the score chart. At the same sum score of 4.5 (♦), given similar specificity, the sensitivity was much lower in the bakers than in laboratory animal, whereas validation in British data showed in between sensitivity.

### Score Chart

The diagnostic model was presented as a score chart. The scores differed for bakers and laboratory animal workers as result of the interaction with type of work (Table 3). The sum scores correspond to predicted risks of being sensitized to occupational HMW allergens.

As shown, bakers and laboratory animal workers with two positive questionnaire predictors but negative serological test results had a sum score of 4.5 (2.0+2.5+0 for bakers and 5.0+1.5-2.0 for laboratory animal workers). This corresponded to a probability of 23% of being sensitized. At this point the sensitivity was 70% and specificity was 80% for laboratory animal workers, but 51% and 80% for bakers (Figure 2).

### External Validity

Of 352 British laboratory animal workers 53 (15%) were sensitized to HMW allergens. Workers in this validation set had shorter working hours/week (38.2 hours, SD 5.2), a lower proportion of workers had work-related allergic symptoms (24%), but a higher proportion had elevated total IgE levels (42%) than observed in the development set ( $p < 0.01$ ).

The calibration of the diagnostic model is shown in Figure 1C. The Hosmer-Lemeshow goodness-of-fit test showed good agreement between the predicted probabilities and observed proportion of sensitization ( $p = 0.786$ ). Figure 2 showed that the diagnostic model discriminated sensitized and non-sensitized workers reasonably well (AUC 0.76; 95% CI 0.68 to 0.83).

**Table 3** Diagnostic model for sensitization to occupational HMW allergens

Variable in the model	Value	Score	
		Bakers	LA workers
Work related symptoms	If exists	2.0	5.0
Total working hours/week	If > 38 hours/week	2.5	1.5
IgE to common allergens	If positive to $\geq 1$ common allergens	2.0	1.0
Total IgE level	If > 100 kU/L	1.5	3.5
Constant		0	-2.0
Sum score		...	...

Sum score	-2	-1	-0.5	0	0.5	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.5	6.0	6.5	7.5	8.0	9.0
Predicted probability of sensitization (%)	3	4	5	6	7	9	11	13	15	18	20	23	31	33	37	47	51	60

The exact formula to calculate the predicted probability is:

The sum product of shrunken regression coefficients and predictor values ( $LP_s$ ) =  $-2.82 + (0.83 \times \text{total working hours/week}) - (0.32 \times \text{total working hours/week} \times \text{type of work}) + (0.69 \times \text{work-related symptoms}) + (1.11 \times \text{work-related symptoms} \times \text{type of work}) + (0.77 \times \text{IgE to common allergen}) - (0.41 \times \text{IgE to common allergen} \times \text{type of work}) + (0.56 \times \text{total IgE level}) + (0.64 \times \text{total IgE level} \times \text{type of work}) - (0.65 \times \text{type of work})$ . With values equal to 1 if predictor is present, and type of work = 1 for laboratory animal workers.

The corresponding predicted probability of sensitization =  $1/(1+\exp(-LP_s))$ .

## DISCUSSION

Our study shows that, it is possible to develop a generic diagnostic model to detect the presence of sensitization to occupational HMW allergens in laboratory animal workers and bakers. The same variables predicted sensitization in the two types of workers with different weighing of the strength of the predictors for different work environments. The model showed adequate external validity in an independent set of laboratory animal workers.

The model was easily applicable since it comprised of four simple questionnaire and routine laboratory test predictors mostly available in occupational health practice. Our results showed that working > 38 hours/week, presence of work-related allergic symptoms, total IgE > 100 kU/L and positive IgE to at least one common allergen to be predictors of the presence of work-related sensitization, defined as class I positive specific IgE to occupational HMW allergens.

Interactions between type of work and work related symptoms, total working hours, and the serologic tests showed that the strength of these variables in predicting the sensitization differed between bakers and laboratory animal workers. For instance, total working hours appeared to be a stronger predictor in bakery setting ( $OR=2.4$ ) than in the laboratory animal setting ( $OR=2.4 \times 0.7=1.7$ ).

Several previous studies have shown that exposure level is a strong predictor of HMW allergens sensitization (1-4). We did, intentionally, not include exposure level in our diagnostic model as we aimed to develop model consisting of simple questionnaires and laboratory tests. Therefore the total working hours may be considered as a simple proxy for cumulative exposure level. Further analyses indeed showed that total working hours differed across type of bakeries and job titles (tasks) and correlated with cumulative allergen exposure.

Another issue related to our modelling strategy concerns the preference of dichotomized variable instead of continuous variable. Dichotomization will somehow

increase the possibility of misclassification for people who are just below or above the cut point. Though, we used dichotomized variable for reason of practical application. Further, the dichotomized variable fitted the data very similar to the RCS, because the increase in risk of sensitization was very steep at 38 hours.

Usually, sensitized workers are detected when they present themselves to the occupational physician because of symptoms. However, to find all sensitized workers the whole population must be studied in a survey like approach, which is less efficient and will result in high expenses.

In our approach, the diagnostic model serves as a triage system to split the population into high and low risk groups. The occupational physician could choose to follow only workers in the high risk group for sequential medical examination and occupational preventive intervention. In this way, the number of uninformative negative serological outcomes can be considerably limited.

An important step before using the diagnostic model, or score chart, is to determine the cut off point of the sum score. In our example, we used a sum score of 4.5. At this point the corresponding predicted probability was 23%, which is higher than the prior probability (16%). If we considered workers with a sum score of 4.5 or more at high risk, we could reduce the number of workers to be referred for further test to 27%. However, at this point, 4 of 10 workers with sensitization will be missed (sensitivity 64 %). With increasing cut off level, the number of workers undergoing further tests is reduced further but the number of false negative diagnosed workers will similarly increase.

Recently another diagnostic model was developed in only Dutch laboratory animal workers (6). The model consisted of only questionnaire variables, specifically for laboratory animal workers that do not apply to bakers. Consequently its use is limited to workers exposed to laboratory animal allergens.

In conclusion, regardless different weights of predictors across type of work, we can use the same variables to predict sensitization in bakers and laboratory animal workers. Our diagnostic model proved to be reliable and discriminated well between workers with high and low risk of sensitization of HMW allergens. Satisfactory validation in 352 British laboratory animal workers showed a good transportability in different geographic area.

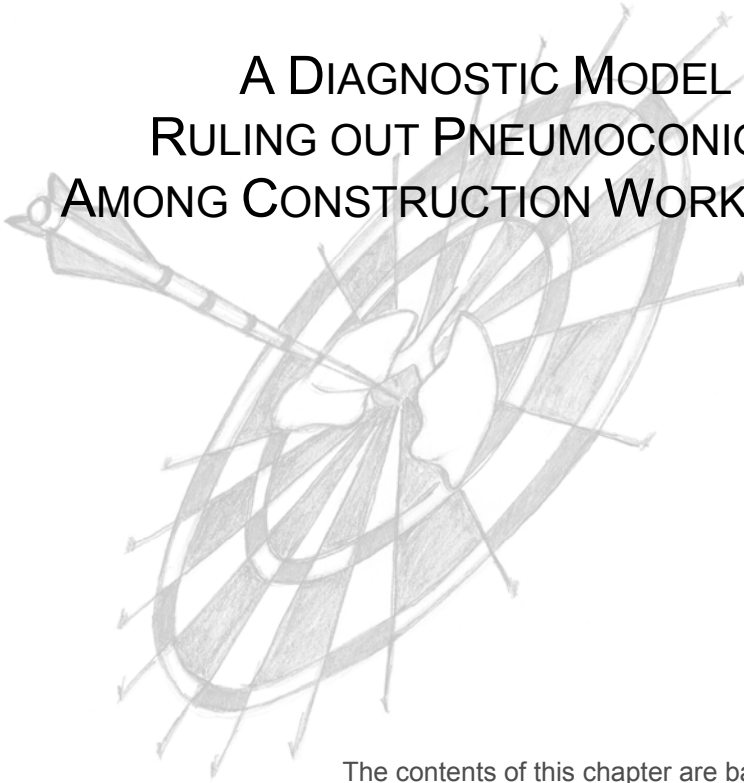
## REFERENCES

1. Houba R, Doekes G, Heederik D. Occupational respiratory allergy in bakery workers: a review of the literature. *Am J Ind Med* 1998;34(6):529-46.
2. Nieuwenhuijsen MJ, Heederik D, Doekes G, Venables KM, Newman Taylor AJ. Exposure-response relations of alpha-amylase sensitisation in British bakeries and flour mills. *Occup Environ Med* 1999;56(3):197-201.
3. Nieuwenhuijsen MJ, Putcha V, Gordon S, Heederik D, Venables KM, Cullinan P, et al. Exposure-response relations among laboratory animal workers exposed to rats. *Occup Environ Med* 2003;60(2):104-8.
4. Jeffrey P, Griffin P, Gibson M, Curran AD. Small bakeries--a cross-sectional study of respiratory symptoms, sensitization and dust exposure. *Occup Med (Lond)* 1999;49(4):237-41.
5. 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.
6. 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.
7. Houba R, Van Run P, Heederik D, Doekes G. Wheat antigen exposure assessment for epidemiological studies in bakeries using personal dust sampling and inhibition ELISA. *Clin Exp Allergy* 1996;26(2):154-63.

8. Cullinan P, Lowson D, Nieuwenhuijsen MJ, Gordon S, Tee RD, Venables KM, et al. Work related symptoms, sensitisation, and estimated exposure in workers not previously exposed to laboratory rats. *Occup Environ Med* 1994;51(9):589-92.
9. Doekes G, Douwes J, Wouters I, de Wind S, Houba R, Hollander A. Enzyme immunoassays for total and allergen specific IgE in population studies. *Occup Environ Med* 1996;53(1):63-70.
10. Heederik D, Venables KM, Malmberg P, Hollander A, Karlsson AS, Renstrom A, et al. Exposure-response relationships for work-related sensitization in workers exposed to rat urinary allergens: results from a pooled study. *J Allergy Clin Immunol* 1999;103(4):678-84.
11. 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.
12. Royston P, Sauerbrei W, Altman DG. Modeling the effects of continuous risk factors. *J Clin Epidemiol* 2000;53(2):219-21.
13. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000;19(8):1059-79.
14. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York; London: Springer; 2001.
15. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1):29-36.
16. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. *Stat Med* 2000;19(24):3401-15.
17. 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;56:826-832.
18. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130(6):515-24.
19. Kruize H, Post W, Heederik D, Martens B, Hollander A, van der Beek E. Respiratory allergy in laboratory animal workers: a retrospective cohort study using pre-employment screening data. *Occup Environ Med* 1997;54(11):830-5.
20. Cullinan P, Lowson D, Nieuwenhuijsen MJ, Sandiford C, Tee RD, Venables KM, et al. Work related symptoms, sensitisation, and estimated exposure in workers not previously exposed to flour. *Occup Environ Med* 1994;51(9):579-83.

# CHAPTER 2.3

## A DIAGNOSTIC MODEL FOR RULING OUT PNEUMOCONIOSIS AMONG CONSTRUCTION WORKERS



The contents of this chapter are based on  
E. Suarhana, KGM Moons, D. Heederik, E. Meijer  
A simple diagnostic model for ruling out pneumoconiosis  
among construction workers  
*Occup Env Med* 2007;64(9):595-601



# A DIAGNOSTIC MODEL FOR RULING OUT PNEUMOCONIOSIS AMONG CONSTRUCTION WORKERS

---

EVA SUARTHANA, KARL MOONS, DICK HEEDERIK, EVERT MEIJER

---

## ABSTRACT

**Background:** Construction workers exposed to silica-containing dust are at risk of developing silicosis even at low exposure levels. Health surveillance among these workers is commonly advised but the exact diagnostic work-up is not specified and therefore may result in abundant unnecessary chest X-ray investigations. We aimed to develop a simple diagnostic model to estimate for an individual worker the probability of having pneumoconiosis from questionnaire and spirometry results in order to accurately rule out workers without pneumoconiosis.

**Methods:** The study was performed using cross sectional data of 1291 Dutch natural stone and construction workers with potentially high quartz dust exposure. A multivariable logistic regression model was developed using chest X-ray with ILO profusion category  $\geq 1/1$  as the reference standard. The model's calibration was evaluated with the Hosmer-Lemeshow (HL) test; the discriminative ability was determined by calculating the area under the receiver operating characteristic curve (ROC area). Internal validity of the final model was assessed by a bootstrapping procedure. For clinical application, the diagnostic model was transformed into an easy-to-use score chart.

**Results:** Age 40 years or older, current smoker, high exposed job, working 15 years or longer in the construction industry, 'feeling unhealthy', and FEV1 were the independent predictors in the diagnostic model. The model showed good calibration (a non-significant HL-test) and discriminative ability (ROC area 0.81, 95% CI 0.74-0.85). Internal validity was reasonable; the optimism corrected ROC area was 0.76. By using a cut-off point with a high negative predictive value the occupational physician can efficiently detect a large proportion of workers with a low probability of having pneumoconiosis and exclude them from unnecessary X-ray investigations.

**Conclusions:** Our diagnostic model is an efficient and effective instrument to rule out pneumoconiosis among construction workers. Its use in health surveillance among these workers can reduce the number of redundant X-ray investigations.

**KEYWORDS:** construction workers, diagnostic model, health surveillance, silicosis

## INTRODUCTION

Silicosis is an interstitial lung disease caused by inhaled crystalline silica that is incurable and may be progressive even after the exposure has ceased.(1) Chronic silicosis is the most common, and typically occurs after 10 years of exposure to relatively low levels of silica. Decrements in lung function or respiratory symptoms are not likely in the early stages of simple silicosis.(2) In the more advanced cases, both obstructive and restrictive lung function effects, as well as decreased diffusion capacity, are more common. The International Agency for Research on Cancer (IARC) has classified silica as a Class I human lung carcinogen.(3) Cancer mortality risk varies, but appears to be highest in smoking workers with silicosis.(4) Chest radiography is the diagnostic investigation in which silicosis is presented with small rounded opacities in the upper and mid zones of the lung. In the construction industry, where quartz exposure arises from drilling, milling, grinding and demolition work, silicosis is often unrecognized.(5)

Recent research indicates that silicosis of grade 1/0 and 1/1 will still occur under current dust standards. Even at the proposed level of  $0.05 \text{ mg/m}^3$  the incidence rate of silicosis 1/1 (small rounded opacities) or greater would be about 10%-20%. It has been calculated that a permissible exposure level of less than  $0.001 \text{ mg/m}^3$  may be required to prevent mild radiographic changes (ILO profusion of  $\leq 1/0$ ) after a lifetime of silica-exposure.(2) Therefore, health monitoring to detect early signs of disease among workers exposed to silica-containing dust is needed.

Guidelines for silicosis surveillance have been published and in general involve questionnaires, physical examination, and additional tests like spirometry and chest X-ray.(6-9) The latter test is obviously more burdensome and costly to execute. The questionnaires are simple to apply and may include useful information for selection of individual workers that are at higher risk of having silicosis and therefore could be included in health surveillance programs. Nevertheless, none of the available protocols for silicosis surveillance specifies the exact diagnostic work-up for workers suspected of having silicosis. For instance, it is widely agreed that questionnaires on occupational and medical (respiratory) history should be collected before one can continue to the next step in the diagnostic work-up. However, it remains unclear which question provides the best diagnostic information on the presence or absence of silicosis. Further, it is indefinite who needs to undergo or who can be excluded from further evaluations.

Multivariable diagnostic prediction models, which consist of simple questions or tests, are often used to rule in or out a certain target condition to avoid unnecessary burdening and costly procedures.(6) We aimed to develop such simple diagnostic model for ruling out pneumoconiosis among construction workers at risk of having silicosis. We therefore used ILO profusion category  $\geq 1/1$  as the reference test. Multivariable logistic regression modelling was used to quantify the independent contribution of different questions and diagnostic tests.(7) Figure 1 illustrates how such model may be applied as a screening instrument in silicosis surveillance. It would enable the occupational physicians to discriminate workers with low probability of pneumoconiosis who do not need further medical investigations. This would obviously increase the efficiency of the surveillance, as it decreases the number of unnecessary chest X-rays.

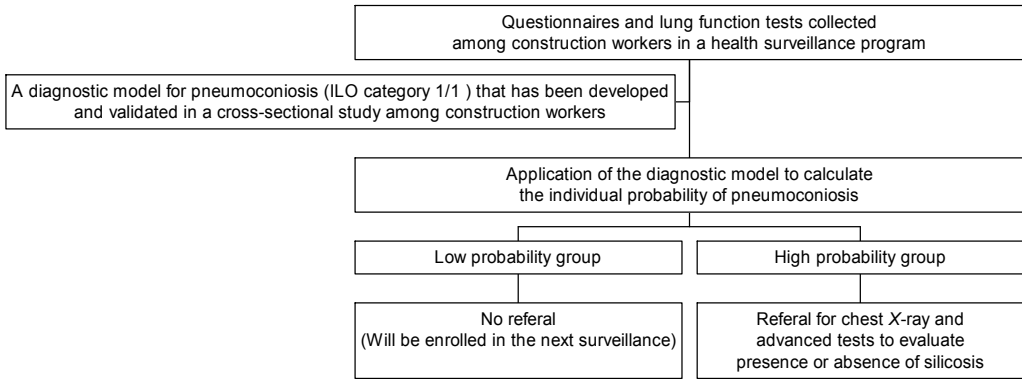


Figure 1 An approach for silicosis surveillance

## METHODS

### Study design and population

We used data from a cross-sectional study among Dutch natural stone and construction workers, 30 years and older.(8) In 1998, 4,173 workers with expected high cumulative exposure to quartz- containing dust registered in the natural stone association and nationwide construction workers database were invited to participate in the study. A questionnaire and invitation to the medical evaluation were sent to the eligible responders' home (n=1,690); and 1,335 (32% of the invited persons) were enrolled. Examinations took place in five locations distributed over the country and consisted of a questionnaire on respiratory symptoms and work history, lung function measurements, and chest X-ray (reference standard). All participating workers signed an informed consent for use of the test results for scientific research. The medical ethical committee of the university approved the study and all procedures were in agreement with European legal requirements with regards to privacy, data storage, and use of X-ray equipment.

### Questionnaire

The prevalence of respiratory symptoms was ascertained with a self-administered questionnaire derived from the British Medical Research Council Respiratory Questionnaire.(9) Respiratory symptoms were considered present if during the preceding 2 years the symptoms lasted for at least three months. Chronic cough was interpreted as either productive or non-productive cough. Shortness of breath was defined as ever being short of breath when walking with people of the same age at normal pace on level ground. Frequent wheezing was defined as wheezing for more than 1 week in the preceding 2 years. We added questions on whether the participants ever had, or have been told that they had, certain respiratory diseases (such as emphysema, pleuritis or tuberculosis), as well as questions on smoking and occupational history. A self-rated health question, "How would you assess your recent health condition?" with the answer choices healthy and unhealthy, was also included.

### Silica exposure assessment

A cumulative exposure index was available for every worker.(8) This exposure proxy was a semi-quantitative measure of the cumulative exposure to silica, which was calculated by multiplying the duration of exposure by an expert exposure index. Three industrial

hygienists, with experience in exposure assessment among construction workers, classified 36 different jobs on a 10 point scale for quartz exposure. The median score of the three experts, weighted for all consecutive and multiple jobs, was used to rank the different past and present occupations of the construction under study. The duration of exposure was calculated by summing up the years worked in jobs with potential mineral dust exposure in the construction industry.

### **Lung function measurements**

Lung function was measured with a pneumotachometer (Masterscreen Pneumo, Jaeger Benelux, Breda, The Netherlands) on the same day as the chest radiographs. The pneumotachometer measures the forced expiratory flow. Trained technicians performed the lung function measurements.(13)

Lung function data were compared with the European Respiratory Society reference values. To compare the actual to the reference lung function levels, we used the standardized residuals (standardized residual = (observed - predicted)/ residual standard deviation). This dimensionless index indicates how far the observed value is removed from the predicted value, and, therefore, how likely it is that the observed lung function occurs in the reference population.(10)

### **Diagnostic outcome (reference standard)**

Chest X-ray indicative for pneumoconiosis (ILO profusion category  $\geq 1/1$ ) was used as the reference standard. The chest radiograph has long been the cornerstone in the diagnosis of silicosis, and the ILO guidelines state that the classification system is to be used for epidemiological survey and routine surveillance of dust-exposed workers.(11)

Posterior-anterior chest X-rays from all individuals were taken in a mobile X-ray unit and read independently by three National Institute of Occupational Safety and Health (NIOSH, Morgantown, USA) "B" readers, according to the ILO guidelines for classification of pneumoconiosis.(11) Profusion score and the predominant shape of the opacities were recorded. Median results of the readings were used.

### **Data analysis**

We first assessed the univariable association between each predictor and the presence of the outcome with binary logistic regression analysis. We also examined whether continuous variables (age, lung function, and cumulative exposure index and total working years) could be transformed into simple categorizations. Cut-off values for categories were chosen based on restricted cubic spline (RCS) functions, provided in S-Plus version 2000 (Mathsoft, Inc., Seattle, WA), at the point where the function showed a change in risk of the outcome.(12) Model  $X^2$  of the models with the continuous and categorized variables were compared to see whether the categorizations were reasonable. To evaluate the association between job titles and the diagnostic outcome, we first assigned one job title for each worker. Workers who had more than 1 job title (n=256) were assigned to a job title with the highest expert exposure index.(8) As this resulted in thirty-six job titles, we further clustered the job titles based on the expert exposure index.

We finally evaluated the univariable association between the clustered job titles and outcome; those with similar regression coefficients were finally grouped into two groups (high versus low exposed) for convenience in practice. We then fitted a multivariable logistic regression model including all potential predictors from the questionnaire and spirometry, based on a univariable p value < 0.50. Secondly, we used a backward stepwise procedure (using p < 0.157 for inclusion) to select a final model with the strongest predictors for absence (or presence) of pneumoconiosis.(12) Extra analysis was performed to compare the diagnostic performance between the final model with all continuous variables preserved in their original form, and with these variables in dichotomized form.

The diagnostic accuracy of the final model was quantified using calibration and discrimination measures. Calibration, the agreement between the predicted probabilities and the observed frequencies of having abnormal chest X-ray indicative for pneumoconiosis, was assessed graphically and tested with the Hosmer–Lemeshow (HL) test (where p-value of 0.10 and higher reflects good agreement). The discriminative ability was determined with the area under the receiver operating characteristic (ROC) curve. The area under the ROC curve (ROC area) shows the relation between false positive rate (1 - specificity) and true positive rate (sensitivity). The ROC area can range from 0.5 (no discrimination) to 1.0 (perfect discrimination). The ROC area reflects the probability that for all possible pairs of workers, in which one worker has pneumoconiosis and one has not, the model indeed assigned a higher probability of having pneumoconiosis to the worker with pneumoconiosis.(13)

We used bootstrapping to assess the internal validity and the amount of over fitting of the model. This bootstrapping procedure gives a correction factor for both the model's ROC area and for the regression coefficients of the predictors in the final model.(14) The regression coefficients of the predictors in the final model were multiplied by this correction factor to prevent the model from yielding optimistic predictions when applied in future (new) workers.

To facilitate application of the final model in practice, the corrected regression coefficients of the predictors in the final diagnostic model were converted to easy-to-use numbers. To derive these numbers, the corrected coefficients were divided by the smallest one, and rounded to the nearest half integer. The discriminative accuracy of this so-developed scoring rule was again assessed. Finally, the sums of the scores were related to their corresponding probabilities of having (a chest X-ray indicative for) pneumoconiosis. Finally, a cut-off point of the sum scores was introduced to divide the worker population into group with low versus high probability of having (chest X-ray indicative for) pneumoconiosis.

Of 1335 available individual data, 44 workers with a missing outcome (chest X-ray result) or with a completely missing questionnaire were excluded, leaving n=1291. Of these, 58 participants (4.5%) had 77 missing values. Although participants with missing value showed similar characteristics as those who had complete values, deletion of subjects with a missing value (so called 'complete case analysis') may still lead to biased result and certainly a loss of power.(15) Therefore, we imputed the missing values by using the linear regression method (with addition of an error term) in SPSS 11.5 for Windows (Statistical Products and Service Solution, Inc, Chicago). The imputation was based on the correlation between each variable with missing values and all other variables as estimated from the 1233 (95.5%) complete dataset; yielding 1291 complete data for further analyses.

## RESULTS

### Population characteristics

ILO profusion categories  $\geq 0/1$ ,  $\geq 1/0$ , and  $\geq 1/1$  were found in 426 (33.0%), 131 (10.1%), and 37(2.9%) workers, respectively. Of 426 X-rays with profusion category  $\geq 0/1$ , we found 38 (8.9%) with primary and 45 (10.6%) with secondary rounded opacities. Table 1 shows that workers with ILO profusion category  $\geq 1/1$  were older, worked longer in the construction industry, had higher cumulative exposure index, and had worse lung function. This group also showed higher percentages of current smoker and workers who felt unhealthy.

**Table 1** General characteristics and univariable association with the outcome

	Workers with chest X-ray profusion ILO category < 1/1 n=1254	Workers with chest X-ray profusion ILO category ≥ 1/1 n=37	OR (95% CI)	p-value
<i>Questionnaire</i>				
Age (years)	41.3 (7.7)	46.1 (7.9)	1.1 (1.03 to 1.12)	< 0.001
Age ≥ 40 years	661 (52.7)	29 (78.4)	3.3 (1.5 to 7.2)	0.003
Male*	1253 (99.9)	37 (100.0)		
Ever smoker	992 (79.1)	33 (89.2)	2.2 (0.8 to 6.2)	0.145
Current smoker	617 (49.2)	25 (67.6)	2.2 (1.1 to 4.3)	0.031
Respiratory symptoms (cough, wheeze, shortness of breath)	456 (36.4)	17 (45.9)	1.5 (0.8 to 2.9)	0.236
Symptoms suggesting BHR	467 (37.2)	12 (32.4)	0.8 (0.4 to 1.6)	0.551
History of lung diseases (emphysema/ pleuritis/ TBC)	252 (20.1)	8 (21.6)	1.1 (0.5 to 2.4)	0.820
Medication for respiratory disease	48 (3.8)	4 (10.8)	0.7 (0.3 to 1.7)	0.457
"Feeling unhealthy"	135 (10.8)	10 (27.0)	3.1 (1.5 to 6.5)	0.003
Job title: low exposed group <sup>‡</sup>	419 (33.3)	5 (13.5)	reference	
high exposed group <sup>‡‡</sup>	836 (66.7)	32 (86.5)	3.2 (1.2 to 8.3)	0.016
Work duration (years)	18.0 (1 to 52)	25.0 (8 to 43)	1.1 (1.03 to 1.10)	< 0.001
Work duration ≥ 15 years	798 (63.6)	32 (86.5)	3.7 (1.4 to 9.5)	0.007
<i>Lung function result</i>				
Standardized residual FEV1	0.2 (-7.0 to 3.5)	-0.2 (-3.6 to 2.0)	0.7 (0.5 to 0.9)	0.003
Standardized residual FEV1 ≤ -1.0	171 (13.6)	12 (32.4)	3.0 (1.5 to 6.2)	0.002
<i>Exposure measurement</i>				
Cumulative exposure index	6.9 (0 to 59.4)	14.0 (1.6 to 32.8)	1.1 (1.0 to 1.1)	< 0.001
Cumulative exposure index ≥ 10.0	452 (36.0)	23 (62.2)	2.9 (1.5 to 5.7)	0.002

Data are presented as median (min to max) or absolute values (%); OR: odds ratio; CI: confidence interval; BHR: bronchial hyper-responsiveness.

\*OR could not be calculated because there was only 1 female worker.

<sup>‡</sup> No construction work, truck driver, production worker, welder, miner, mechanic, painter, crane driver, foundation worker, worker cleaning-up asbestos, gypsum brick layer, finishing mechanic, tuck pointer, carpenter, insulator, tiler, floorer, bricklayer, unskilled personnel, plasterer, bricklayer assistant, concrete worker, grinder-road construction worker, or railway-road construction worker.

<sup>‡‡</sup> Concrete repairman, concrete blaster, concrete driller and grinder, terrazzo worker, pile-top crusher, natural stone worker, recess miller, tuck pointer chasing out mortar between bricks, rubble cleaner, recess cutter or demolition worker.

The restricted cubic spline (RCS) functions clearly showed an increase in risk of having X-ray indicative for pneumoconiosis at age of 40 years or older, working in the construction industry 15 years or longer, a cumulative exposure index of 10 or higher, and a standardized residual FEV1  $\leq -1.0$ . The model  $X^2$  of the continuous versus dichotomized variables for age, working years, cumulative exposure index, and standardized residual FEV1 were 12.6 vs. 10.2, 14.1 vs. 9.5, 13.2 vs. 10.1, and 7.8 vs. 8.2, respectively. Working 15 years or longer in the construction industry significantly increased the probability of pneumoconiosis by almost four times (odds ratio, OR=3.7), and workers over 40 years or having a decreased lung function had triple the probability of pneumoconiosis (see Table 1).

### Model development

The initial multivariable model included seven questionnaire predictors (univariable p-values  $\leq 0.50$ ; Table 1). Of these, only five, i.e. age 40 years or older, current smoker, high exposed job title, working 15 years or longer in the construction industry, and 'feeling unhealthy' appeared independently ( $p < 0.157$ ) related to the presence or absence of pneumoconiosis (Table 2, first column). Table 2 also shows that the cumulative exposure index did not give additional value to the diagnostic information provided by the five questionnaire items; the odds ratio of the cumulative exposure index was far from significant and the ROC area of both models were the same (second column). Addition of the lung function to the reduced questionnaire model (third column) slightly increased the ROC area (delta ROC area of 0.02) and significantly increased the model  $X^2$  of the questionnaire model (deviance=7.8, df=1,  $p=0.005$ ). The calibration plot of the reduced questionnaire + lung function model showed good calibration, confirmed by a non-significant HL-test ( $p=0.20$ ). Therefore, the questionnaire + lung function model was chosen as the final model. The bootstrapping procedure yielded a correction factor of 0.82 for the regression coefficients of the final model, which indicated a reasonable internal validity. The corrected AUC was 0.76 (instead of 0.81).

**Table 2** The strength of the predictors for chest X-ray indicative for pneumoconiosis (ILO profusion category  $\geq 1/1$ )

Predictors	Questionnaire model	Questionnaire + Exposure model	Questionnaire + Lung function model	
	OR (95% CI)	OR (95% CI)	$\beta^\ddagger$	OR (95% CI)
Age $\geq 40$ years	2.3 (1.0 to 5.4)	2.3 (1.0 to 5.4)	0.72	2.3 (1.0 to 5.4)
Current smoker	2.5 (1.2 to 5.1)	2.5 (1.2 to 5.1)	0.70	2.4 (1.1 to 4.9)
High exposed job title $^{\ddagger\ddagger}$	4.1 (1.5 to 10.6)	3.9 (1.3 to 11.6)	1.14	4.0 (1.5 to 10.5)
Working $\geq 15$ years	3.3 (1.2 to 9.1)	3.2 (1.1 to 9.4)	1.00	3.4 (1.2 to 9.3)
'Feeling unhealthy'	2.6 (1.2 to 5.7)	2.6 (1.2 to 5.7)	0.84	2.8 (1.3 to 6.0)
Cumulative exposure index $\geq 10.0$	-	1.1 (0.5 to 2.4)	-	-
Standardized residual FEV1 $\leq -1.0$	-	-	0.91	3.0 (1.5 to 6.3)
ROC area of the model	0.79 (0.74 to 0.85)	0.79 (0.74 to 0.85)		0.81 (0.75 to 0.86)

OR: odds ratio; CI: confidence interval

$^\ddagger$  Regression coefficients after multiplication by the correction or shrinkage factor obtained from the bootstrapping procedure.

$^{\ddagger\ddagger}$  See the legend of the Table 1.

The probability of having a chest X-ray result that is indicative for the presence of pneumoconiosis defined as ILO profusion category  $\geq 1/1$  can be estimated using the following formula:  $P(\text{pneumoconiosis}) = 1/(1 + \exp(-((0.72 \times \text{age} \geq 40 \text{ years}) + (0.7 \times \text{current smoker}) + (1.14 \times \text{high exposed job title}) + (1.0 \times \text{work in the construction industry} \geq 15 \text{ years}) + (0.84 \times \text{'feeling unhealthy'}) + (0.91 \times \text{standardized residual FEV1} \leq -1.0) - 6.33)))$ . Each predictor is valued as 1 when present and 0 when absent.

There was no significant difference in the diagnostic performance between the model with all continuous variables preserved in their original form and the model with dichotomized variables. The model with continuous variables showed model  $X^2$  of 44.0 with 6 degrees of freedom, ROC area of 0.809, and HL-test  $p=0.514$ , whereas model with dichotomized variables showed model  $X^2$  of 44.9 with 6 degrees of freedom, ROC area of 0.805, and HL-test  $p=0.465$ .

### Score chart

After multiplication by the correction (shrinkage) factor obtained from the bootstrapping procedure, the corrected regression coefficients of the independent predictors of the final model were converted into a simple-to-use score system or score chart (Table 3). The predictive accuracy of this scoring system was good;  $p$ -value of the HL-test was 0.39 and the ROC area was 0.787 (95% CI 0.729 to 0.845). Using the adjusted regression coefficients from Table 2 (see formula legend Table 2), we calculated the predicted probabilities of pneumoconiosis corresponding to the different sum scores (Table 3, lower part). As an example how to use this chart, a 45 year old non smoking concrete driller who worked in the construction industry for 10 years, felt healthy, but had a standardized residual FEV1 of -1 would have a sum score of 3.75 (1+0+1.5+0+0+1.25). This corresponded to a probability of pneumoconiosis of 2%.

**Table 3** Diagnostic model for chest X-ray indicative for pneumoconiosis and the corresponding predicted probability

Variable in the model	Value	Score
Age	$\geq 40$ years	1.0
Smoking habit	Current smoker	1.0
High exposed job title	Concrete repairman, concrete blaster, concrete driller and grinder, terrazzo worker, pile-top crusher, natural stone worker, recess miller, tuck pointer chasing out mortar between bricks, rubble cleaner, recess cutter or demolition worker	1.5
Work duration in the construction industry	$\geq 15$ years	1.5
Self-rated health	'Feeling unhealthy'	1.25
Standardized residual FEV1	$\leq -1.0$	1.25
	Sum score	...
Sum score	< 3   3   3.75   4.0   4.25   4.75   5.25   6.25   7.5	
Predicted probability of outcome (%)	0   1   2   2.5   3   5   8   19   45	

Table 4 displays the diagnostic accuracy parameter plus the corresponding proportions of the detected pneumoconiosis cases and unnecessary referrals for each sum score threshold. For example, using the sum score threshold of 3.75 or higher as cut-off point for referral for chest X-ray, one refers 567 (43.9%) workers. As the aim of our model was to rule out pneumoconiosis, the negative predictive value was 99.4% (720/724) and the likelihood ratio of a negative test (LR-) was 0.2 ((4/37)/(720/1254)). For comparison, not using the model and referring all workers for chest X-ray would in fact result in 1254 unnecessary negative X-rays (as only 37 had a positive X-ray).

**Table 4** The diagnostic accuracy across different cut-off points for referral for chest X-ray investigation

Sum scores cut-off	Number of workers (%) <sup>*</sup> per sum score category	Number of workers with chest X-ray ILO profusion category $\geq$ 1/1 (n=37) n (%) <sup>†</sup>	Number of workers without chest X-ray ILO profusion category $\geq$ 1/1 (n=1254) n (%) <sup>‡</sup>	Sensitivity (%)	Specificity (%)	NPV (%)
$\geq$ 2	1065 (82.5)	37 (3.5)	1028 (96.5)	100.0	18.0	100.0
$>$ 3	684 (53.0)	37 (5.4)	647 (94.6)	100.0	48.4	100.0
$\geq$ 3.75	567 (43.9)	33 (5.8)	534 (94.2)	89.2	57.4	99.4
$>$ 4.0	494 (38.3)	31 (6.3)	463 (93.7)	83.8	63.1	99.2
$\geq$ 4.25	293 (22.7)	22 (7.5)	271 (92.5)	59.5	78.4	98.5
$>$ 4.75	270 (20.9)	21 (7.8)	249 (92.2)	56.8	80.1	98.4
$\geq$ 5.25	119 (9.2)	13 (10.9)	106 (89.1)	35.1	91.5	98.0

<sup>\*</sup> Proportion of all workers (n=1291)

<sup>†</sup> Proportion of workers with positive X-ray within the sum score category

<sup>‡</sup> Proportion of workers with negative X-ray within the sum score category

## DISCUSSION

Our study shows that the diagnostic model for pneumoconiosis ILO profusion category  $\geq$  1/1 effectively rules out pneumoconiosis and reduces a substantial number of unnecessary referrals for chest X-ray investigations.

The main motivation for (multivariable) diagnostic research is to determine whether simple diagnostic tests already predict the presence or absence of the target disease without having to perform the more invasive and costly reference test, with acceptable misclassifications. The motive of diagnostic research is simply to decrease patient burden and health care expenses and by no means to explain causality.(7, 16) In our study, we focused on optimal prediction of the absence of pneumoconiosis in order to decrease unnecessary X-ray referrals.

Diagnostic studies are inherently cross sectional, and the test results under study are commonly not causal factors (i.e. not part of the causal pathway). In fact, most test results are actually the consequence of disease presence. This explains the inclusion of the variables “feeling unhealthy” and “current smoker” in our final diagnostic model. These associations thus do not necessarily express a causal relation between smoking and pneumoconiosis. We only used the information carried by the smoking habit to estimate the probability of having a positive X-ray. The OR of 2.4 for current smoker only means that current smokers, compared to non-smokers, have 2.4 times higher probability of having a positive X-ray, without any reference to causality. The OR for “feeling unhealthy” should also be interpreted as: workers who feel unhealthy have 2.8 times higher probability of having a positive X-ray than workers who feel healthy. Furthermore, “feeling unhealthy” as a global assessment of an individual’s health perspective, although not disease specific, has also been proven to be an independent and strong predictor of mortality in community studies.(17)

The selected predictors in our diagnostic model, including age and declining lung function have been mentioned in previous studies.(5, 18-21) Here again, age is not a causal factor, but rather captures information about duration of exposure and potential progression of disease and thus has independently contributed to the prediction of pneumoconiosis. Some intuitive predictors such as work duration and job with high exposure to silica containing dust were selected in our model as well. The cumulative exposure index indeed showed a significant univariable association with a chest X-ray result that was indicative for pneumoconiosis. However, adding this variable to the multivariable model that already

included work duration and job title, did not improve the predictive accuracy of the model. Apparently, its information (that was reflected in the univariable analysis) was already provided by work duration and job title. Several exposure studies indeed demonstrated that the number of years of working in the construction industry could be used as a surrogate measure for cumulative exposure to silica.(4, 25) The high exposed job consisted of job titles with high quartz exposure level and, therefore, could be considered a proxy for exposure as well.(22)

We chose ILO profusion category  $\geq 1/1$ , as an abnormal chest X-ray indicative for pneumoconiosis, as our reference standard. Choosing profusion category  $\geq 1/0$  would lead to defining a category in which the absence of small opacities was seriously considered. This would result in significantly more misclassification. Plain chest radiographs may be insensitive to early changes of the lung parenchyma. High-resolution computed tomography (HRCT) is superior to chest X-ray in identifying early parenchymal lesions.(23) Nevertheless, codifying abnormalities according to ILO classification by conventional radiographic assessment remains the first diagnostic step in epidemiological studies and in health surveillance programs of dust exposed workers.(11) Besides, the increased radiation dose, as well as the added expense and time involved, do not justify the use of HRCT for surveillance.

Silicosis occurrence is not surprising among workers exposed to silica-containing dusts in various industries and occupations. Patients may be free of symptoms with abnormalities identified by chest X-ray during medical screening. However, regarding the low prevalence, it is not efficient to perform a chest X-ray as a routine test in the population at risk since this will yield many avoidable negative outcomes.

Our diagnostic prediction model can be applied by the occupational physician to decide whether a worker should have a chest X-ray investigation or not. For groups of workers with low probability, there will be no further action. They will be enrolled in the next surveillance round. For groups of workers with a high probability, they should be referred for chest X-ray and possibly advanced medical tests (i.e. diffusion capacity and HRCT) to confirm the presence or absence of pneumoconiosis or silicosis.

The diagnostic model showed good diagnostic accuracy (calibration, discrimination, and internal validity). However, to use it as a screening tool, one should carefully choose the cut-off point above which workers should be referred for chest X-ray. Screening must lead to a high level of case detection (high sensitivity), and at the same time a reasonably low level of unnecessary X-ray referrals. So, the choice of a cut-off point must be based on an acceptable proportion of missed cases and of unnecessary referrals. As shown in Table 4, a higher cut-off leads to a lower sensitivity but at the gain of investigating less workers. Policy makers should, therefore, balance between the number of missed cases and cost reduction gained by minimizing the number of referrals for advanced diagnostic tests. For example, not referring workers with sum scores lower than 4.75 will save the expenses for unnecessary chest X-rays up to 80% of the subjects. This cut-off point also has a high negative predictive value of 98.4%, which means that more than 98% of those who are not referred would indeed have a negative chest X-ray (if tested). However, 16 of all 37 cases will be missed. Given the slowly progressive nature of the disease and the fact that surveillance is repeatedly conducted overtime, we could expect that the missed cases would be captured in the next surveillance. However, if the aim is to detect as many cases as possible, this cut-off value will not be the first choice. One then might use a lower cut-off point to reduce the number of missed cases. For instance, a sum scores of 3.0 or higher yields 100% sensitivity, which means that all cases will be captured, but at the expense of referring half of all workers. A cut-off point in between the earlier discussed options is 3.75, with 89.2% sensitivity, 99.4% negative predictive value, and not referring 56.1% of all workers. Doing so, the use of this diagnostic model as an initial screening instrument, will surely increase the efficiency of health surveillance in construction workers.

The small number of cases relative to the high number of potential predictors studied is the limitation of our analysis. For diagnostic studies, no exact formula for sample size calculation exists. However, the general rule is that per candidate predictor variable there should be at least 10 events (1 to 10 rule).(24) We had 9 candidate predictors with 37 cases (ratio 1 to 4). When limited positive cases are available to develop a model, 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 than split-sample or cross validation methods.(28) It turned out that the model had a reasonable internal validity (we obtained a correction factor of 0.82; the closer the correction factor is to 1, the less optimism). Nevertheless, an external validation in a new population is required to confirm the performance of the model and its transportability into all construction workers. Another important point from the modelling aspect is the dichotomization of various continuous variables. Dichotomization increases the potential for misclassification and we may lose important information conveyed by a variable.(29) Nevertheless we dichotomized those variables for reason of simplicity in practice. The cut-off values were chosen based on restricted cubic spline function (RCS), at the point where the function showed an observable change in risk of the outcome.(12) Additional analysis showed that the model  $X^2$  of both forms were comparable and there was no significant difference in the diagnostic performance between the model with all continuous variables preserved in their original form and the model with dichotomized variables.

In conclusion, we derived a diagnostic model for pneumoconiosis that can be applied for health surveillance on a large scale in natural stone and construction workers. The model comprises simple questionnaire items and routine lung function, which are widely available in occupational health settings. With our approach, the efficiency of the health surveillance can be increased considerably by decreasing a large number of unnecessary referrals for chest X-ray investigations. Yet, external validation of the model is recommended before it can be used with confidence in all construction workers

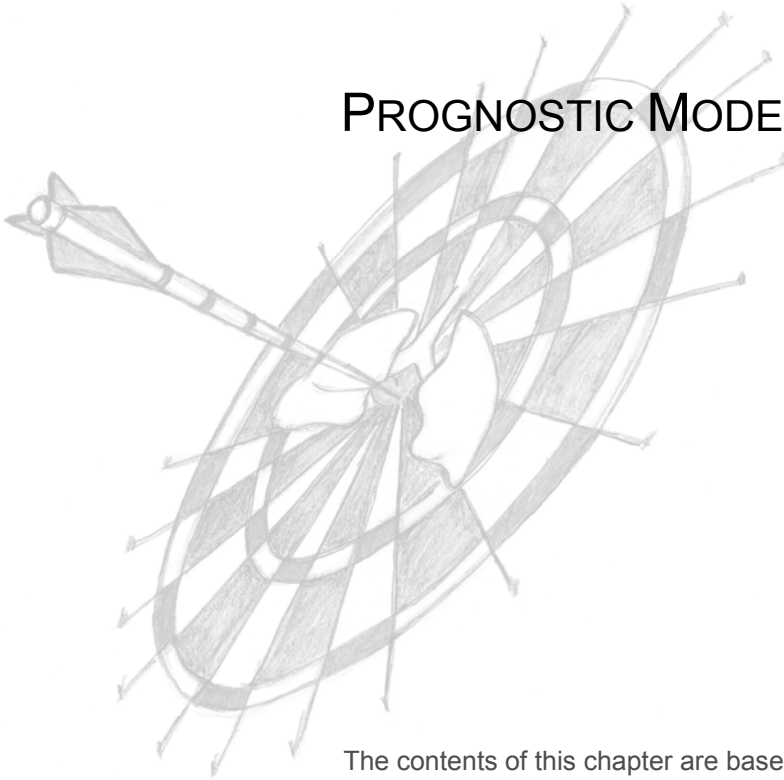
## REFERENCES

1. WHO. WHO fact sheet no 238. 2000.
2. Greaves IA. Not-so-simple silicosis: a case for public health action. *Am J Ind Med* 2000;37(3):245-51.
3. IARC. IARC Summary & Evaluation; 1997.
4. t Mannelje A, Steenland K, Attfield M, Boffetta P, Checkoway H, DeKlerk N, et al. Exposure-response analysis and risk assessment for silica and silicosis mortality in a pooled analysis of six cohorts. *Occup Environ Med* 2002;59(11):723-8.
5. Wagner GR. Asbestosis and silicosis. *Lancet* 1997;349(9061):1311-5.
6. 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.
7. Moons KG, Biesheuvel CJ, Grobbee DE. Test research versus diagnostic research. *Clin Chem* 2004;50(3):473-6.
8. Tjoe Nij E, Burdorf A, Parker J, Attfield M, van Duivenbooden C, Heederik D. Radiographic abnormalities among construction workers exposed to quartz containing dust. *Occup Environ Med* 2003;60(6):410-7.
9. Medical Research Council Committee on the Aetiology of Chronic Bronchitis. Standardized questionnaire on respiratory symptoms. *Br Med J* 1960;2:1665.
10. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. 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;16:5-40.
11. ILO. Guidelines for the Use of ILO International Classification of Radiographs of Pneumoconioses. Revised Edition 2000 ed. Geneva: ILO; 2002.
12. 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.

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. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59(10):1087-91.
16. Moons KGM, Grobbee DE. Diagnostic studies as multivariable, prediction research. *J Epidemiol. Community Health* 2002;56:337 - 338.
17. Benyamini Y, Idler EL, Leventhal H, Leventhal EA. Positive affect and function as influences on self-assessments of health: expanding our view beyond illness and disability. *J Gerontol B Psychol Sci Soc Sci* 2000;55(2):P107-16.
18. Kreiss K, Greenberg LM, Kogut SJ, Lezotte DC, Irvin CG, Cherniack RM. Hard-rock mining exposures affect smokers and nonsmokers differently. Results of a community prevalence study. *Am Rev Respir Dis* 1989;139(6):1487-93.
19. Ng TP, Chan SL. Lung function in relation to silicosis and silica exposure in granite workers. *Eur Respir J* 1992;5(8):986-91.
20. Cowie RL. The influence of silicosis on deteriorating lung function in gold miners. *Chest* 1998;113(2):340-3.
21. Cowie RL, Mabena SK. Silicosis, chronic airflow limitation, and chronic bronchitis in South African gold miners. *Am Rev Respir Dis* 1991;143(1):80-4.
22. ACOEM. Evidence based statements: Medical surveillance of workers exposed to crystalline silica. 2005 [cited 2005; Available from: <http://www.acoem.org/guidelines/article.asp?ID=82>
23. ATS. Diagnosis and initial management of nonmalignant diseases related to asbestos. *Am J Respir Crit Care Med* 2004;170:691-715.
24. 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.

# CHAPTER 3

## PROGNOSTIC MODELS



The contents of this chapter are based on  
E Suarhana, JL Malo, D Heederik,  
H Ghezzi, J L'Archevêque RT, D Gautrin

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

*Submitted in a revised version*



---

# WHICH TOOLS BEST PREDICT THE INCIDENCE OF WORK-RELATED SENSITIZATION AND SYMPTOMS?

---

EVA SUARTHANA, JEAN-LUC MALO, DICK HEEDERIK,  
HEBERTO GHEZZO, JOCELYNE L'ARCHEVÊQUE RT, DENYSE GAUTRIN

---

## ABSTRACT

**Background:** We assessed the baseline value of a questionnaire used alone or in combination with skin-prick testing (SPT) to common allergens and/or bronchial responsiveness (BR) testing with methacholine in predicting the occurrence of (1) sensitization to laboratory animal (LA) allergens and (2) respiratory symptoms after 32 months of training.

**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 for asthma and allergic symptoms at baseline composed the final questionnaire model for the occurrence of occupational sensitization 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 sensitization, but not for symptoms at work.

**Conclusions:** Questionnaire is a good tool to predict the incidence of occupational sensitization and symptoms; additional tests improved the specificity of the prediction for LA sensitization.

**KEYWORDS:** allergy, bronchial hyperreactivity, occupational diseases, prognosis, questionnaire

## INTRODUCTION

The attributable risk of occupation for asthma in population-based studies is close to 10% (1) (2) (3) (4), 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 (5). 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 (6). Wherein 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 sensitization to laboratory animal (LA) allergens was proposed.(7) 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 sensitization in three years of follow up. Nevertheless, the predictive value of these tools in regards 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. Our purpose was to investigate the natural history of occupational allergy and asthma.(8-11) Students in animal health technology program showed a high incidence (8.9 per 100 person-years (PY) of sensitization to LA allergens.(10) This group also showed an incidence of probable OA of 2.7 per 100 PY.(11)

We aimed at developing prediction models for incidence of specific IgE-sensitization to and respiratory symptoms in contact with high-molecular-weight agents after 32 months training program in the same population of animal health apprentices by using a questionnaire alone or in conjunction with clinical tests, such as skin-prick testing and bronchial responsiveness testing at entry. 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 program 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 program. The tests were subsequently administered yearly until the end of their program. 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.

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 sensitization to LA allergen at the initial visit were also included and considered at risk of developing a new sensitization to another LA allergen. For the purpose of developing prognostic models for

sensitization 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 (n=314). Prognostic models for symptoms at work were derived from apprentices who had neither rhinoconjunctivitis nor chest symptoms at work at the initial visit (n=296).

### Potential predictors

We used questionnaire items, clinical information on atopic status –skin-prick tests reactivity to common allergens– and BR to methacholine at baseline as potential prognostic predictors. 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.(12) We used principal components analysis (PCA) to reduce the number of questionnaire predictors to be included in the multivariate logistic regression analysis.(13) PCA is often used to identify underlying variables, or factors, that explain the 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 one or greater (i.e. 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 one or greater 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. 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.(14) The procedure for performing the bronchial challenge test was modified as described elsewhere to take into account the absence of an onsite physician.(15) Bronchial responsiveness was considered measurable when the level of the methacholine concentration caused a decrease in FEV<sub>1</sub> of 20% (PC<sub>20</sub>) of 32 mg/ml or lower – a measurable PC<sub>20</sub>. 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 sensitization 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 32<sup>nd</sup> month visit, but no positive skin reactions to any of these LA allergens tested at the initial visit. 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 32<sup>nd</sup> month visit, but no symptoms at the initial visit.

### Data analysis

We developed the prognostic models for sensitization to 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 multivariable 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 sensitization to LA allergens.(16) 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  $\chi^2$  of the 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> models were compared to the 1<sup>st</sup> model's  $\chi^2$  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 sensitization, was assessed graphically and tested with the Hosmer–Lemeshow (H-L) test, wherein a p-value of 0.10 and higher reflects good agreement.(17) The discriminative ability was determined with the area under the receiver operating characteristic (ROC) curve. The area under the ROC curve (ROC area) shows the relationship between a false positive rate (1-specificity) and a true positive rate (sensitivity). 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.(18)

The internal validity of the models was assessed using the bootstrapping procedure. Random bootstrap samples were drawn with replacement from the population consisting of Canadian apprentices (1000 replications). 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.(19) The regression coefficients of the predictors in the final model were multiplied by this correction factor to prevent the model from producing optimistic predictions (i.e. 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 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 (Statistical Package for Social Sciences, Chicago Il) and S-Plus 6 for Windows (Insightful Corp).

## RESULTS

There were 49 of 314 (15.6%) and 45 of 296 (15.2%) apprentices who developed new sensitization to LA allergens and symptoms at work, respectively, at the 32<sup>nd</sup> month of the training program.

### **Predictors of incidence of sensitization 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 sensitization to LA allergens (Table 1). SPT based atopy or BR were both associated with the endpoints with comparable strength.

**Table 1** Distribution and strength of univariable association between the predictors and sensitization to LA Allergens

	Total (n=314)	Not-sensitized (n=265)	Sensitized (n=49)	OR (95% CI)
Age mean (standard deviation) 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 for asthma *	139 (44.3)	106 (40.0)	33 (67.3)	3.1 (1.6 to 5.9) **
Symptoms indicative for 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 <sub>20</sub> (≤ 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.

OR (95% CI): odds ratio (95% confidence interval); COPD: chronic obstructive pulmonary diseases; SPT: skin-prick test; PC<sub>20</sub>: level of the methacholine concentration that caused a decrease in FEV<sub>1</sub> of 20%.

\* 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; awaken 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 eleven common inhalants.

\*\* Univariable p-value < 0.50.

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 for asthma and allergic symptoms appeared as independent predictors of the development of LA sensitization 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 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 allergen

**Table 2** Associations between the predictors and sensitization to laboratory animal allergens from multivariable regression modelling

	Questionnaire only (Model 1)		Questionnaire + SPT (Model 2)		Questionnaire + BR (Model 3)		Questionnaire + SPT + BR (Model 4)	
	$\beta$	OR (95% CI)	$\beta$	OR (95% CI)	$\beta$	OR (95% CI)	$\beta$	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 <sub>20</sub> ( $\leq$ 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	

SPT: skin-prick testing; BR: bronchial responsiveness to methacholine;  $\beta$ : regression coefficient after multiplication by correction factor from bootstrapping procedure; OR (95% CI): odds ratio (95% confidence interval).

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

ROC area (95%CI) Measurable PC<sub>20</sub> 0.69 (0.60 to 0.77)

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

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

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

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  $X^2$  (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  $X^2$  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  $X^2$  (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 for asthma and allergic symptoms remained in the final questionnaire model (Table 4, Model 1).

**Table 3** Distribution and the strength of univariable association between the predictors and symptoms at work

	Total (n=296)	No symptoms (n=251)	Had symptoms at work (n=45)	OR (95% CI)
Age mean (standard deviation) 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 for asthma	125 (42.2)	96 (38.2)	29 (64.4)	2.9 (1.5 to 5.7) **
Symptoms indicative for 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 <sub>20</sub> (< 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. OR (95% CI): odds ratio (95% confidence interval); COPD: chronic obstructive pulmonary diseases; SPT: skin-prick test; PC<sub>20</sub>: level of the methacholine concentration that caused a decrease in FEV<sub>1</sub> of 20%.  
 \*\* Univariable p-value < 0.50.

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 BH testing alone was significantly lower than the questionnaire model (Table 4, legend). 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  $X^2$  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 sensitization (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 sensitization, 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

**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)	
	$\beta$	OR (95% CI)	$\beta$	OR (95% CI)	$\beta$	OR (95% CI)	$\beta$	OR (95% CI)
Intercept	-3.15		-3.32		-3.45		-3.40	
Symptoms indicative for 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 <sub>20</sub> ( $\leq$ 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	

SPT: skin-prick testing; BR: bronchial responsiveness to methacholine;  $\beta$ : regression coefficient after multiplication by correction factor from bootstrapping procedure; OR (95% CI): odds ratio (95% confidence interval).

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

ROC area (95%CI) Measurable PC<sub>20</sub> 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:

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

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

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.

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

	Number of apprentices in the group (%) <sup>*</sup>	Number of apprentices with sensitization (n=49) n (%) <sup>‡</sup>	Number of apprentices without sensitization (n=265) n (%) <sup>‡‡</sup>	Sensitivity (%)	Specificity (%)	NPV (%)
Probability >=0.10						
Model 1	195 (62.1)	44 (22.6)	151 (77.4)	89.8	43.0	95.8
Model 2	176 (56.1)	41 (23.3)	135 (76.7)	83.7	49.1	94.2
Model 3	187 (59.6)	44 (23.5)	143 (76.5)	89.8	46.0	96.1
Model 4	159 (50.6)	41 (25.8)	118 (74.2)	83.7	55.5	94.8
Probability >=0.20						
Model 1	84 (26.8)	27 (32.1)	57 (67.9)	55.1	78.5	90.4
Model 2	83 (26.4)	31 (37.3)	52 (62.7)	63.3	80.4	92.2
Model 3	59 (18.8)	26 (44.1)	33 (55.9)	53.1	87.5	91.0
Model 4	90 (28.7)	33 (36.7)	57 (63.3)	67.3	78.5	90.4
Probability >=0.30						
Model 1	84 (26.8)	27 (32.1)	57 (67.9)	55.1	78.5	90.4
Model 2	56 (17.8)	24 (42.9)	32 (57.1)	49.0	87.9	90.3
Model 3	59 (18.8)	26 (44.1)	33 (55.9)	53.1	87.5	91.0
Model 4	48 (15.3)	24 (50.0)	24 (50.0)	49.0	90.9	90.6

NPV: negative predicted value; Model 1: questionnaire alone; Model 2: questionnaire and SPT; Model 3: questionnaire and BR testing; Model 4: questionnaire, SPT, and BR testing.

<sup>\*</sup> Proportion of all apprentices (n=314)

<sup>‡</sup> Proportion of apprentices with symptoms at work within the probability group

<sup>‡‡</sup> Proportion of apprentices without symptoms at work within the probability group

To apply the models in daily practice, we transformed the regression questionnaire models for both endpoints into easy-to-use score charts (Appendices 2 and 3). As an example on 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 sensitization 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 sensitization to LA allergens and respiratory symptoms at work using data from a prospective study of apprentices.(8-11) We found that the questionnaire models had a good accuracy in predicting incidence of both endpoints. Addition of SPT to common allergens and/or BR testing increased the specificity of the prediction for LA sensitization, but not for symptoms at work.

**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

	Number of apprentices in the group (%) <sup>*</sup>	Number of apprentices with symptoms at work (n=45) n (%) <sup>‡</sup>	Number of apprentices without symptoms at work (n=251) n (%) <sup>‡‡</sup>	Sensitivity (%)	Specificity (%)	NPV (%)
Probability >=0.10						
Model 1	118 (39.9)	37 (31.4)	81 (68.6)	82.2	67.7	95.5
Model 2	133 (44.9)	38 (28.6)	95 (71.4)	84.4	62.2	95.7
Model 3	143 (48.3)	38 (26.6)	105 (73.4)	84.4	58.2	95.4
Model 4	124 (41.9)	37 (29.8)	87 (70.2)	82.2	65.3	95.3
Probability >=0.20						
Model 1	118 (39.9)	37 (31.4)	81 (68.6)	82.2	67.7	95.5
Model 2	92 (31.1)	33 (35.9)	59 (64.1)	73.3	76.5	94.1
Model 3	86 (29.1)	32 (37.3)	54 (62.8)	71.1	78.5	93.8
Model 4	100 (33.8)	36 (36.0)	64 (64.0)	80.0	74.5	95.4
Probability >=0.30						
Model 1	67 (22.6)	25 (37.3)	42 (62.7)	55.6	83.3	91.3
Model 2	67 (22.6)	25 (37.3)	42 (62.7)	55.6	83.3	91.3
Model 3	86 (29.1)	32 (37.3)	54 (62.8)	71.1	78.5	93.8
Model 4	57 (19.3)	23 (40.4)	34 (59.6)	51.1	86.5	90.8

NPV: negative predicted value; Model 1: questionnaire alone; Model 2: questionnaire and SPT; Model 3: questionnaire and BR testing; Model 4: questionnaire, SPT, and BR testing.

<sup>\*</sup> Proportion of all apprentices (n=296)

<sup>‡</sup> Proportion of apprentices with symptoms at work within the probability group

<sup>‡‡</sup> Proportion of apprentices without symptoms at work within the probability group

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 sensitized from non sensitized apprentices to that from SPT or BR testing alone (ROC area of 0.73 vs. 0.69 and 0.69, respectively). For symptoms at work, the discriminative ability of the questionnaire model was significantly higher than SPT as well as BR testing alone (ROC area of 0.78 vs. 0.65 and 0.61, respectively). We used a measurable PC<sub>20</sub> to define bronchial responsiveness because extra analysis prior to the modelling showed that a measurable PC<sub>20</sub> had the strongest univariable association with the outcomes as compared to stricter definitions, for example PC<sub>20</sub> 16 mg/ml or lower.

A ROC area of 0.73 meant that by using the questionnaire model we could correctly assign a higher probability to a sensitized apprentice in 73% pairs of apprentices in which one apprentice is sensitized to LA allergens and one is not(18). This figure was equal to what was demonstrated by the questionnaire model for sensitization to LA allergens that was developed in a population of Dutch workers (ROC area of 0.73; 95%CI 0.73 to 0.89). In the quoted study(7), the addition of total IgE and sensitization 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, the addition of SPT and/or bronchial responsiveness testing significantly improved the model  $\chi^2$  of the questionnaire model for LA sensitization. The ROC areas did not increase significantly, but as illustrated in Table 5 addition of these tests increased the specificity of the prediction model. Nevertheless, none of SPT, BR testing or their combination improved the questionnaire model for symptoms at work.

Adding information from SPT and BR testing yielded marginal to moderate additional predictive information; 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 patients with disease manifestation, for example in the clinical setting, may yield different models with different strengths of association between predictors and the end point. 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 health care and social system, and legal context where the model will be applied, one may prefer to use a complete questionnaire-based model (counselling) or a questionnaire and additional tests based models (rigid surveillance).

As for the outcomes, we used a loose definition of respiratory symptom at work - developed any of nasojunctival 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 sensitization to LA allergens. We did the principal components analysis (PCA) to reduce the number of predictors to be included in the multivariate logistic regression analysis.(13) After the models were developed, we did the bootstrapping procedure to check whether the models were reasonably valid and would not produce an optimistic estimate in a new population (i.e. too high or too low). This procedure has been shown to be superior than split-sample or cross validation methods.(20) 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 (provide 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 generalizability into other groups of animal health apprentices.(21, 22)

To the best of authors' knowledge, these are the first prognostic models for LA sensitization and respiratory symptoms at work developed in an animal health apprentices setting. Such models enable the quantification of the probability of developing sensitization to LA allergens or respiratory symptoms at work for individuals and thus, they may support career counselling and health monitoring programs 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 sensitization 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 would be 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 in a high probability group of developing sensitization 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 (i.e. bronchial methacholine challenge test) in every worker. As shown in Table 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 sensitization to LA allergens, we will save the expenses for unnecessary test 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 because we can expect that missed cases will be captured in the next surveillance round. However, if the aim is to detect as many cases as possible, a lower cut-off value might be of interest.

In conclusion, we developed prognostic models to predict the occurrence of sensitization to LA allergens and symptoms at work in animal health apprentices. The

questionnaire model alone is a good tool to predict incidence of occupational sensitization and symptoms. Addition of SPT and/or bronchial responsiveness testing improved the specificity of the prediction for LA sensitization. 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. It would also be interesting to develop similar models in populations exposed to other high- and low-molecular weight agents causing occupational asthma in further study.

## REFERENCES

1. Blanc P, Toren K. How much asthma can be attributed to occupational factors? *Am J Med* 1999;107:580-587.
2. Kogevinas M, Anto J, Sunyer J, Tobias A, Kromhout H, Burney P, et al. Occupational asthma in Europe and other industrialised areas: a population-based study. *Lancet* 1999;353:1750-1754.
3. Becklake M, Chan-Yeung M, Malo J. Epidemiological Approaches in Occupational Asthma. In: *Asthma in the Workplace*, 3rd edition. Bernstein IL, Chan-Yeung M, Malo JL, Bernstein DI, eds. Francis & Taylor, New York, NY 2006:37-85.
4. Kogevinas M, Zock J, Jarvis D, Kromhout H, Lillienberg L, Plana E, et al. Exposure to substances in the workplace and new-onset asthma: an international prospective population-based study (ECRHS-II). *Lancet* 2007;370:336-341.
5. Bernstein I, Bernstein D, Chan-Yeung M, Malo J. Definition and classification of asthma in the workplace. In: *Asthma in the Workplace*. 2006;Bernstein IL, Chan-Yeung M, Malo JL, Bernstein DI, eds. 3rd ed., Taylor & Francis, New York:1-8.
6. Bernstein D, Campo P, Baur X. Clinical Assessment and Management of Occupational Asthma. In: *Asthma in the workplace*, 3rd ed. 2006;Bernstein IL, Chan-Yeung M, Malo JL, Bernstein DI., eds. Taylor & Francis., New York:161-178.
7. Meijer E, Grobbee D, Heederik D. A strategy for health surveillance in laboratory animal workers exposed to high molecular weight allergens. *Occup Environ Med* 2004;61:831-837.
8. 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):1841-1847.
9. Gautrin D, Ghezzi H, Infante-Rivard C, Malo JL. Natural history of sensitization, symptoms and occupational diseases in apprentices exposed to laboratory animals. *Eur Respir J* 2001;17(5):904-8.
10. Gautrin D, Ghezzi H, Infante-Rivard C, Malo JL. Incidence and determinants of IgE-mediated sensitization in apprentices. A prospective study. *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1222-8.
11. Gautrin D, Infante-Rivard C, Ghezzi H, Malo JL. Incidence and host determinants of probable occupational asthma in apprentices exposed to laboratory animals. *Am J Respir Crit Care Med* 2001;163(4):899-904.
12. 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.
13. Bollen K. Structural equations with latent variables. New York: John Wiley & Sons; 1989.
14. Sterk PJ, Fabbri LM, Quanjer PH, Cockcroft DW, O'Byrne PM, Anderson SD, 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.
15. Troyanov S, Malo JL, Cartier A, Gautrin D. Frequency and determinants of exaggerated bronchoconstriction during shortened methacholine challenge tests in epidemiological and clinical set-ups. *Eur Respir J* 2000;16(1):9-14.
16. 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.

17. Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley and Sons, Inc; 1989.
18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143(1):29-36.
19. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med* 1990;9(11):1303-25.
20. Peduzzi P, Concato J, Kemper E, Holford T, Feinstein A. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373-1379.
21. Bleeker S, Moll H, Steyerberg E, Donders A, Derksen-Lubsen G, Grobbee D, et al. External validation is necessary in prediction research: A clinical example. *J Clin Epidemiol* 2003;56(8):826-832.
22. Justice A, Covinsky K, Berlin J. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515-524.

## Appendix 1

Clusters of correlating symptoms and percentage of explained variance as determined by principal component analysis (factor loadings after orthogonal varimax rotation between brackets)

<b>1. Respiratory symptoms at work</b>	<b>2. Symptoms indicative for asthma</b>	<b>3. Allergic symptoms</b>
Explained variance 25.1%	Explained variance 13.2%	Explained variance 10.4%
Do you ever get itchy, red, and/or watery eyes? (0.65)	Have you had wheezing or whistling in the last 12 months? (0.79)	Do you develop eye or nasal or respiratory or skin symptoms when exposed to common allergen such as dust mite, animal hair or pollen? (0.87)
Do you ever get runny or stuffy nose or sneezing? (0.49)	Have you had an attack of shortness of breath in the last 12 months? (0.80)	Do you ever have an itchy runny nose or sneezing even when you do not have a cold? (0.90)
Do you ever feel shortness of breath? (0.82)	Do you ever start to cough induced by exercise, strenuous work, cold air, heavy smell, smoke, or dust? (0.68)	Have you ever had "hay fever"? (0.72)
Do you ever start to cough? (0.86)	Asthma confirmed by physician (0.69)	
Do you ever start to wheeze? (0.85)		
<b>4. Symptoms indicative for COPD</b>		<b>5. Skin symptoms</b>
Explained variance 6.5%		Explained variance 5.8%
Do you sometimes wake with a feeling of tightness in your chest first thing in the morning? (0.62)		Have you ever had eczema (0.68)
In the last few months, have you been awoken by cough at night? (0.69)		Have you ever had urticaria or hives (0.73)
Do you usually bring up phlegm from your chest first thing in the morning for at least 3 months each year? (0.49)		
Have you ever had chronic bronchitis confirmed by doctor? (0.52)		

**Appendix 2**

Application of questionnaire prognostic model for incidence of sensitization to laboratory animal allergens 32 months after entry of training in an animal health technology program

**Score Chart**

Predictors	Value	Score		
Symptoms indicative for asthma	Present if give "YES" answer to at least one of the following questions: - Have you had wheezing or whistling in the last 12 months? - Have you had an attack of shortness of breath in the last 12 months? - Ever diagnosed asthma by physician - Do you ever start to cough induced by exercise, strenuous work, cold air, heavy smell, smoke, or dust?	1.0		
Allergic symptoms	Present if "YES" answer to at least one of the following questions: - Do you develop eye or nasal or respiratory or skin symptoms when exposed to common allergen such as dust mite, animal hair or pollen? - Do you ever have an itchy runny nose or sneezing even when you do not have a cold? - Have you ever had "hay fever"?	2.0		
<b>Sum score</b>		...		
<b>Sum score</b>	0	1	2	3
Predicted probability of sensitization to LA allergens (%)	3	7	18	37

**Comparison of the accuracy of the selected cut-off points of the sum score**

Sum score	Number of apprentices (%) <sup>a</sup> in the group	Number of apprentices with sensitization (n=49) n (%) <sup>b</sup>	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
>=1	195 (62.1)	44 (22.6)	89.8	43.0	22.6	95.8
>= 2	140 (44.6)	38 (27.1)	77.6	61.5	27.1	93.7
3	84 (26.8)	27 (32.1)	55.1	78.5	32.1	90.4

PPV= Positive predictive value; NPV= Negative predictive value

<sup>a</sup> Proportion of all apprentices (n=314)

<sup>b</sup> Proportion of apprentices with occupational sensitization within the sum score category

### Appendix 3

Application of questionnaire prognostic model for incidence of symptoms at work 32 months after entry of training in an animal health technology program

#### Score Chart

Predictors	Value	Score		
Symptoms indicative for asthma	Present if give "YES" answer to at least one of the following questions: <ul style="list-style-type: none"> <li>- Have you had wheezing or whistling in the last 12 months?</li> <li>- Have you had an attack of shortness of breath in the last 12 months?</li> <li>- Ever diagnosed asthma by physician</li> <li>- Do you ever start to cough induced by exercise, strenuous work, cold air, heavy smell, smoke, or dust?</li> </ul>	1.0		
Allergic symptoms	Present if "YES" answer to at least one of the following questions: <ul style="list-style-type: none"> <li>- Do you develop eye or nasal or respiratory or skin symptoms when exposed to common allergen such as dust mite, animal hair or pollen?</li> <li>- Do you ever have an itchy runny nose or sneezing even when you do not have a cold?</li> <li>- Have you ever had "hay fever"?</li> </ul>	2.0		
<b>Sum score</b>		...		
<b>Sum score</b>	0	1	2	3
Predicted probability of symptoms at work (%)	1	4	23	44

#### Comparison of the accuracy of the selected cut-off points of the sum score

Sum score	Number of apprentices (%) <sup>a</sup> in the group	Number of apprentices with symptoms at work (n=45) n (%) <sup>b</sup>	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
>=1	176 (59.5)	41 (23.3)	91.1	46.2	23.3	96.7
>= 2	118 (39.9)	37 (31.4)	82.2	67.7	31.4	95.5
3	67 (22.6)	25 (37.3)	55.6	83.3	37.3	91.3

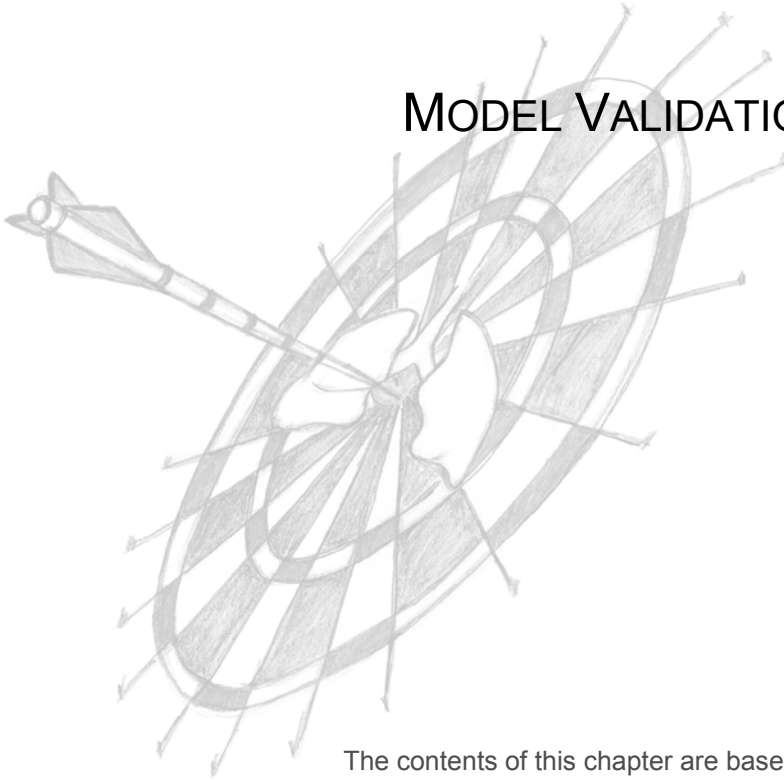
PPV= Positive predictive value; NPV= Negative predictive value

<sup>a</sup> Proportion of all apprentices (n=296)

<sup>b</sup> Proportion of apprentices with symptoms at work within the sum score category

# CHAPTER 4

## MODEL VALIDATION



The contents of this chapter are based on  
E Suarathana, E Meijer, D Heederik, H Ghezzi, JL Malo, D Gautrin  
The generalizability of the Dutch diagnostic model  
for laboratory animal allergens sensitization in Canadian apprentices  
*Provisionally accepted for publication in the Journal of Clinical Epidemiology*



# EXTERNAL VALIDATION OF DIAGNOSTIC MODEL FOR SENSITIZATION TO LABORATORY ANIMAL ALLERGENS IN CANADIAN APPRENTICE

---

EVA SUARTHANA, EVERT MEIJER, DICK HEEDERIK,  
HEBERTO GHEZZO, JEAN-LUC MALO, DENYSE GAUTRIN

---

## ABSTRACT

**Background:** In general, prediction models show a lower performance in populations other than where the model was derived. We aimed at assessing the transportability of an existing diagnostic questionnaire model for the sensitization to laboratory animal allergens.

**Methods:** 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 (H-L) 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 (H-L test p-value <0.001). The other methods yielded models with good calibration (p-value > 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 factors from the bootstrapping procedure were > 0.90).

**Conclusions:** Once updated, the diagnostic model is valid and can be applied with reasonable performance in an animal health apprentice setting.

**KEYWORDS:** diagnostic model, occupational sensitization, questionnaire, screening, validity

## 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 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 laboratory animal (LA) allergens among Dutch laboratory workers exposed to rats, mice and other rodents.(2) They demonstrated that a diagnostic model based upon questionnaire items could be accurately used to predict the presence of workers at high- or at 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 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)

The utility of predictive models depends on how well they perform when applied to a population who 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 that reason, external validation of the model is necessary to address the accuracy of a model in different, but related workers.(5-7) Canadian apprentices at 1.4 months after entry showed a different distribution of personal and exposure characteristics compared to 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.

## METHODS

### Populations

The existing diagnostic questionnaire model for the sensitization to laboratory animal (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, skin-prick tests (SPT), and lung function tests were collected from all workers. The self-administered questionnaire was based upon a Dutch version of an internationally accepted respiratory questionnaire.(9) The questionnaire included questions on age, gender, respiratory problems, personal & family history of allergic symptoms, and smoking history. Additional questions were asked about employment history, work duration with laboratory animals, and allergic symptoms due to working with laboratory animals.

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 non-specific provocation tests upon the beginning of their three-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.

### 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 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 laboratory animals. 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 hyper-responsiveness (BHR) were considered present if subjects experienced respiratory problems induced by exercise, strenuous work, very cold air, heavy smell, smoke, or dust.

### 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) as well as 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.(1, 2) Sensitization to LA allergens was defined as a positive SPT response to any occupational allergens.

### Data analysis

Of 417 eligible individuals, 3 (0.7%) apprentices with no skin-prick test result were excluded, leaving n=414 (99.3%) with complete data.

#### *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:

$$P(\text{sensitization}) = 1 / (1 + \exp(-(-1.82 + \text{history of asthma attacks} * 0.98 + \text{history of allergic symptoms} * 0.87 + \text{allergic symptoms during work} * 1.03 + \text{exposed to rats} \geq 20 \text{ hours/week} * 0.79 + \text{male gender} * 0.46))).(2)$$

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 done using the logistic regression function provided in SPSS version 14.0 for Windows (Statistical Package for Social Sciences, Chicago II): predictors from the Dutch model were frozen in the first block and subsequently, predictors with univariable p-value <0.5 from Canadian setting were added in the second block. Backward stepwise selection was applied with an inclusion criterion of  $p < 0.157$  in the second block, and the overall  $\chi^2$  of the block was evaluated.

The diagnostic accuracy of the diagnostic model was quantified using calibration and discrimination measures (*Hmisc* and *Design* library, function “*fit*” and “*val.prob*” in S-Plus 6 for windows (Insightful Corp)). 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-value of 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)

Method 3 and 4 were followed by bootstrapping (S-Plus *Hmisc* and *Design* library, function “*fit*”, “*validate*” and “*val.prob*”) 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.

#### *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 a p-value < 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 above.

#### *Model Application*

For screening purposes, a cut-off 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 cut-off 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 (S-Plus *Hmisc* and *Design* library, function *nomogram*). 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.

## 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).

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

	Dutch workers n=472	Canadian apprentices n=414
Age years <sup>a</sup>	34.4 (9.4)	18.3 (3.3)
Exposure duration years <sup>a</sup>	8.3 (8.9)	
Time exposed to rats hours/week <sup>a</sup>	8.5 (10.4)	0.13 (0.15)
Work with rats >20 h/week	59 (12.5)	0 (0.0)
Male gender	293 (62.1)	57 (13.8)
Allergic symptoms during the past 12 months <sup>b</sup>	131 (27.8)	191 (41.6)
Allergic symptoms during work <sup>c</sup>	152 (32.2)	72 (17.4)
Asthma attacks	66 (14.0)	48 (11.6)
Symptoms of asthma <sup>d</sup>	45 (9.5)	75 (18.1)
Symptoms suggestive for BHR <sup>e</sup>	192 (40.8)	192 (46.4)
Personal atopic history <sup>f</sup>	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)
<b>Outcome</b>		
Positive SPT to LA allergens <sup>g</sup>	129 (27.3)	57 (13.8)

Data presented as n (%) unless otherwise stated

<sup>a</sup> Mean (standard deviation)

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

<sup>c</sup> Experience chest tightness, running nose or sneezing, running or itching eyes, and/or itching skin on contact with laboratory animals

<sup>d</sup> Present if there were at least two relevant symptoms (wheezing, chest tightness, cough, and dyspnea)

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

<sup>f</sup> Present in the case history of eczema, urticaria, or hay fever

<sup>g</sup> Present if there is a positive skin reactivity to rat/mouse urine and/or to rabbit/guinea pig dander

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 five percent of the Dutch worked  $\geq 20$  hours/week with rats while all Canadian apprentices were exposed less than 1 hour/week 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 to 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).

### Diagnostic questionnaire model derived from the Canadian apprentices

We fitted a full model comprising age, gender, 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-value=0.298).

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

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

Data presented as n (%) unless otherwise stated

<sup>a</sup> Odds ratio (95% confidence interval).

<sup>b</sup> Mean (standard deviation)

### External validation of the Dutch diagnostic questionnaire model

Recalibration of the model's intercept (Method 2) gave a new intercept of -2.86 (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 hours/week), the third method yielded comparable regression coefficients of the predictors as compared to 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  $X^2$  deviance=11.8, df=2). The internal validity of the revised model was good; the correction factor from the bootstrapping procedure was 0.94.

**Table 3** The Dutch Questionnaire model when validated in Canadian apprentices across different validation methods, compared to the Canadian Questionnaire model

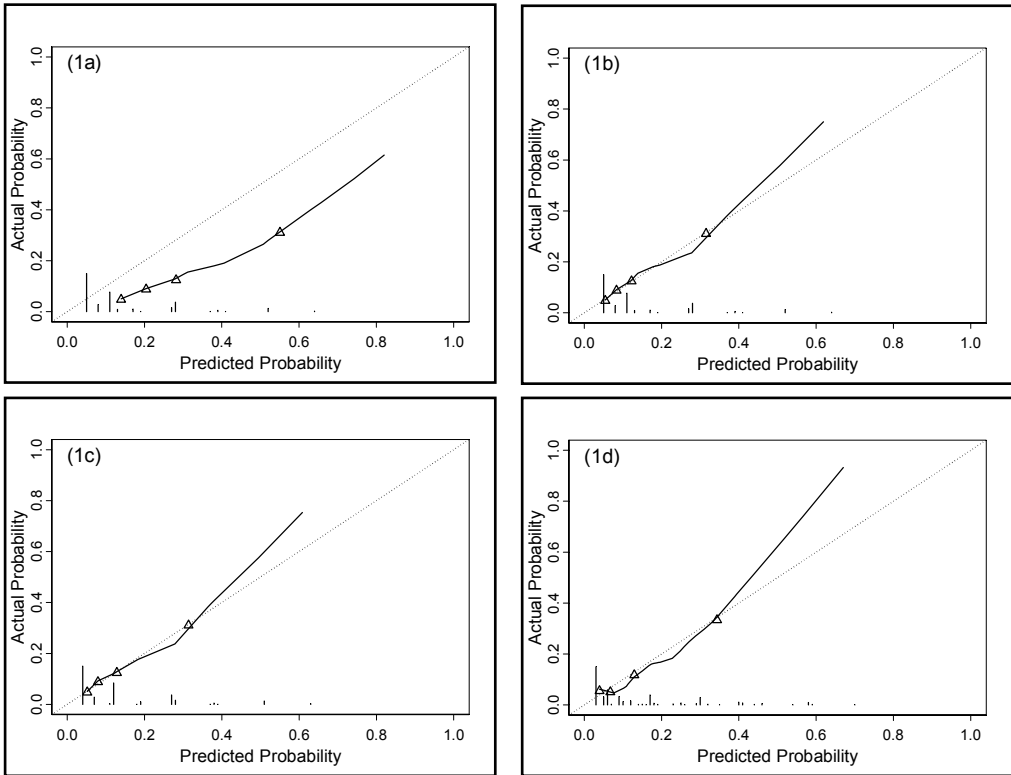
	Validation of the Dutch Model				Canadian Model
	No update	Recalibrate the intercept	Refit the Model	Model revision	
	(Method 1)	(Method 2)	(Method 3)	(Method 4)	
	$\beta^a$	$\beta^a$	$\beta^b$	$\beta^b$	$\beta^b$
Intercept	-1.82	-2.86	-2.91	-3.19	-3.11
Work with rats >20 h/week	0.79	0.79	<sup>c</sup>		
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

<sup>a</sup> Regression coefficient

<sup>b</sup> Regression coefficient after multiplication by correction factor from bootstrapping procedure

<sup>c</sup> Not applicable

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-value < 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 method were good (the HL-test p-value >0.1).



**Figure 1** Calibration plots of the Dutch questionnaire model when applied without adjustment (1a), the intercept was recalibrated (1b), the model was refitted (1c), and the model was revised (1d) 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-quantiles of predicted probability. Distribution of the predicted probabilities is indicated with vertical lines at the bottom.

### Model application

To generate case specific advice and support decision making by health professionals, several cut-off points of the predicted probability were selected and their diagnostic accuracy was compared across the updated models (Table 4). A higher cut-off leads to a higher specificity and PPV, but at the cost of lower sensitivity, and vice versa. A cut-off 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 cut-off points. Therefore, the revised model was converted to a nomogram (Fig. 2). The total scoring points corresponded with 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 asthma attacks and atopy, has a total number of points of 145 (0+0+0+45+0+100), which corresponds to a predicted probability of sensitization of 0.15.

**Table 4** Comparison of diagnostic accuracy of the selected cut-off point of the predicted probability produced by different validation methods

<b>Probability Cut-off</b>	<b>Number of apprentices in the probability group (%)<sup>a</sup></b>	<b>Number of apprentices with sensitization (n=57) n (%)<sup>b</sup></b>	<b>Number of apprentices without sensitization (n=357) n (%)<sup>c</sup></b>	<b>Sensitivity (%)</b>	<b>Specificity (%)</b>	<b>PPV<sup>d</sup> (%)</b>	<b>NPV<sup>e</sup> (%)</b>
<b>≥ 0.10</b>							
Method 2	204 (49.3)	45 (22.1)	159 (77.9)	78.9	55.5	22.1	94.3
Method 3	204 (49.3)	45 (22.1)	159 (77.9)	78.9	55.5	22.1	94.3
Method 4	192 (46.4)	44 (22.9)	148 (77.1)	77.2	58.5	22.9	94.1
<b>≥ 0.15</b>							
Method 2	102 (24.6)	32 (31.4)	70 (68.6)	56.1	80.4	31.4	92.0
Method 3	102 (24.6)	32 (31.4)	70 (68.6)	56.1	80.4	31.4	92.0
Method 4	135 (32.6)	40 (29.6)	95 (70.4)	70.2	73.4	29.6	93.9
<b>≥ 0.20</b>							
Method 2	90 (21.7)	29 (32.2)	61 (67.8)	50.9	82.9	32.2	91.4
Method 3	89 (21.5)	29 (32.6)	60 (67.4)	50.9	83.2	32.6	91.4
Method 4	86 (20.8)	32 (37.2)	54 (62.8)	56.1	84.9	32.6	92.4

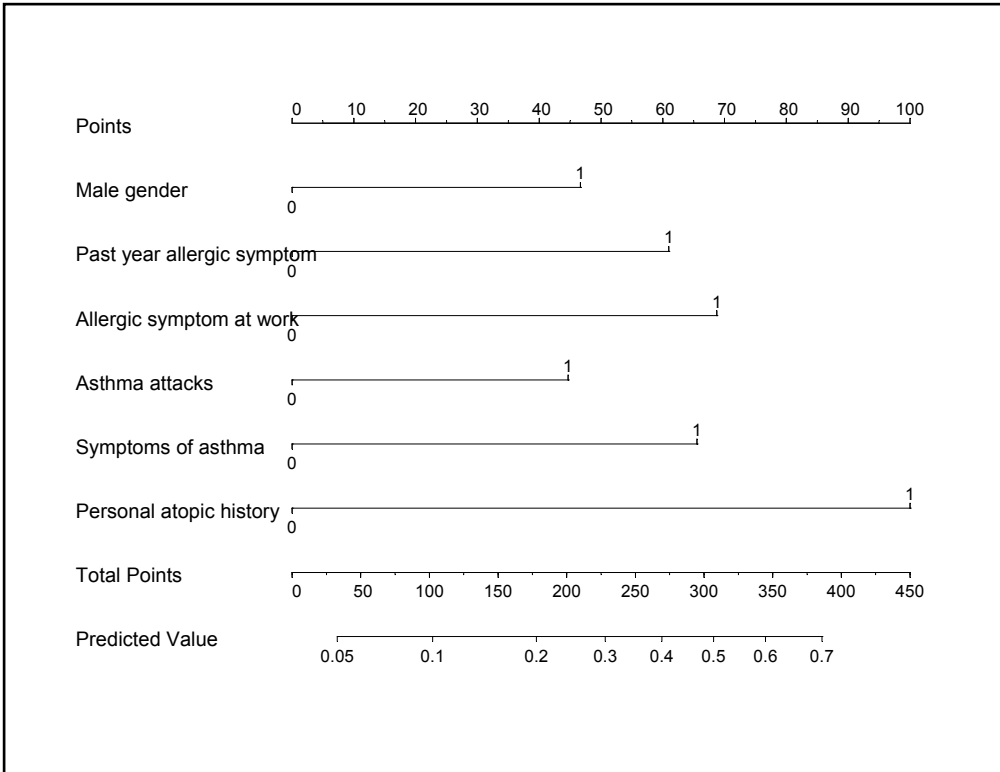
<sup>a</sup> Proportion of all apprentices (n=414)

<sup>b</sup> Proportion of apprentices with occupational sensitization within the sum score category

<sup>c</sup> Proportion of apprentices without occupational sensitization within the sum score category

<sup>d</sup> Positive predictive value

<sup>e</sup> Negative predictive value



**Figure 2** Nomogram for manual calculation of the predicted probability of sensitization to LA 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 asthma attacks and atopy, has a total number of points of 145 (0+0+0+45+0+100) which corresponds to a predicted probability of wheat sensitization of 0.15.

## DISCUSSION

We aimed at assessing the transportability of an earlier obtained diagnostic questionnaire model for sensitization to laboratory animal 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)

### Issues on modeling and external validation

It is generally acknowledged that regression models used in diagnostic prediction research performs better on data from which the model is derived then 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 Fig.1a clearly demonstrates that when the Dutch model was 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 since it used a crude classification with a cut-point of 20 hours/week exposure, while all the Canadian apprentices worked with rats less than 20 hours/week 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%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  $p$ -value  $> 0.10$  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  $X^2$ . However, the ROC area of the revised model was only slightly higher compared to the ROC area obtained by the simpler methods. This could be explained by changes in regression coefficients after the inclusion of the new predictors.

Re-estimating the intercept as well as the regression coefficients of predictors yielded very similar estimations when compared to the original values. This was not surprising since 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 8 potential predictors in the full model; event per variable ratio 7 to 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.(16) 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 less optimism.

### **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 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 cut-off 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 cut-off which offers a higher specificity and PPV, but at the cost of lower sensitivity.

The PPV for this test is comparable to 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.<sup>(17)</sup> Interestingly, although both screening tests shared similar diagnostic properties, they were applied in a very different context. In general, false positive will lead to unnecessary stress and healthcare burden, whereas false negative 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, while 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 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. Since occupational sensitization is not a disease, but rather a precondition which is strongly associated with development of occupational respiratory symptoms, it is acceptable to speculate that false negative cases would 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 laboratory animal allergens can increase the efficiency of health surveillance programs in these apprentices.

## REFERENCES

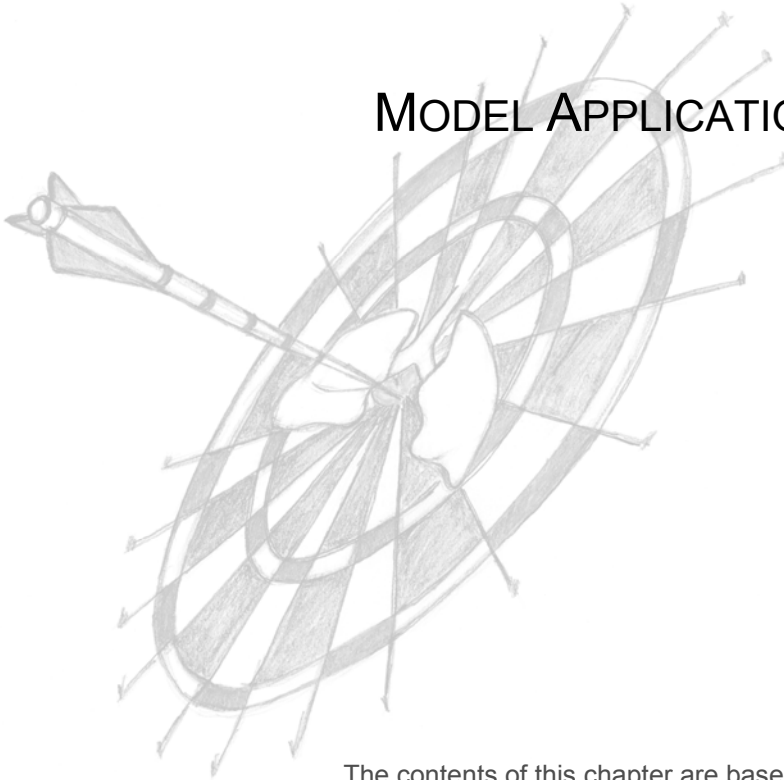
1. Gautrin D, Infante-Rivard C, Dao T, Magnan-Larose M, Desjardins D, Malo J. Specific IgE-dependent sensitization, atopy, and bronchial hyperresponsiveness in apprentices starting exposure to protein-derived agents. *Am. J. Respir. Crit. Care Med.* 1997;155(6):1841-1847.
2. Meijer E, Grobbee DE, Heederik 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. Suarathana 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;56:826-832.
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. Standardized questionnaire on respiratory symptoms. *Br Med J* 1960;2:1665.
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. 159-171.
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. 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.
17. 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.



# CHAPTER 5

## MODEL APPLICATION



The contents of this chapter are based on  
E Meijer, E Suarhana, J de Monchy, F van Rooy, J Rooijackers, T Meijster,  
J Jacobs, E van Otterloo, J Spithoven, V Zaat, DE Grobbee, D Heederik  
Diagnostic research in occupational respiratory allergy:  
*Results from a nationwide surveillance among bakery workers.*  
*Submitted in a revised version*



---

# DIAGNOSTIC RESEARCH IN OCCUPATIONAL RESPIRATORY ALLERGY

*RESULTS OF A NATIONWIDE SURVEILLANCE AMONG BAKERY WORKERS*

---

EVERT MEIJER, EVA SUARTHANA, JAN DE MONCHY, FRITS VAN ROOY,  
JOS ROOIJACKERS, TIM MEIJSTER, JOSE JACOBS, EEF VAN OTTERLOO,  
JACK SPITHOVEN, VANESSA ZAAT, DIEDERICK E. GROBBEE, DICK HEEDERIK

---

## ABSTRACT

**Background:** A covenant between the Dutch government, responsible branch organizations (from traditional and industrial bakeries, flour mills, and baking product manufacturers) and unions agreed on reducing occupational exposure to flour dust and related allergens, in order to reduce the burden of occupational related allergic diseases.

**Methods:** A three-phase national health surveillance was developed. In the preparatory phase, a diagnostic questionnaire model to estimate the probability of sensitisation to wheat and/or  $\alpha$ -amylase allergens was developed. In the 1<sup>st</sup>-phase this questionnaire model was distributed and used to classify workers into three categories: low, intermediate and high probability of developing occupational sensitisation. To evaluate the performance of the diagnostic model, a substudy was carried out in randomly selected traditional and industrial bakeries. All workers in the selected bakeries were asked to complete an extended questionnaire and to draw blood for serology testing. In the 2<sup>nd</sup>-phase, the intermediate probability group will be evaluated further by occupational physicians. The high probability group will be referred to a specialized occupational respiratory health clinic, whereas the low probability group will be enrolled in the next surveillance cycle. In the last phase, intervention for individual worker and their workplace will be carried out by a multidisciplinary team.

**Results:** A total of 5,325 workers participated in the 1<sup>st</sup> phase of surveillance. There were 3,059 (57.4%) workers with a low, 1,282 (24.1%) with an intermediate, and 984 (18.5%) with a high probability of sensitisation to wheat and/or  $\alpha$ -amylase. Workers in the high probability group had the highest percentages of allergic and respiratory symptoms than workers in the low and intermediate group; similar pattern was found in all sectors. Evaluation of 766 workers in the substudy from whom serum samples were available showed that 70% of the sensitized workers were in the intermediate and high probability group. Workers with a high probability reported the highest percentages of allergic symptoms in the past 12 months, medication use for respiratory problems, doctor visit for allergic complaints, absenteeism and changed job due to allergic symptoms.

**Conclusion:** This national three-phase surveillance among bakers in the Netherlands showed that workers with different risks of having an (occupational) allergic health problem can effectively be captured by applying the diagnostic rule for sensitization to wheat and/or  $\alpha$ -amylase allergens. This approach reduces costs, thereby increases feasibility and is applicable for small and medium sized enterprises. Early detection of subgroups of workers with an intermediate or high probability of sensitization improved the expected benefit of diagnosis and treatment and outweighed the expected potential harm of no diagnosis and no treatment.

**KEYWORDS:** diagnostic research, diagnostic model, decision rule, occupational allergy, bakery workers.

## INTRODUCTION

Work in the baking industry is an established risk factor for several allergic conditions. Asthma, allergic rhinitis, conjunctivitis, and dermatitis may all result from contact with flour and enzyme allergens.(1) Although individuals with high exposure levels are more likely to have serious medical complaints and disability, even workers with intermittent or occasional exposures may be affected. Results from two cross-sectional studies performed over the last decade in the Netherlands (1996, 2001) have shown about 20% to 28% bakers sensitized to wheat flour or fungal  $\alpha$ -amylase, 35%-45% allergic complaints, and about 10% asthmatics.(2-4)

The Dutch government, responsible branch organizations (traditional and industrial bakeries, flour mills, and baking product manufacturers), and labour unions agreed to collaborate in reducing the burden of these occupational related allergic diseases. They committed on reducing the occupational exposure levels to flour dust and related allergens to feasible low levels and on implementing a health surveillance program to trace and refer sensitized bakers with allergic diseases. However, the baking industry in the Netherlands consists of about 85 industrial bakeries of which 60% have more than 50 employees, and a large number of traditional bakeries (n~3,000). Sixty percent of the traditional bakeries have less than 5 employees. In general, traditional bakeries have a poor coverage by occupational health care services. Therefore, a simple diagnostic instrument for early detection of sensitized bakers was requested that was evidence based and at low costs. The instrument should be the first step in detecting different risk groups that should be diagnosed further by occupational physicians and in specialized respiratory clinics.

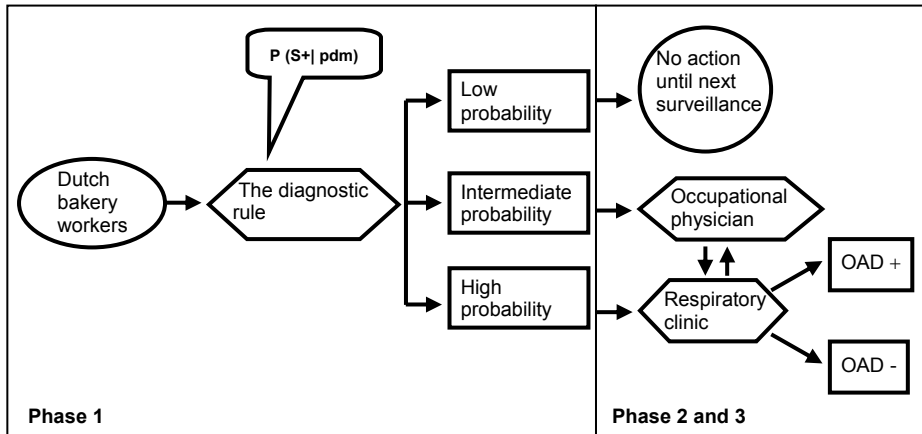
Guidelines for medical surveillance have been published and they usually include baseline pre-placement medical evaluation and testing, periodic follow-up questionnaires, interval medical examinations, and selective use of single tests.(5, 6) However, none of these approaches make use of prediction models in which personal and work related characteristics are applied to estimate the individual probability of the presence (diagnostic) or occurrence (prognostic) of an outcome that is closely related to the disease(s) of interest. Because the diagnosis of whatever allergic disease (especially bronchial asthma) can only be made at an individual level in a clinical setting, precursors of the disease of interest have to be detected first. By detecting this precursor, progress can be made in decreasing disease prevalence if a clinical diagnosis can be made earlier and more efficiently. With this approach, sequential diagnostic investigations are only required in workers with an elevated risk of having the chosen precursor and thus, leaving a considerable amount of workers in which no further medical investigations is needed.

This paper describes the overall design of three-phase nationwide respiratory allergy surveillance. Prior to surveillance, a diagnostic questionnaire model to estimate the probability of sensitisation to wheat and/or  $\alpha$ -amylase allergens was developed. This model was applied for risk stratification in 5,325 workers participated in surveillance. Detailed results from 766 randomly selected workers, in whom extensive information about respiratory and allergic disease, IgE serology, and tasks and jobs were collected, are presented. Preliminary results from clinical evaluations are reported as well.

## METHODS

### Study population and method

A three-phase nation wide surveillance program among traditional, industrial bakeries, milling industries, and baking products manufacturers has been developed according to a phased approach as was described earlier in workers exposed to laboratory animal allergens.(7)



**Figure 1** Surveillance for respiratory allergy among Dutch bakers.  $P(S+| pdm)$  is the predicted probability of occupational sensitization with the application of the diagnostic model for wheat and/or  $\alpha$ -amylase sensitization. In phase one, workers will be stratified based on their disease probability. In phase two, different advice is given to workers in each probability group. In phase three, clinical work up will be carried out by an occupational physician or a respiratory specialist. OAD: occupational allergic disease.

Figure 1 shows the phased approach:

#### 1. Risk stratification

A specially developed short questionnaire was distributed to all bakery workers that were potentially exposed to flour dust. With the results of this questionnaire and the application of the diagnostic model for sensitization to wheat and/or fungal  $\alpha$ -amylase, the individual probability of sensitization was estimated. On the basis of these calculations, every worker was classified in one of the three categories: low, intermediate and high probability (score) of being sensitised.

#### 2. Results and advice

The results of the calculated risk estimates were communicated by letter to every worker individually and their occupational physician. Workers with a low total score were told their probability to be sensitized was very low and no further action was needed; they will be enrolled in next surveillance. Workers with an intermediate as well as a high score were advised to contact occupational physicians. Workers with intermediate scores were evaluated further (more explorative clinical and work history, as well as serology) by occupational physicians. Occupational physicians have the possibility to refer workers to a specialized occupational respiratory health clinic in case a serious allergic disease is suspected that needs further clinical investigations. As for workers with a high probability of being sensitized, occupational physicians were advised to refer them directly to a specialized respiratory physician from one of the two clinics serving all bakers across the country.

#### 3. Intervention

Workers who visited the clinic underwent clinical history taking, physical examination, serology tests, spirometry, and non-specific bronchial hyper responsiveness (NSBH) to histamine. On indication peak expiratory flow recordings (PEFR) and NSBH were measured after a continuous period of at least two weeks both at and off work. Workers with an established diagnosis were referred back to their occupational physician. The objective was to develop an intervention program for these workers. The program should focus on hygienic interventions in the workplace as well as on a personal level, job

rotation and adequate medical treatment. An important goal was to not remove all affected workers directly from exposure, but to find intelligent solutions at all levels. However, this intervention program has recently started but no results can be presented yet.

### **Preparatory phase**

#### *The diagnostic model*

A diagnostic model to estimate the individual probability of sensitization to wheat and/or fungal  $\alpha$ -amylase allergens (class 2 IgE serology) was previously developed from 390 Dutch bakers who participated in an earlier survey(4), according to an approach as proposed by Meijer et al.(7) For practical application, the diagnostic model was transformed into a score chart as follow:  $Sum\ scores = (asthma\ symptom*2) + (rhinitis\ symptom*2) + (conjunctivitis\ symptom*1) + (during\ work\ symptom*1.5)$ . Each symptom is valued as 1 when present and 0 when absent. The risk stratification based on the sum scores was as follow: 0 to 1: low probability; 1.5 to 3.0: intermediate probability; 3.5 or higher: high probability of being sensitized. The model development and the score chart are described more extensively in the Appendix.

#### *The short questionnaire*

A short self-administered questionnaire was developed by incorporating the four predicting questions from the diagnostic model together with additional questions on respiratory symptoms, allergy, and bronchial hyper responsiveness(8, 9) univariably associated with the outcome (sensitization to wheat and/or  $\alpha$ -amylase). The questionnaire was also enriched with questions on absenteeism, medication use, doctor's visit, change in job title due to allergic symptoms, and smoking habit. The questionnaire was produced in a 2-pages scannable form. The first page contained basic information and job history (job title, task, percentage of time spent as bread baker, date start working in the current job, total weekly working hours, and shifts). The second page consisted of the above mentioned 19 questions.

Workers from traditional bakeries, industrial bakeries, milling industries, and baking product manufacturers in the Netherlands were invited by the branch organisations to participate in this national program for early detection of occupational allergic diseases. All companies were visited by instructed consultants who delivered the short questionnaire to the bakers. Bakers were asked to complete the questionnaire and return it by regular mail to the investigators' office.

### **Evaluation of the results**

The application of the diagnostic model was evaluated with data derived from a substudy among 340 and 28 randomly selected traditional and industrial bakeries. In this substudy, all workers in the selected bakeries were asked to complete a long questionnaire, which corresponded to the questionnaire that was used to develop the model, and blood for IgE serology was drawn. For workers who had not returned the short questionnaire earlier, an extra questionnaire was administered with items on non-participation. 766 workers participated in the first phase of surveillance as well as in the substudy and therefore, completed both the short and the long questionnaires, and had blood drawn for IgE serology. Misclassification for sensitization was easily calculated in these workers by comparing the results of IgE serology with the predicted probabilities derived from the short questionnaire.

For the clinical evaluation of the referred workers, occupational asthma (OA) was considered present when there was a history of work-related asthmatic symptoms, sensitisation to a workplace agent, bronchial hyperresponsiveness at work defined as a 20% fall in FEV<sub>1</sub> associated with a histamine dose of 2,5 mg or less (PD<sub>20</sub>) and serial PEF<sub>R</sub> showing a work-related pattern and/or at least one doubling dose increase in PD<sub>20</sub>

comparing work and away from work.(10) By now, over 100 workers have been referred to the specialized clinic and is ongoing for others. Preliminary results from the first 75 bakers who completed the examination procedures will be presented.

### **Serology**

Specific IgE antibodies against wheat flour,  $\alpha$ -amylase, and five common allergens (house dust mite, cat fur, dog fur, grass pollen, birch pollen) were measured with an earlier developed and modified enzyme immunoassay (EIA).(11) An optical density (OD) of 492 exceeding the OD +0.05 of the reagent blank (no serum control) was considered as a positive IgE serology to common allergen. An OD +0.1 or higher was defined as a positive IgE serology to wheat or  $\alpha$ -amylase allergens. In the clinic, specific IgE antibodies against wheat flour,  $\alpha$ -amylase, and five common allergens were measured with a commercial immunoassay (Pharmacia CAP system, Pharmacia Diagnostics, Sweden). Titres 0.35 kU/l or higher were considered positive sensitisation. The EIA has been compared earlier to the CAP assay for fungal  $\alpha$ -amylase and wheat allergens. For  $\alpha$ -amylase, the agreement was good across the measurement range. For wheat, the sensitivity was lower, especially at lower titres, but the specificity was high. The overall agreement was satisfactory, and good at higher titres.(11)

### **Data Analyses**

All statistical analyses were performed with SPSS 15.0 for Windows (Statistical Package for Social Sciences, Chicago Il). Prevalence rates were compared using the  $\chi^2$  (or Fischer's exact) test. Differences between mean were compared using Student's t (or Mann-Whitney) test or ANOVA. The diagnostic model was developed using stepwise backward logistic regression analysis.(12) Guidelines to develop and evaluate the diagnostic accuracy (i.e. calibration and discrimination) and internal validity of a model were described elsewhere.(12-14) Statistical significance was defined as p-value of less than 0.05 (two tailed).

## **RESULTS**

Between November 2004 and November 2006, 6714 workers from 1637 traditional bakeries and 1760 workers from 74 industrial bakeries were registered (Table 1). They cover almost all industrial and half of the traditional enterprises throughout the country. Of 1637 registered traditional bakeries, 1189 (72.7%) participated in surveillance program whereas all of the registered industrial bakeries participated. About 60% and 80% of the workers from the participating traditional and industrial bakeries, respectively, sent back the short questionnaire. In the flour milling (n=10) and baking products industries (n=13) almost sixty percent of their workers joined this program. Non-participation analysis among 86 bakers showed that one-third forgot to send back and one third said not to have received the questionnaire. The percentage of workers worried that the results could be used in a disadvantageous way by their employer was low (3%).

Most of the workers were male. Traditional bakers showed statistically lower mean sum scores and mean predicted probabilities of sensitization compared to the other industries. No statistically significant differences in observed sensitization rates were found in 766 traditional and industrial bakers who participated in the substudy. Application of the short questionnaire showed the highest rate (60.5%) of workers with low scores in traditional bakers and the highest rate (21.3%) of high scores in industrial bakers (Table 1).

Overall, 57.4% had a low probability of sensitization and was informed that further medical investigations were not required (Table 2). These workers were statistically significantly older of age and worked longer than workers with an intermediate or high score.

They also reported significantly lower rates of allergic symptoms, doctor's visit (4.4%), medication use (2.4%), or absenteeism (0.2%) than workers with intermediate and high score.

**Table 1** Questionnaire responses among different industries enrolled in the surveillance program.

	Traditional	Industrial	Milling	Baking products	Total
Registered employers n	1,637	74	-	-	1747*
Registered employees n	6,714	1,760	-	-	8474*
Participating employers n	1,189	74	10	13	1,286
Employees in participating bakeries n	5,398	1,760	686	552	8,396
Participating employees n (%)	3,214 (59.5)	1,398 (79.4)	398 (58.0)	320 (58.0)	5,325
Female	347 (10.9)	72 (5.2)	26 (6.7)	54 (17.5)	499 (9.5)
Mean sum scores (SE)	1.3** (0.03)	1.6 (0.05)	1.5 (0.09)	1.4 (0.10)	1.4 (0.03)
Low Score n (%)	1,946 (60.5)	721 (51.6)	213 (54.2)	179 (55.9)	3,059 (57.4)
Intermediate Score n (%)	706 (22.0)	379 (27.1)	109 (27.7)	88 (27.5)	1,282 (24.1)
High Score n (%)	562 (17.5)	298 (21.3)	71 (18.1)	53 (16.6)	984 (18.5)
Mean predicted probability sensitization for wheat and/or $\alpha$ -amylase (SE)	0.17** (0.002)	0.19 (0.004)	0.20 (0.007)	0.19 (0.007)	0.18 (0.002)
Observed sensitization rate (wheat and/or $\alpha$ -amylase) (SE) n=766	0.19 (0.02)	0.21 (0.02)	-	-	

Data presented as n (%) unless otherwise stated. SE: standard error.

\* Numbers not known in milling and baking products industries

\*\* Traditional bakers have a statistically significant lower predicted probability of sensitization compared to the industrial bakers.

Somewhat less than 20% (984 of 5,325) of the workers with a high probability were advised to visit the occupational respiratory health clinic. Almost 25% of the bakers in the intermediate group could reliably be informed that their probability of sensitization was elevated and that further medical investigations should be performed by their occupational physicians. About 43% of the workers in the high score group reported skin symptoms, 66% reported atopic symptoms, 20% reported bronchial hyper responsiveness, 48% used medication for respiratory problems, 39 % visited their physician for allergic complaints, 8 % absenteeism and 9 % changed job due to allergic symptoms (Table 2).

Table 3 outlines the results from 766 substudy participants who completed a more extensive questionnaire and in whom IgE serology to wheat flour,  $\alpha$ -amylase and five common allergens was measured. The mean predicted probability of sensitization to wheat and/or fungal  $\alpha$ -amylase in all bakers was lower (18.0%) than the observed sensitization rate (20.4%), at a borderline statistically significant level ( $p=0.07$ ). However, the predicted sensitization rate to wheat and/or fungal  $\alpha$ -amylase in the high score group was higher (43%) than the observed rate (39.7%). The wheat sensitization rate was 31% in bakers with a high score, indicating the lesser importance of  $\alpha$ -amylase sensitization (17.2%). In the low

score group the predicted rate (9%) of sensitization to wheat and/or  $\alpha$ -amylase was significantly lower than the observed sensitization rate (12.7%).

**Table 2** General characteristics and questionnaire responses across low, intermediate and high score groups

	Low score ( $\leq 1.0$ )	Intermediate Score (1.5 - 3.0)	High score ( $\geq 3.5$ )	Total
<b>Questionnaire n (%)*</b>	3,059 (57.4)	1,282 (24.1)	984 (18.5)	5,325
<b>General characteristics</b>				
Age mean (SE)	40.3 (0.2)**	37.7 (0.3)	38.6 (0.4)	39.3 (0.2)
Female	299 (9.9)	118 (9.3)	82 (8.4)	499 (9.5)
Work duration mean (SE)	13.8 (0.2)**	11.7 (0.3)	12.9 (0.3)	13.1 (0.14)
<b>Questionnaire responses</b>				
Skin symptoms in the last 12 months	380 (12.6)	369 (29.6)	407 (42.6)	1156 (22.1)
Respiratory or eye/nose symptoms in contact with common allergens <sup>†</sup>	209 (6.8)	448 (35.2)	642 (65.8)	1299 (24.5)
Symptoms suggestive for BHR <sup>‡‡</sup>	24 (.8)	52 (4.1)	144 (14.9)	220 (4.2)
Use of medication to improve respiratory complaints in the last 12 months (e.g. inhalants)	74 (2.4)	186 (14.6)	468 (48.1)	728 (13.7)
Doctor visit for allergic complaints in the last 12 months	134 (4.4)	197 (15.5)	379 (38.7)	710 (13.4)
Absenteeism due to allergic symptoms in the last 12 months	6 (0.2)	34 (2.7)	79 (8.2)	119 (2.2)
Change of function or task due to respiratory symptoms	16 (0.5)	19 (1.5)	83 (8.5)	118 (2.2)

Data presented as n (%) unless otherwise stated. SE: standard error.

\* Proportion of all workers (n=5,325)

\*\* Statistically significantly ( $p < 0.05$ ) different from the intermediate and high score group

<sup>†</sup> Considered present if experienced respiratory problem in contact with dust, plants, or pets; or after ingestion of certain food.

<sup>‡‡</sup> Considered present if experienced respiratory problem due to change in temperature, mist, and/or cooking smell.

**Table 3** Results of IgE serology across different score groups from the substudy in traditional and industrial bakers

	Low score ( $\leq 1.0$ )	Intermediate Score (1.5 - 3.0)	High score ( $\geq 3.5$ )	Total
n(%)*	387 (50.5)	205 (26.8)	174 (22.7)	766 (760**)
IgE sensitisation to any of 5 common allergens (HDM, cat fur, dog fur, grass, berch)	85 (22.1)	85 (41.9)	99 (57.6)	269 (35.4)
HDM sensitization	64 (16.6)	57 (28.1)	65 (37.8)	186 (24.5)
Total IgE > 100 kU/L	79 (20.6)	58 (28.7)	70 (40.9)	207 (27.4)
IgE sensitisation to wheat allergens	21 (5.4)	23 (11.2)	54 (31.0)	98 (12.7)
IgE sensitisation to $\alpha$ -amylase allergens	31 (8.0)	25 (12.2)	30 (17.2)	86(11.2)
IgE sensitisation to wheat and/or $\alpha$ -amylase allergens	49 (12.7)	38 (18.5)	69 (39.7)	156 (20.4)
Mean predicted probability sensitization to wheat and/or $\alpha$ -amylase allergens (SE)	0.09 (0.000)	0.20 (0.001)	0.43 (0.004)	0.18 (0.002)

Data presented as n (%) unless otherwise stated. SE: standard error; HDM: House dust mite.

\* Proportion of all workers (n=766)

\*\* n=760 for common allergens

### Clinical Findings

Preliminary clinical results of 75 referred bakers (70 traditional and 4 industrial bakers) showed 40% wheat and 10%  $\alpha$ -amylase sensitized individuals (Table 4). Rye serology was only assessed in workers who expressed that they experienced symptoms when exposed to rye (n=14). Eight out of 14 (57%) bakers were sensitized to rye allergen. Of these 8 workers, 6 (75%) were also sensitized to wheat allergens. Seventy three percent (n= 55) of the referred individuals were bread bakers or bread baker and confectioner as well.

**Table 4** Preliminary results of clinical investigations of 75 referred bakers

	Low score ( $\leq 1.0$ )	Intermediate Score (1.5 - 3.0)	High score ( $\geq 3.5$ )	Total
n	2	25	47	75
Female	0	1 (4.0)	5 (10.6)	6 (8.0)
Bread baker	0 (0.0)	9 (34.6)	17 (65.4)	26
Confectioner	1 (5.3)	8 (28.6)	10 (52.6)	19
Bread baker and confectioner	1 (3.6)	8 (28.6)	19 (67.9)	28
IgE sensitisation				
Wheat	1/2 (50.0)	9/25 (36.0)	19/46 (41.3)	29/73 (39.7)
$\alpha$ -amylase	0 (0.0)	0 (0.0)	7 (15.2)	7/73 (9.6)
Rye (Only in rye exposed workers. n=14)	0/1 (0.0)	2/3 (67.0)	6/10 (60.0)	8/14 (57.1)
Soya	0 (0.0)	2 (8.0)	1 (2.2)	3/73 (4.1)
Common allergens	2/2 (100.0)	8/25 (32.0)	24/46 (52.2)	34/73 (46.6)
Non specific bronchial hyper responsiveness at work				
Histamine PD20 2,5 mg or less	0/2 (0.0)	8/25 (32.0)	20/42 (47.6)	28/69 (40.6)
Clinical diagnoses (n= 54)				
Clinical diagnosis of asthma (all causes)	0/2 (0.0)	5/18 (27.8)	20/34 (58.8)	25/54 (46.3)
Occupational asthma	-	2/5 (40.0)	7/20 (35.0)	9/25 (36.0)
Rhino-conjunctivitis (all causes)	1/2 (50.0)	16/18 (88.9)	29/34 (85.3)	46/54 (85.2)
Work-related allergic rhinitis	1/1 (100.0)	9/14 (64.3)	18/29 (62.1)	29/45 (64.4)

Data presented as n (%)

A clinical diagnosis of an allergic disease was set in 54 of the 75 referred individuals (73%). In 21 individuals, other (respiratory) diseases were diagnosed with so far unknown work relatedness. Allergic rhino-conjunctivitis was diagnosed in 85% (46/54) of the individuals of whom 63.0% (29/46) reported their rhinitis to be associated with their work. Work related respiratory symptoms were mostly reported in the high score group (52.2%), while other work-related symptoms were equally spread over intermediate and high score groups (data not shown). Asthma was diagnosed in 25 (46.3%) bakers, of whom 9 (36%) bakers showed to have occupational asthma; most of them (90%) were detected in the high probability group. Six out of 25 asthma cases (24%) could be assigned as having work-aggravated asthma.

The evaluation of dermatological problems was only minimally assessed and concentrated on whether or not a serious dermatitis was present at the time of clinic visit. Nevertheless, 10 patients showed some type of dermatitis. No additional dermatological testing was performed in these workers. Only 3 workers (4%) showed no actual allergic disease, such as rhino-conjunctivitis, asthma, or dermatitis. One of them had a history of asthma, and one suffered from hyperventilation.

## DISCUSSION

In this occupational respiratory allergy surveillance among 5,325 bakers in the Netherlands, a short questionnaire, containing four diagnostic predictors for sensitization to wheat and/or fungal  $\alpha$ -amylase allergens was used for medical decision making. The questionnaire aimed at early detection of workers sensitized to wheat or  $\alpha$ -amylase allergens. The results show that bakery workers can be categorized effectively into three groups with different risks for occupational allergic diseases: a high risk group in which a detailed clinical evaluation was needed to set a diagnosis of (occupational) allergy more accurately, an intermediate risk group in which medical follow up by occupational physicians was essential for health protection and medical follow-up, and a low risk group comprising about 60% of the population in which medical investigations could be held back.

Prediction models for various outcomes have been developed in the last decades regularly in clinical practice and are increasingly being used to predict the presence or occurrence of a disease or outcome.(15-17) These models can be used to rationalize the process of decision making for individual patients or to stratify patients by disease severity to create risk groups. Surprisingly, in occupational medicine only a few prediction models have been developed recently for early estimation of the presence of an occupational disease.(7, 18-22) In this study we report the results of the application of a diagnostic model, the stratification into 3 risk groups, and of the referral to specialized occupational respiratory clinics.

### Diagnostic phase

Given the large number of workers spread all over the country, we presented the questionnaire to the individual bakers by trained consultants visiting the bakery. With this approach we tried to moderate non-response. The results of the questionnaire were transformed into total scores to predict more accurately the presence of sensitization in every individual worker. The application of the scoring rule enabled to exclude workers with a low probability from further medical investigations and concentrate advanced medical tests on workers with a high probability. This approach led to a substantial reduction of advanced medical tests to only 22% of the total investigated workers. Therefore, the choice of cut off points to classify the workers into high, intermediate, or low probability groups is crucial and determines referral policy and misclassification rate. A balance must be sought between an acceptable proportion of missed cases and unnecessary referrals. We used a cut-off point of 1.5 that corresponds to an overall sensitization rate to wheat and/or fungal amylase of 20% in the total population and classified these individuals below this score into the low score group.(4) Serology results from the substudy showed that 40% (69/174) workers with a high score were eventually sensitized, and consequently 60% not sensitized. However, misclassification of workers wrongly assigned as "diseased" (sensitized) is less dramatic because these workers for the most part show substantial medical problems that need to be diagnosed further.

On the other hand, there are also workers with a low score that are wrongly classified as "non-diseased" (not sensitized). This is a serious problem that cannot be denied. Although the low score group was a large group (almost 60% of all workers) and was told to have a low probability of being sensitized, still 49 (12.7%) of them were sensitized to wheat and/or fungal  $\alpha$ -amylase. This means that 31% of all sensitized workers were missed by the decision rule. Extrapolation to workers in the low score group leads to probably 394 (12.7% x 3,059) sensitized individuals that are not captured by the decision rule; this is a substantial number of workers. However these workers show considerably less doctor visits, medication use, absenteeism, and change in job title caused by allergic symptoms and thus are likely to have less severe sensitization. These workers were statistically significantly older and worked longer in their present job than workers in the

intermediate and high score groups. It appears that these bakers, although sensitized, remain in their jobs without having serious symptoms for already quite some time.

By contrast, 70% of the sensitized bakers come from the intermediate and high score groups. Extrapolation to the intermediate and high score group leads to 634 (28% x 2,266) probably sensitized individuals that can be captured for further medical investigations. Furthermore, workers in the high score group showed 10 to 20 times higher rates of symptoms suggestive for BHR, allergic symptoms, doctor's visits, medication use, absenteeism, and change in job title due to allergic symptoms than their colleagues in the low score group. Bakers in the intermediate score group worked statistically significantly shorter in their present job. This might be an indication that these workers may be more likely to leave their job earlier, which needs to be confirmed in follow-up studies.

### **Clinical evaluation**

The results of the clinical evaluation in 75 referred bakers show that the group with a high probability of sensitisation to wheat and/or  $\alpha$ -amylase had high rates of allergic diseases, especially rhino-conjunctivitis (85%) and asthma (59%). Occupational asthma was diagnosed in 35% (7/20) of the asthmatics in the high score group, and in 40% (2/5) of the intermediate score group.

Work-aggravated asthma was detected in 24% asthmatics in the high score group (data not shown). Nevertheless, PEFr's could be performed in only a minority (n=9) of asthma patients because of constraints put by the employer. Therefore, cases of work-exacerbated asthma may have been missed so far. This is reported by Bolen et al(23) who described that self-reported symptoms and use of medication fail to identify work-related asthma exacerbations as determined by serial peak exploratory flow measurements.

Rhino-conjunctivitis was a common problem: it was found in 85% of workers in the intermediate and high probability groups. A clear association with work could be established in 64%, indicating a health problem that needs full hygienic attention by advising about control measures, working practices, and cleaning processes. Besides, 20 out of 47 patients (42.6%) with rhino-conjunctivitis were found to have coexisting asthma. Therefore, medical attention should focus on accurate diagnosing and adequate prescription of medication.

### **General reflection**

A diagnostic rule was used as a decision tool to generate case specific advice, to support decision making about individual workers (patients) by health professionals, the workers themselves or others concerned about them. It was developed with a clear relevance to the "bakers world" in which a clinical need for early and accurate diagnoses of allergic diseases was urgently needed. Although the model was developed on the best available knowledge, derived from a well-designed study, failures and difficulties for various reasons can be expected. One of these problems concerns the misclassification of the outcome of interest and non-detected diseased workers.

One might wonder whether it is acceptable to miss 394 probably sensitised individuals out of 5,325 participating bakers at the gain of minimizing the number of individuals to be evaluated by occupational physicians or to be referred to the respiratory clinic. This 60% reduction comprised about 3,000 individuals in whom no further medical investigations were advised. Although, they include the 394 sensitized workers, the reported symptoms are significantly less severe than workers with higher scores who were advised to participate in a medical follow up. So it may be concluded that they belong to a group of workers sensitized to wheat and/or  $\alpha$ -amylase allergens with less serious health problems that did not influence their work capability harmfully at that moment.

To appreciate the level of misclassification, we also analyzed how many clinical asthma cases would be missed if we excluded workers with a low score from the sequential medical work up. Clinical asthma was considered present if workers gave positive answers to at least two of the following questions: "Did you wheeze the last 12 months?", "Are you

allergic or have you been allergic to one or more allergens?”, “Do you experience breathing problems when you are exposed to baking and frying smells?”, and “Did you use medication to improve your breathing problems (inhalants, aerosols, pills) in the last 12 months?”. An analysis among the clinical cases showed that responses to these questionnaire items were strongly associated with a diagnosis of asthma. We estimated that 51 asthma cases would be present in the low score group ( $51/3,059 = 1.7\%$ ). This potential non-detection of asthmatics is a serious problem, because asthma is an already established disease that should properly be treated. We do not have an adequate solution to this problem other than to invite these workers individually to visit their occupational physician and refer them to the respiratory health clinic. In the near future, additional diagnostic models, for different respiratory endpoints may be used in parallel, or our model for sensitization may be replaced by other more advanced models. However, such models will need to be developed on the basis of information collected in this system.

In conclusion, this national surveillance program for occupational respiratory allergy among bakers in the Netherlands showed that early detection of sensitization to wheat allergens can be improved and formalized by using an approach based on diagnostic research. Workers with different risks of having an (occupational) allergic disease can effectively be identified by using sensitization to wheat and/or fungal amylase allergens as a precursor of occupational allergic diseases. The questionnaire based prediction model enabled selecting different risk groups in a large scale at presumably low expenses. By selecting 50% of bakery workers with a low risk of sensitization, more advanced medical investigations can be held back in this group. The probably missed sensitized workers may be detected in a next round. It simultaneously showed that detection of subgroups of workers with an intermediate or high probability of sensitization improved the expected benefit of diagnosis and treatment and outweighed the expected potential harm of no diagnosis and no treatment.

## REFERENCES

1. Bernstein DI. Allergic reactions to workplace allergens. *Jama* 1997;278(22):1907-13.
2. Houba R, Heederik D, Doekes G. Wheat sensitization and work-related symptoms in the baking industry are preventable. An epidemiologic study. *Am J Respir Crit Care Med* 1998;158(5 Pt 1):1499-503.
3. Houba R, Heederik DJ, Doekes G, van Run PE. Exposure-sensitization relationship for alpha-amylase allergens in the baking industry. *Am J Respir Crit Care Med* 1996;154(1):130-6.
4. Oostenbrink JH TJ, Tempels Z, Heide S, Steketee HA, Kerkhof M, Monchy JGR. Aard en omvang van beroepsgebonden klachten bij werknemers in bakkerijen, meelfabrieken en grondstoffenindustrie. Groningen: Academisch Ziekenhuis Groningen; 2002 24 Dec 2002.
5. Cullinan P, Tarlo S, Nemery B. The prevention of occupational asthma. *Eur Respir J* 2003;22(5):853-60.
6. Nicholson PJ, Cullinan P, Taylor AJ, Burge PS, Boyle C. Evidence based guidelines for the prevention, identification, and management of occupational asthma. *Occup Environ Med* 2005;62(5):290-9.
7. 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.
8. 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.
9. van der Lende R, Orie NG. The MRC-ECCS questionnaire on respiratory symptoms (use in epidemiology). *Scand J Respir Dis* 1972;53(4):218-26.
10. Bernstein DI, Campo P, Baur X. Clinical assessment and management of occupational asthma. In: Bernstein IL, Chan-Yeung M, Malo J, Bernstein DI, editors. *Asthma in the Workplace*. 3rd ed. New York: Taylor & Francis; 2006. p. 172.
11. Doekes G, Douwes J, Wouters I, de Wind S, Houba R, Hollander A. Enzyme immunoassays for total and allergen specific IgE in population studies. *Occup Environ Med* 1996;53(1):63-70.

12. 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.
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. Hosmer D, Lemeshow S. *Applied logistic regression*. New York: John Wiley and Sons, Inc; 1989.
15. Wells PS, Anderson DR, Bormanis J, Guy F, Mitchell M, Gray L, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet* 1997;350(9094):1795-8.
16. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97(18):1837-47.
17. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PW, Levy D. Profile for estimating risk of heart failure. *Arch Intern Med* 1999;159(11):1197-204.
18. 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.
19. Suarathana 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.
20. Kuijpers T, van der Windt DA, van der Heijden GJ, Twisk JW, Vergouwe Y, Bouter LM. A prediction rule for shoulder pain related sick leave: a prospective cohort study. *BMC Musculoskelet Disord* 2006;7:97.
21. Duijts SF, Kant IJ, Landeweerd JA, Swaen GM. Prediction of sickness absence: development of a screening instrument. *Occup Environ Med* 2006;63(8):564-9.
22. Suarathana E, Moons KG, Heederik D, Meijer E. A simple diagnostic model for ruling out pneumoconiosis among construction workers. *Occup Environ Med* 2007;64(9):595-601.
23. Bolen AR, Henneberger PK, Liang X, Sama SR, Preusse PA, Rosiello RA, et al. The validation of work-related self-reported asthma exacerbation. *Occup Environ Med* 2007;64(5):343-8.

## APPENDIX

### Development of the diagnostic model

A diagnostic model was derived from data from a cross sectional study in 390 Dutch bakery workers. Workers were asked to complete a self-administered questionnaire which was derived from the IUATLD (ref. 8) and the MRC-ECCS (ref. 9). The questionnaire consisted of employment (job, tasks) data, history of lower and upper respiratory symptoms, allergic symptoms due to common allergens, symptoms suggesting bronchial hyperresponsiveness, work-related upper and lower respiratory symptoms, skin symptoms, absenteeism, medication use, changes in tasks or jobs, and smoking habits. Informed consent was obtained from all participating workers. Subsequently blood was drawn for serological testing to detect specific IgE antibodies against wheat and  $\alpha$ -amylase allergens. Specific IgE antibodies to allergens were measured with a commercial immunoassay (Pharmacia CAP system, Pharmacia Diagnostics, Sweden). Sensitization was defined as class II ( $\geq 0.7$  kU/l) positive IgE serology.

No exposure measurements were included in the model development because of practical reasons. Multivariable regression analysis with backward stepwise selection was done to develop the model. This procedure was followed by bootstrapping of the final model to adjust for over-optimism. Guidelines on the development and evaluation of prediction models are described by Harrell (ref. 12). The final diagnostic model consisted of four questionnaire items and is presented in the following table. The model showed good calibration (H-L test p-value=0.748) and reasonable discrimination (ROC area 0.73 (0.67 to 0.79)).

Scoring rule for sensitization to wheat allergen and the corresponding predicted probability

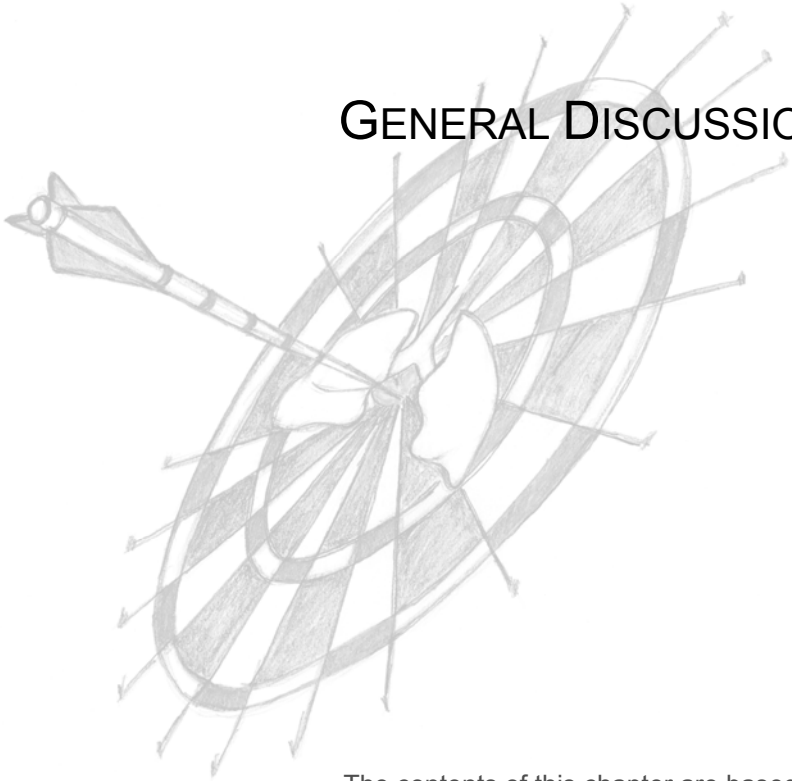
Predictors	Answer	Score
"Have you ever had asthma in the past 12 months?"	If yes	2
"Have you ever had allergic rhinitis including hay-fever?"	If yes	2
"Have you ever had itchy and/or red eyes in the past 12 months?"	If yes	1
"Do you experience more of the following symptoms during work: shortness of breath, chest tightness, itchy eyes, itchy nose, and/or sneezing?"	If yes	1.5
	Sum scores	Max 6.5
<b>Sum score</b>	<b>0</b>	<b>1</b>
	<b>2</b>	<b>3.5</b>
	<b>4.5</b>	<b>5.5</b>
	<b>6.5</b>	
Predicted probability (%)	9	14
	20	31
	42	53
	64	

To calculate the predicted probability of sensitization:  $P(\text{sensitization}) = 1 / (1 + \exp(-(-2.32 + 0.92 * \text{asthma} + 0.90 * \text{rhinitis} + 0.46 * \text{conjunctivitis} + 0.62 * \text{during work symptom})))$



# CHAPTER 6

## GENERAL DISCUSSION



The contents of this chapter are based on  
E. Suarhana, E. Meijer, D.E. Grobbee, D. Heederik  
Predicting occupational diseases  
*Submitted*



---

# PREDICTING OCCUPATIONAL DISEASES

---

EVA SUARTHANA, EVERT MEIJER, DIEDERICK E. GROBBEE, DICK HEEDERIK

---

Prediction research is relatively new to the occupational health field(1-6) while it is well established to support decision making in clinical medicine(7-9). In prediction research, prediction models are developed to estimate the individual probability of the presence (diagnostic model) or future occurrence (prognostic model) of an outcome (i.e. disease). These models are used to assist in clinical decisions for individual patients, or to stratify patients into risk groups for different categories of disease severity.(10, 11) In daily practice physicians produce implicit probability estimates using their clinical knowledge and experience in combination with tests results to diagnose a disease. A diagnostic model improves the diagnosis by allowing objective and standardized quantification of the individual probability of having a disease while avoiding as much as possible the use of (invasive) advanced and costly reference test.

As an example from clinical practice, Wells and colleagues developed a diagnostic model that comprised of patients' history and physical examination to predict the prior test probability of deep vein thrombosis (DVT) in secondary care patients. They demonstrated that the use of this model in combination with impedance plethysmography (IPG) can safely rule out the presence of DVT. This approach largely reduced patients' burden and health care cost by withholding a costly venography.(7) Prediction of a 10-years risk of coronary heart disease (CHD) using the Framingham scores is a well known example of prognostic prediction.(8) A simplified model comprised of gender, age, smoking habit, diabetes status, lipid profile, and blood pressure can accurately predict a later CHD in outpatients free of disease. Such prediction allows physician to take preventive action in order to avert CHD in patients with a high probability.

As an analogue, diagnostic prediction in the field of occupational health will enable the (occupational) physician to accurately predict in a cost-effective manner the probability of occupational diseases for every individual worker in a large population and stratify them into different risk groups. As an example, in the traditional approach to diagnose workers with silicosis, a series of diagnostic tests (i.e. extensive questionnaires, chest X-rays, and lung function evaluations) will be administered to any worker exposed to silica dust in a periodic surveillance.(12-14) In a recent study among construction workers with a low prevalence of early signs of disease (2.9% of these workers had a positive X-ray indicative of pneumoconiosis)(15), a diagnostic model could identify accurately workers with a low probability of pneumoconiosis and rule them out from further investigations.(6) On the basis of the prevalence it could be expected beforehand that there will be a high number of negative X-ray reading results and thus, application of the model improved diagnostic efficiency over no model.(6)

## Issues in Model Development

When developing prediction models, several issues of concern exist. The first consideration is the choice of the relevant outcome of interest. With regard to occupational allergy, the development of the prediction models for sensitization to occupational allergens is controversial because sensitization is not considered a disease. However, sensitized subjects, when continuously exposed to such allergens, are at high risk of developing severe occupational allergic diseases (OAD) such as occupational asthma.(16) Therefore, with the development of diagnostic models we aimed at early identification of individuals with

a high probability of having or developing OAD.(2, 3) The same explanation can be given to the development of the diagnostic model for pneumoconiosis: to detect silicosis in early phase.(6)

The second consideration is the choice of predictors to be included in the model. In contrast to etiologic studies, prediction studies utilize the associations between predictors and the outcome to estimate the probability of having or developing an outcome and are not aiming at explaining causal associations.(10, 11) Predictors might be obtained from a standard anamnesis, a physical examination, or additional testing. In the diagnostic model for pneumoconiosis, being a “current smoker” was selected as important independent predictor in the model with an odds ratio (OR) of 2.4. The OR means that current smokers, compared to non-smokers, have 2.4 times higher probability of having a positive chest X-ray, without any reference to causality between smoking and pneumoconiosis.(6) Furthermore, time sequence (a cause should precede an effect) is not a requirement in diagnostic studies. These studies are inherently cross-sectional and most predictors are actually the consequence of disease presence. In the same pneumoconiosis model, “feeling unhealthy” was also selected as a rather strong predictor with an OR of 2.8. Etiologically seen, having pneumoconiosis is expected to precede the presence of feeling unhealthy. In a diagnostic perspective, an unhealthy feeling was strongly associated and considerably increased the probability of having an X-ray indicative for pneumoconiosis.(6)

In practice, no diagnosis is established by a single test result and each test result is judged together with other (previous) test results. Therefore, multivariable regression analysis is used to evaluate the diagnostic or prognostic value of every single test independent to the presence of other predictors.(17) To stay close to the diagnostic sequence in practice, one would start modelling with candidate predictors from questionnaires and then evaluate if adding information from other tests can improve the performance of the questionnaire model. The inclusion of other tests into the model will be strongly dependent on the availability of resources (e.g. funding, skilled nurses and/or technicians, series of tests under study) as well as practicability and feasibility of performing such tests on a large scale and routine basis. Depending on the health care and insurance systems, more or less invasive tests may or may not be considered acceptable in an occupational health care setting. For example, a questionnaire based prognostic model for OAD might be sufficient for job and career counselling among trainees exposed to HMW allergens, while a more burdening or invasive test (i.e. bronchial challenge tests) might not be acceptable for this purpose at the time disease is still absent.(18)

### **Issues in Model Application**

After a model is developed, it is important to evaluate how good it performs in a new, but related, population.(19-21) In general, prediction models show a lower performance in populations other than where the model was derived from. For example, the DVT model developed in secondary care patients can not simply be generalized to primary care patients due to the different domain.(22) For that reason, external validation of the model is necessary to address its generalizability before one can use it with confidence.(19-21)

After assuring that a model is valid and produce accurate predictions, we can transform it into an easy-to-use score chart or nomogram to facilitate its use in practice. Nevertheless, when a computerized recording system is available, a prediction model can easily be incorporated to automatically produce the predicted probability of a disease.

The next important issue in model application is to determine (a) probability thresholds to stratify individuals into risk categories. This determination will influence misclassification on the population level: How many healthy individual will be falsely assigned to high risk group (false positive) and how many diseased individuals will be falsely assigned to low risk group (false negative). In general a higher threshold leads to a lower percentage of high risk group; the specificity is higher (lower false positive rate), but at the cost of lower sensitivity (higher false negative rate). The main motive for diagnostic research

in occupational health is to increase the efficiency of surveillance by avoiding unnecessary tests from workers with a low probability of having occupational disease. Nevertheless, the choice for a threshold must be based on the balance between the proportion of missed cases and reduction of unnecessary diagnostic tests. It will also depend on the severity of the disease, disease prognosis when missed, or alternatively the improvement of prognosis when detected. When the outcome is a slowly progressing non-lethal disease (i.e. silicosis), a decision maker may choose a high threshold by speculating that the missed cases can be captured in next round of surveillance.

A careful consideration of ethical and legal aspect should also be addressed when applying a prediction model in occupational practice. In most cases, one should limit application of prognostic models for prevention purposes in the workplace. For example, a model that predicts the incidence of occupational sensitization may allow an occupational physician to prevent the development of OAD. With the help of occupational hygienists, further exposure to the relevant occupational exposure among workers with a high probability can be limited (i.e. job rotation and enforcement of the use of (personal) protective equipment). Nevertheless, employers might want to use such model in pre-employment selection or make a progressive intervention such as change of job without signs of disease. Unless the model produced an excellent prediction, this action is ethically and legally unacceptable since we are imposing decisions on individuals who are disease free and might not develop disease in the future.

### **The Future of Prediction Modelling in Occupational Health**

In conclusion, prediction modelling is a novel approach in occupational health that can strongly enhance early detection of occupational diseases. Nevertheless, investigators should carefully design the model and clearly state the context where and how it can be used. In the field of occupational respiratory allergic diseases, the opportunity to develop diverse prediction models in populations exposed to high- and/or low-molecular weight agents is wide open. Recent study illustrated that it is possible to develop a generic model for sensitization to high molecular weight allergens in bakery and laboratory animal workers.<sup>(3)</sup> Nevertheless, it remains an open question if we can develop one universal model for occupational sensitization and symptoms. Future studies are also needed to evaluate the cost-effectiveness of the application of these models in occupational health practice more explicitly.

### **REFERENCES**

1. 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.
2. 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.
3. Suarhana 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. Kuijpers T, van der Windt DA, van der Heijden GJ, Twisk JW, Vergouwe Y, Bouter LM. A prediction rule for shoulder pain related sick leave: a prospective cohort study. *BMC Musculoskelet Disord* 2006;7:97.
5. Duijts SF, Kant IJ, Landeweerd JA, Swaen GM. Prediction of sickness absence: development of a screening instrument. *Occup Environ Med* 2006;63(8):564-9.
6. Suarhana E, Moons KG, Heederik D, Meijer E. A simple diagnostic model for ruling out pneumoconiosis among construction workers. *Occup Environ Med* 2007;64(9):595-601.
7. Wells PS, Anderson DR, Bormanis J, Guy F, Mitchell M, Gray L, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet* 1997;350(9094):1795-8.

8. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97(18):1837-47.
9. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PW, Levy D. Profile for estimating risk of heart failure. *Arch Intern Med* 1999;159(11):1197-204.
10. 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.
11. Moons KG, Biesheuvel CJ, Grobbee DE. Test research versus diagnostic research. *Clin Chem* 2004;50(3):473-6.
12. ILO. Technical and ethical guidelines for workers' health surveillance. Geneva: ILO; 1997.
13. NIOSH. Occupational Respiratory Disease Surveillance. 2004 2004 [cited 2005; Available from: <http://www.cdc.gov/niosh/topics/surveillance/ords/StateBasedSurveillance/SENSORSilicosis.htm#A>
14. ACOEM. Evidence based statements: Medical surveillance of workers exposed to crystalline silica. 2005 2005 [cited 2005; Available from: <http://www.acoem.org/guidelines/article.asp?ID=82>
15. Tjoe Nij E, Burdorf A, Parker J, Attfield M, van Duivenbooden C, Heederik D. Radiographic abnormalities among construction workers exposed to quartz containing dust. *Occup Environ Med* 2003;60(6):410-7.
16. Bernstein DI. Allergic reactions to workplace allergens. *Jama* 1997;278(22):1907-13.
17. Moons KGM, Grobbee DE. Diagnostic studies as multivariable, prediction research. *J. Epidemiol. Community Health* 2002;56:337 - 338.
18. Suarhana E, Malo J, Heederik D, Ghezzi H, Gautrin D. Which tools best predict the incidence of occupational IgE sensitization and respiratory symptoms? Predictors of occupational allergy. Submitted 2008.
19. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130(6):515-24.
20. 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;56:826-832.
21. Toll D, Janssen K, Vergouwe Y, Moons K. Validation, updating and impact of clinical prediction rules: a review. *Journal of Clinical Epidemiology* 2007; Accepted for publication.
22. Oudega R, Hoes AW, Toll DB, Moons KG. The value of clinical findings and D-dimer tests in diagnosing deep vein thrombosis in primary care. *Semin Thromb Hemost* 2006;32(7):673-7.

# SUMMARY

Prediction research is relatively new to the occupational health field while it is well established to support decision making in clinical medicine. In prediction research, prediction models are developed to estimate an individual's probability of the presence or future likelihood of occurrence of an outcome (i.e. disease of interest or its related condition). These models are used to assist clinical decision making for individuals, or to stratify individuals into risk groups with different likelihood for developing disease or disease severity. Prediction models enable objective and standardized quantification of the probability of having or developing a disease without performing (invasive) advanced and costly reference test. This thesis aims at demonstrating the development, validation, and application of prediction models for occupational lung diseases.

All models described in this thesis were developed using multivariable logistic regression analysis with backward stepwise selection. The agreement between the predicted probabilities and the observed frequencies of the outcome (calibration) was evaluated. The model's ability to discriminate individuals with and without the outcome (discrimination) was assessed by the ROC area. Bootstrapping was used to assess the internal validity of the model. This procedure produces a ROC area corrected for optimism and a shrinkage factor. The regression coefficients of the predictors in the model were multiplied by the shrinkage factor to prevent over optimism when the model applied to new subjects. Finally, to facilitate their application in practice, all models were transformed into easy-to-use scoring rules.

In **chapter 2.1** we developed and validated a diagnostic questionnaire model to predict sensitization to wheat allergens in bakery workers. Six predictors composed the model: type of bakery, nasoconjunctival symptoms in the last 12 months, asthma symptoms in the last 12 months, shortness of breath and wheeze, work-related upper and lower respiratory symptoms. The model showed a good discrimination (ROC area 0.76) and internal validity (the shrinkage factor was 0.89 and the corrected ROC area was 0.75). External validation demonstrated a reasonable discrimination (ROC area 0.69), but poor calibration. Adjustment of the intercept improved the model's calibration. In its form as an easy-to-use scoring rule, the model enables manual calculation of the individual probability of being sensitized to wheat allergens. Our findings support the idea that simple questionnaires can reasonably predict the presence of wheat sensitization, and can be used in all bakery workers exposed to wheat flour.

In **chapter 2.2** we presented the development of a generic diagnostic model for sensitization to high molecular weight (HMW) allergens using simple questionnaire items and routine laboratory tests. The model was developed in pooled data from Dutch laboratory animal (LA) and bakery workers. The final model included the predictors: number of working hours/week, work-related respiratory symptoms, total IgE, and IgE to common allergen. We found significant interactions between the type of work and other predictors, which resulted in different scores for LA workers and bakers. The discriminative ability of the model in the in laboratory animal was 0.80, whereas in bakers it was 0.70. Overall, the internal validity of the model was good (optimism-corrected AUC of 0.76 and a shrinkage factor of 0.95). External validation in British LA workers showed that the model was satisfactory calibrated and discriminated workers at high and low risk of being sensitized (ROC area 0.76). We concluded that it was possible to develop a generic model for prediction of sensitization to occupational HMW allergens. However, the weighing of predictors differed across specific working environments.

**Chapter 2.3** described the development a diagnostic model to estimate the probability of having pneumoconiosis, defined as chest X-ray with ILO profusion category  $\geq 1/1$ . The model was derived from a cross sectional study among 1291 Dutch natural stone and construction workers with potentially high quartz dust exposure. Age 40 years or older, current smoker, high exposed job, working 15 years or longer in the construction industry, 'feeling unhealthy', and standardized residual FEV1  $\leq -1.0$  were the independent predictors in the diagnostic model. The accuracy (calibration and discrimination) of the model was good (ROC area 0.81). After correction for optimism the ROC area was 0.76. This model was aimed to accurately rule out workers with a low risk of having pneumoconiosis. By choosing a cut-off point with a high negative predictive value the occupational physician can eventually detect this large proportion and exclude them from unnecessary chest X-ray investigations. Therefore, its use might improve the efficiency of pneumoconiosis detection, especially in a working population with a low prevalence.

In **chapter 3** we used baseline responses to a questionnaire, skin-prick tests (SPT) to common allergens, and bronchial responsiveness (BR) testing with methacholine to develop prognostic models to predict work-related sensitization and respiratory symptoms in animal health apprentices, after 32 months of training. Four 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. Symptoms indicative for asthma and allergic symptoms at baseline composed the final questionnaire model for the occurrence of occupational sensitization and symptoms. The questionnaire models for both outcomes showed good discrimination (ROC areas were 0.73 and 0.78, respectively) and calibration. Addition of SPT and/or BR testing increased the specificity of the questionnaire model to predict sensitization, but not for symptoms at work. Therefore, we concluded that the questionnaire model was a good tool to predict the incidence of occupational sensitization and symptoms. Additional tests improved the specificity of the prediction model for LA sensitization.

In **chapter 4** we assessed the validity and explored the possibility to update an existing diagnostic questionnaire model for detection of sensitization to LA allergens. The model was developed in Dutch LA workers and externally validated in Canadian animal health apprentices. Several approaches were used to externally validate the model: (1) no update; (2) update of the intercept of the model; (3) update of the intercept and all regression coefficients; and (4) revision of the model by excluding existing predictor(s) and including potential new predictor(s). The third and fourth methods were followed by a bootstrapping procedure. We found that when the model was applied with no update, the model's discrimination was adequate (ROC area was 0.74 vs. the original ROC area of 0.76), but the calibration was poor. The updated models showed good calibration and reasonable discrimination (ROC area ranged between 0.73 and 0.75). These findings suggested that once updated, the diagnostic model is valid and can be applied with reasonable performance in an animal health apprentice setting.

In **chapter 5** we demonstrated the application of a diagnostic model to predict sensitization to wheat and/or  $\alpha$ -amylase allergens in a nationwide surveillance program to detect occupational respiratory allergy among workers from the baking and flour producing industries. The diagnostic model with four simple questionnaire items was transformed into a scoring rule, and was used to stratify 5,325 workers into three risk categories: low, intermediate and high. There were 18.5% workers with a high risk who were advised to visit a specialized occupational respiratory health clinic, whereas 24.1% workers with an intermediate risk were advised to visit their occupational physicians for further diagnostic work up. For workers in the low risk group (57.4%), no or less far-reaching medical

investigations are needed. Clinical investigations showed that workers in the high probability group had the highest percentages of doctor visit, medication use, absenteeism, and change in job title due to allergic symptoms. The application of the diagnostic rule, simultaneously, showed that detection of subgroups of workers, with an intermediate or high probability improved the expected benefit of diagnosis and treatment.

In the **discussion**, some main issues in model development and application are explored. Predictors to be included in a prediction model can be obtained from questionnaire, physical examination, laboratory test, imaging, or other additional tests. To stay close to the diagnostic sequence in practice, it is attractive to start a model with questionnaire items. The inclusion of other tests into the model will be strongly dependent on the availability of resources and the context where the model will be applied. After assuring that a model is valid and produces accurate predictions, it is important to determine probability thresholds to stratify individuals into risk categories. In general a higher threshold leads to a lower percentage of high risk group; the specificity is higher (lower false positive rate), but at the cost of lower sensitivity (higher false negative rate). The choice for a threshold must be based on the balance between the proportion of missed cases and reduction of unnecessary diagnostic tests. Some other considerations in model application were also described, including potential misuse of the model that has to be carefully avoided, as well as the transformation of the prediction model into a user friendly tool.

This thesis shows that predicting lung diseases in the context of occupational health care and practice is possible. The use of the prediction tools assists the decision making process and would hopefully reduce expenses. Application of the prediction models has not been fully explored but efforts to increase the use of predictive models deserve strong support.



# SAMENVATTING

---

## VOORSPELLEN VAN BEROEPSGEBONDEN LONGZIEKTEN

---

Voorspellend onderzoek wordt al langer wordt toegepast bij het ondersteunen van beslissingen in de klinische geneeskunde maar is nieuw op het terrein van beroepsziekten. In predictieonderzoek worden modellen ontwikkeld om de individuele kans op de aanwezigheid van een ziekte of uitkomst (diagnostiek) of het optreden ervan in de toekomst (prognostiek) te voorspellen. Deze modellen worden gebruikt voor de ondersteuning van klinische beslissingen en om individuen in te delen in risicogroepen voor verschillende ziekten. Predictiemodellen maken het mogelijk om op een objectieve en gestandaardiseerde wijze de individuele kans op het hebben of ontwikkelen van een ziekte te bepalen zonder uitgebreide (invasieve) en dure testen te moeten uitvoeren. Dit proefschrift beschrijft de ontwikkeling, validatie en toepassingen van predictiemodellen voor beroepsgebonden longziekten.

Alle modellen, zoals in dit proefschrift beschreven, zijn ontwikkeld met behulp van multivariate logistische regressie met terugwaartse selectie. De overeenkomst tussen de waargenomen en de voorspelde prevalenties van de ziekte (calibratie) wordt geëvalueerd. De discriminatie van het model wordt bepaald door de oppervlakte onder de Receiver Operating Characteristic (ROC) curve te berekenen. De interne validatie van het model wordt bepaald door middel van de “bootstrap” methode om te corrigeren voor “overfitting” van het model. De regressie coëfficiënten van de predictoren (voorspellende variabelen) van het model zijn vervolgens gecorrigeerd om te voorkomen dat het model te optimistische voorspellingen zou doen bij nieuwe werknemers. Tot slot, om de klinische toepassing te vereenvoudigen, zijn alle modellen omgezet in eenvoudig bruikbare score regels.

In **hoofdstuk 2.1** wordt beschreven hoe een diagnostisch vragenlijstmodel is ontwikkeld en gevalideerd om sensibilisatie voor tarweallergenen te voorspellen. Het model omvat zes predictoren: type bakkerij, oog en neusklachten gedurende de laatste 12 maanden, astma klachten in de laatste 12 maanden, kortademigheid en piepen, klachten van de bovenste en onderste luchtwegen gedurende het werk. Het model toonde goede discriminatie (ROC oppervlak 0,76) en interne validatie (de correctiefactor was 0,89 en het gecorrigeerde ROC oppervlak was 0,75). Externe validatie toonde een redelijke discriminatie (ROC oppervlak 0,69), maar een slechte calibratie. Correctie van de intercept verbeterde de calibratie van het model. In de vorm van een eenvoudige score-kaart geeft het de mogelijkheid om handmatig de individuele kans op sensibilisatie voor tarwe allergenen te berekenen. Ons onderzoek ondersteunt de opvatting dat een eenvoudige vragenlijst betrouwbaar de aanwezigheid van tarwe sensibilisatie kan voorspellen en kan worden ingezet bij alle bakkers blootgesteld aan tarwemeel.

In **hoofdstuk 2.2** wordt de ontwikkeling van een algemeen diagnostisch model voor de sensibilisatie voor allergenen met een hoog moleculegewicht beschreven door gebruik te maken van eenvoudige vragenlijstonderdelen en routine laboratorium testen. Het model werd ontwikkeld met samengevoegde data van Nederlandse proefdierwerkers en bakkers. Het model bevatte de volgende predictoren: aantal werkuren per week, werkgerelateerde respiratoire klachten, totaal IgE en IgE voor algemene allergenen. We vonden significante interactie tussen het type werk en de andere predictoren. Dit resulteerde in verschillende scores voor proefdierwerkers en bakkers. De discriminatie van het model was voor

proefdierwerkers 0,80 en voor bakkers 0,70. De interne validatie van het model was goed (de voor optimisme gecorrigeerde ROC oppervlak was 0,76 bij een correctiefactor van 0,95). Externe validatie in Britse proefdierwerknemers toonde dat het model voldoende discrimineerde en kalibreerde voor hoog en laag sensibilisatierisico (ROC oppervlak 0,76). In deze study is aangetoond dat het mogelijk is een algemeen model voor het voorspellen van sensibilisatie tegen beroepsgebonden allergenen met een hoog moleculegewicht te ontwikkelen. Echter de weging van de predictoren verschilt tussen de verschillende arbeidsomgevingen.

**Hoofdstuk 2.3** beschrijft een diagnostisch model om de kans op pneumoconiose te voorspellen. Pneumoconiose is hierbij gedefinieerd als afwijkingen op de X-thorax overeenkomstig ILO categorie  $\geq 1/1$ . Het model is ontwikkeld uit een cross-sectionele studie onder 1291 Nederlandse bouwvakkers en natuursteenwerkers met mogelijk hoge kwartsblootstelling. Leeftijd ouder dan 40 jaar, actief roken, hoge blootstelling, meer dan 15 jaar werkzaam in de bouwnijverheid, zich "ongezond voelen" en de FEV1 (gestandaardiseerd residu  $\leq -1.0$ ) zijn de onafhankelijke predictoren in het model. Het voorspellend vermogen (calibratie en discriminatie) van het model was goed (ROC oppervlak 0,81). Na correctie voor optimisme was het ROC oppervlak 0,76. Doel van het model was om betrouwbaar werknemers met een geringe kans op pneumoconiose te detecteren. Door een afkappunt met een hoog negatief voorspellende waarde te kiezen kan de bedrijfsarts deze grote groep werknemers een onnodig röntgenonderzoek van de thorax besparen. Daardoor kan dit model de efficiency van het opsporen van pneumoconiose verbeteren, zelfs in een populatie met een lage prevalentie.

In **hoofdstuk 3** zijn vragenlijstgegevens, huidpriktesten (HPT) voor algemene allergenen en bronchiale hyperreactiviteit (BHR) voor methacholine gebruikt om bij leerling proefdierwerkers, na 32 maanden, een predictie model te ontwikkelen voor het voorspellen van werkgerelateerde sensibilisatie en luchtwegklachten. Er zijn 4 modellen ontwikkeld voor elk eindpunt: vragenlijst (1), vragenlijst en HPT (2), vragenlijst en BR-test (3) en vragenlijst, SPT en BR-test (4). Het uiteindelijke model van werkgerelateerde sensibilisatie en luchtwegklachten bevatte: klachten die wezen op astma en allergieklachten reeds aanwezig bij aanstelling. De vragenlijstmodellen voor beide eindpunten vertoonden goede discriminatie (ROC oppervlak respectievelijk 0,73 en 0,78) en goede calibratie. Het toevoegen van HPT en/of BHR test resultaten verhoogde de specificiteit van het vragenlijstmodel om sensibilisatie te voorspellen, maar niet voor arbeidsgerelateerde klachten. Om deze reden kon worden vastgesteld dat het vragenlijstmodel een goed instrument is om de kans op werkgerelateerde sensibilisatie en klachten te voorspellen. Toegevoegde testen verbeterden de specificiteit van het predictiemodel voor sensibilisatie tegen proefdier allergenen.

In **hoofdstuk 4** modificeren en valideren we een predictiemodel dat de sensibilisatie voor proefdier allergenen voorspelt. Het model is ontwikkeld in Nederlandse proefdierwerkers en is extern gevalideerd in Canadese leerling proefdierwerkers. Er zijn verschillende benaderingen gekozen om het model extern te valideren: (1) geen up-date; (2) up-date van de *intercept* van het model; (3) up-date van de *intercept* en alle regressiecoëfficiënten en (4) revisie van het model door alle bestaande voorspellers uit te sluiten en/of nieuwe voorspellers in te voegen. De derde en vierde methode werden uitgevoerd door middel van een zogenaamde *bootstrapping* procedure. We vonden dat, wanneer het model werd toegepast zonder up-date, de discriminatie van het model voldoende was (ROC oppervlak 0,74 terwijl het originele model een ROC oppervlak had van 0,76), maar dat de calibratie slecht was. De ge-up-date modellen vertoonden een goede calibratie en een redelijke discriminatie (ROC oppervlak tussen 0,73 en 0,75). Deze uitkomsten suggereren dat een geactualiseerd diagnostisch model geschikt is en kan worden toegepast met redelijke resultaten bij leerling proefdierwerker.

In **hoofdstuk 5** hebben we aangetoond dat de toepassing van een diagnostisch model voor de sensibilisatie voor tarwe- en/of  $\alpha$ -amylase-allergenen in een nationaal gezondheidsonderzoek naar allergische luchtwegaandoeningen onder werknemers in de bakkers- en meelverwerkende industrieën. Het diagnostische model met vier eenvoudige vragen werd omgezet in score regel en werd gebruikt om 5325 werknemers in drie categorieën in te delen: laag, gemiddeld en hoog risico. 18,5% van de werknemers had een hoog risico, zij zijn verwezen naar een gespecialiseerde polikliniek voor arbeidsgebonden longaandoeningen. De 24,1 % van de werknemers, bij wie een gemiddeld risico werd vastgesteld, zijn verwezen naar hun bedrijfsarts voor nader onderzoek van hun klachten. De werknemers met een laag risico (57,4%) zijn niet verder verwezen, omdat geen aanvullend onderzoek noodzakelijk is. Klinisch onderzoek toonde aan dat de groep werknemers met het hoogste risicoprofiel tevens het hoogste percentage dokters bezoek, medicijn gebruik, verzuim, en werk verandering vanwege allergische klachten toonde. Het toepassen van de diagnostische regel liet eveneens zien dat door het opsporen van werknemers met een gemiddeld of hoog risicoprofiel de te verwachte voordelen van diagnose en behandeling kon worden verbeterd.

In de **discussie** wordt een aantal algemene zaken met betrekking tot modelontwikkeling en toepassing besproken. Voorspellers voor een predictiemodel kunnen worden verkregen via een vragenlijst, lichamelijk onderzoek, laboratorium onderzoek, röntgenopnames, en andere aanvullende testen. Om zo dicht mogelijk bij het gebruikelijke diagnostische proces te blijven is het aantrekkelijk om te starten met een vragenlijst. Het toevoegen van andere testen aan het model zal sterk afhangen van de beschikbaarheid van middelen en context waarin het model wordt toegepast. Nadat is vastgesteld dat een model geschikt is en betrouwbare voorspellingen doet, is het van belang de afkappunten te bepalen om individuen te verdelen in de verschillende risicogroepen. In het algemeen leidt een hoger afkappunt tot een lager percentage werknemers in de hoogste risicogroep, een grotere specificiteit (minder vals positieven) met als nadeel een lagere sensitiviteit (meer vals negatieven). De keuze voor een bepaald afkappunt moet zijn gebaseerd op een evenwicht tussen de hoeveelheid gemiste waarnemingen en het verminderen van onnodige diagnostische testen. Een aantal andere overwegingen voor modeltoepassingen zijn eveneens in dit hoofdstuk beschreven, zoals mogelijk misbruik dat moet worden vermeden, als het omzetten van het regressiemodel in een gebruiksvriendelijk instrument.

Dit proefschrift toont aan dat het voorspellen van beroepsgebonden longziekten door middel van predictiemodellen mogelijk is. Het gebruik van voorspellers bevordert het besluitvormingsproces en vermindert hopelijk de kosten. De toepasbaarheid van voorspellende modellen is niet volledig onderzocht maar een toename van het gebruik van predictiemodellen moet sterk worden bevorderd.



# RANGKUMAN

---

## PREDIKSI KELAINAN PARU AKIBAT KERJA

---

Penelitian diagnostik dan prognostik merupakan hal yang relatif baru di bidang kesehatan kerja walaupun penelitian ini sering dilakukan untuk membantu pembuatan keputusan di ilmu kedokteran klinis. Dalam disertasi ini, model statistik dibuat untuk mengestimasi risiko terjadinya suatu *outcome* (bisa berupa penyakit atau keadaan yang berhubungan) pada individu. Model ini digunakan untuk membantu penetapan keputusan klinis terhadap individu atau untuk stratifikasi risiko atau kecenderungan terjadinya suatu *outcome*. Model ini memungkinkan kuantifikasi probabilitas terjadinya penyakit saat ini atau di masa mendatang secara obyektif tanpa harus menjalani test dan prosedur yang membutuhkan biaya yang besar. Disertasi ini difokuskan pada pembuatan, validasi, dan penerapan model diagnostik dan prognostik untuk memprediksi kelainan paru akibat kerja.

Semua model dalam tesis ini dibuat menggunakan analisis *multivariable logistic regression* dengan *backward stepwise selection*. Dilakukan penilaian kesesuaian antara probabilitas yang diprediksi dan frekuensi *outcome* yang diobservasi (kalibrasi). Kemampuan diskriminasi model untuk membedakan individu yang memiliki dan tidak memiliki *outcome* dievaluasi dengan *receiver operating characteristic (ROC) area*. Area ROC dapat berkisar antara 0,5 (diskriminasi buruk) sampai 1,0 (diskriminasi sempurna). Pengujian validitas internal model dilakukan dengan menggunakan prosedur *bootstrapping*. Prosedur ini menghasilkan area ROC yang terkoreksi dan faktor koreksi (*shrinkage factor*). Koefisien regresi model prediksi ini kemudian dikalikan dengan faktor koreksi untuk menghindari *overfitting* (estimasi terlalu tinggi atau terlalu rendah) ketika model diterapkan pada subyek baru. Pada akhirnya, untuk memudahkan penerapan model dalam praktik, semua model diubah menjadi sistem skor yang mudah digunakan.

Di **bab 2.1** dikembangkan sebuah model diagnostik dalam bentuk kuesioner untuk mengetahui tingkat sensitivitas terhadap gandum pada pekerja pabrik roti. Model ini menggunakan 6 prediktor yang terdiri dari: tipe perusahaan (tradisional dan non-tradisional), gejala nasojunctival selama 12 bulan terakhir, gejala asma selama 12 bulan terakhir, mengi, gangguan saluran pernapasan bagian atas dan bawah akibat kerja. Model ini menunjukkan kemampuan diskriminasi yang baik (area ROC 0,76) dan validitas internal yang memuaskan (faktor koreksi 0,89 dan area ROC yang dikoreksi 0,75). Validasi eksternal menunjukkan diskriminasi yang cukup baik (area ROC 0,69), tapi kalibrasi tidak memuaskan. Dilakukan beberapa metode *update* untuk mengetahui apakah model bisa ditingkatkan performanya. Re-estimasi *intercept* ternyata mampu memperbaiki kalibrasi model. Dalam bentuk sistem skor, model ini bisa digunakan untuk memprediksi risiko sensitivitas terhadap alergen gandum secara manual. Penemuan ini mendukung ide bahwa kuesioner yang sederhana bisa digunakan untuk memprediksi ada tidaknya sensitivitas terhadap gandum. Setelah dilakukan validasi eksternal pada kelompok pekerja roti yang berbeda, model ini bisa diterapkan secara lebih luas lagi pada populasi pekerja pabrik roti lainnya.

Di **bab 2.2**, dikembangkan model diagnostik generik untuk memprediksi sensitivitas terhadap alergen dengan berat molekul besar dengan menggunakan kuesioner yang sederhana dan uji laboratorium rutin. Model ini dikembangkan dengan menggunakan kumpulan data pekerja laboratorium hewan dan pekerja pabrik roti di Belanda. Model final terdiri dari lama waktu bekerja setiap minggunya, gangguan pernapasan di tempat kerja, total IgE dan sensitivitas terhadap alergen umum (debu rumah, serbuk bunga, bulu kucing,

bulu anjing). Ditemukan interaksi yang sangat kuat antara jenis pekerjaan dan prediktor dalam model, dimana interaksi ini menghasilkan skor prediktor yang berbeda antara pekerja laboratorium hewan dan pekerja pabrik roti. Kemampuan diskriminasi (area ROC) model diagnostik pada pekerja laboratorium hewan sebesar 0,80, sedangkan pada pekerja pabrik roti sebesar 0,70. Secara keseluruhan, validitas internal model ini baik (area ROC yang dikoreksi 0,76 dan faktor koreksi 0,95). Validitas eksternal pada pekerja laboratorium hewan di Inggris menunjukkan bahwa kalibrasi model ini cukup memuaskan (Hosmer-Lemeshow test  $p=0,786$ ) dan dapat mendiskriminasi pekerja dengan risiko tinggi dan rendah (area ROC 0,76). Dapat disimpulkan bahwa pengembangan model generik untuk memprediksi sensitisasi alergen dengan berat molekul besar mungkin dilakukan. Walaupun demikian, bobot prediktor dapat berbeda untuk lingkungan kerja yang berbeda.

Di **bab 2.3** dijelaskan tentang pengembangan model diagnostik untuk mengestimasi probabilitas pneumoconiosis, sebagaimana diindikasikan oleh hasil Roentgen dada dengan profusi kategori  $> 1/1$ . Model ini dikembangkan berdasarkan data pekerja konstruksi yang berpotensi terpajan debu silika. Prediktor independen yang digunakan dalam model diagnostik ini adalah: usia 40 tahun ke atas, perokok, bekerja di tempat yang tingkat pajanannya tinggi, bekerja di industri konstruksi selama 15 tahun atau lebih, “merasa kurang sehat”, dan  $FEV_1$ . Model ini menunjukkan kalibrasi dan diskriminasi yang baik (uji HL tidak bermakna). Hasil validasi model secara internal cukup baik; area ROC yang dikoreksi 0,76. Model ini ditujukan untuk mendeteksi pekerja yang probabilitas pneumoconiosis-nya rendah sehingga kelompok ini tidak perlu diikutkan dalam pemeriksaan Roentgen dada. Dokter perusahaan dapat menggunakan nilai ambang yang memiliki *negative predictive value* yang tinggi untuk memastikan bahwa kelompok yang bersiko rendah memang tidak akan menunjukkan hasil Roentgen yang positif. Penggunaan model ini sangat membantu meningkatkan efisiensi pendeteksian pneumoconiosis, terutama untuk populasi pekerja yang tingkat prevalansinya rendah.

Di **bab 3** dijelaskan proses pengembangan model prognostik untuk prediksi sensitisasi dan gangguan pernapasan yang terjadi pada pekerja magang setelah menjalani magang selama 32 bulan di bidang *animal health technology*. Model dibuat menggunakan informasi kuesioner, *skin-prick testing* (SPT) terhadap alergen yang umum, dan uji hiperreaktivitas bronkus (BR) dengan *methacholine*. Empat model logistik regresi dibuat untuk tiap *outcome*, yaitu: (1) kuesioner; (2) kuesioner dan SPT; (3) kuesioner dan uji BR. Gejala-gejala asma dan alergi pada awal magang terpilih sebagai prediktor dalam model kuesioner final untuk memprediksi sensitisasi dan gangguan pernapasan di tempat kerja. Kedua model kuesioner tersebut menunjukkan diskriminasi yang baik (area ROC berturut-turut 0,73 dan 0,78) dan kalibrasi yang baik pula ( $p > 0,10$ ). Penambahan informasi SPT dan/atau BR meningkatkan spesifisitas model kuesioner untuk sensitisasi, tetapi tidak untuk gangguan pernapasan akibat kerja. Disimpulkan bahwa model kuesioner dapat dipakai sebagai alat untuk memprediksi terjadinya sensitisasi atau gangguan pernapasan di tempat kerja; test tambahan (SPT dan/atau BR) meningkatkan spesifisitas prediksi sensitisasi terhadap alergen hewan laboratorium.

Di **bab 4** dilakukan uji validasi model diagnostik untuk mendeteksi sensitisasi terhadap alergen hewan laboratorium dan dilakukan evaluasi apakah model ini bisa ditingkatkan performanya. Model ini dibuat berdasarkan data pekerja laboratorium hewan di Belanda dan diuji validitas eksternalnya pada pekerja magang di bidang *animal health technology* di Kanada. Uji validitas eksternal dilakukan dengan menggunakan beberapa pendekatan, yaitu: (1) tanpa update; (2) rekalisasi *intercept*; (3) re-kalibrasi *intercept* dan koefisien regresi; dan (4) revisi model dengan mengeksklusi prediktor yang ada atau memasukkan prediktor baru yang potensial. Metode ketiga dan keempat diikuti oleh prosedur *bootstrapping*. Kemudian kalibrasi dan diskriminasi model ini diuji kembali untuk

mengetahui ketepatan akurasi. Ketika model ini diterapkan tanpa *update*, diskriminasinya cukup baik (area ROC 0,74 vs. area ROC asal 0,76), tetapi hasil kalibrasi tidak memuaskan (uji H-L test  $p < 0,001$ ). Model yang sudah di-*update* menunjukkan kalibrasi yang baik ( $p > 0,10$ ) dan hasil diskriminasi yang baik (area ROC berkisar antara 0,73 dan 0,75). Temuan ini membuktikan bahwa setelah di-*update*, model diagnostik yang dibuat pada populasi pekerja bisa diterapkan pada pekerja magang di bidang *animal health technology* dengan hasil yang baik.

Di **bab 5** diilustrasikan penerapan model diagnostik untuk prediksi sensitisasi alergen gandum dan enzim pada para pekerja pabrik roti dan industri tepung. Model diagnostik ini dibuat dengan menggunakan empat kuesioner sederhana yang kemudian diubah menjadi sistem skor, yang kemudian digunakan untuk menstratifikasi 5.325 pekerja ke dalam tiga kelompok risiko: 18,5% pekerja dengan tingkat risiko tinggi dirujuk ke klinik khusus untuk gangguan paru akibat kerja, sedangkan 24,1% pekerja dengan risiko sedang dirujuk ke dokter perusahaan untuk diperiksa lebih lanjut. Untuk grup pekerja dengan risiko rendah (57,4%), tidak ada rujukan yang perlu dilakukan, tetapi mereka akan diikutsertakan dalam program surveilans berikutnya. Pemeriksaan secara klinis menunjukkan bahwa grup pekerja dengan probabilitas tinggi mempunyai persentase tinggi dalam hal kunjungan ke dokter, penggunaan obat, absen kerja, dan pindah pekerjaan dikarenakan gejala alergi. Penerapan model diagnostik ini secara konsisten menunjukkan bahwa kelompok pekerja yang mempunyai probabilitas tinggi atau sedang dapat terdeteksi secara akurat sehingga pemeriksaan dan pengobatan bisa dilakukan secara dini.

Pada **bab diskusi**, dibahas beberapa persoalan dalam pengembangan dan penerapan model diagnostik dan prognostik. Prediktor dapat berasal dari kuesioner, pemeriksaan fisik, uji laboratorium, pencitraan, atau uji tambahan lainnya. Mengikuti praktik di klinik, model yang terdiri dari kuesioner saja lebih diutamakan. Penggunaan uji tambahan bergantung kepada tersedianya sumber daya di mana suatu model akan diterapkan. Setelah model ini dipastikan menghasilkan prediksi yang akurat, penentuan ambang probabilitas diperlukan untuk membuat kategori risiko. Secara umum, semakin tinggi nilai ambang yang dipilih, semakin kecil persentase grup berisiko tinggi; spesifisitas lebih tinggi (*false positive rate* lebih rendah), tetapi sensitivitas lebih rendah (*false negative rate* lebih tinggi). Penentuan nilai ambang harus didasari atas keseimbangan antara persentase misklasifikasi dan reduksi uji diagnostik yang tidak diperlukan. Pertimbangan lain yang harus diperhatikan dalam penerapan model ini adalah potensi implikasi negatif yang harus dihindari, begitu juga transformasi model regresi ke dalam bentuk yang mudah digunakan.

Tesis ini mengilustrasikan bahwa prediksi kelainan paru akibat kerja mungkin dilakukan. Evaluasi penerapan model prediksi ini belum sepenuhnya dieksplorasi, namun demikian usaha untuk meningkatkan penggunaan model prediksi di bidang kesehatan kerja patut terus diupayakan.



# AFFILIATION OF CONTRIBUTORS

**Denyse Gautrin, Heberto Ghezso, Jean-Luc Malo, Jocelyne L'Archevêque RT**  
Center for Asthma in the Workplace, Hôpital du Sacré-Coeur de Montréal, Canada

**Dick Heederik, Eef van Otterloo, Evert Meijer, Frits van Rooy, Jack Spithoven, Jose Jacobs, Tim Meijster, Vanessa Zaat**  
IRAS (Institute for Risk Assessment Sciences), Environmental Epidemiology Division, Utrecht University, The Netherlands.

**Dick Heederik, Diederick E. Grobbee, Karel G.M. Moons, Yvonne Vergouwe**  
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands.

**Frits van Rooy**  
Arbo Unie Expert Centre for Chemical Risk Management, Utrecht, the Netherlands.

**Frits van Rooy, Jos Rooijackers, Vanessa Zaat**  
Netherlands Expertise Centre for Occupational Respiratory Disorders, Utrecht, the Netherlands.

**Jan de Monchy**  
Department of Allergology, University Hospital Groningen, Groningen, the Netherlands

**Mark Nieuwenhuijsen**  
Department of Occupational and Environmental Medicine, Imperial College (NHLI), London, United Kingdom



# LIST OF PUBLICATIONS

**E. Suartha**, F.A. Tarigan, F. Kaligis, A. Sandra, D. Purwanto, S. Hadi. [Hypertension prevalence and the related nutrition factors among housewives in Jakarta]. *Maj Kedok Indon* 2001; 51 (5).

**E. Suartha**, Y. Vergouwe, M. Nieuwenhuijsen, D. Heederik, DE Grobbee, E. Meijer. Diagnostic rule for sensitization in workers exposed to occupational high molecular weight allergens. *Am J Ind Med.* 2005 Sep;48(3):168-74.

A Widodo, AM Anggraeni, A Halim, A Maureen, B Landy, **E Suartha**, S Adjie. [Knowledge, behavior, and attitude towards pregnancy and its complication among non primigravida pregnant women]. *Maj Kedok Indon* 2005; 55(10) 630-638.

**E. Suartha**, KGM Moons, D. Heederik, E. Meijer. A simple diagnostic model for ruling out pneumoconiosis among construction workers. *Occup Env Med* 2007;64(9):595-601

Sunanto, **E Suartha**. [Epidemiology for clinicians]. *J Kardiologi Indon* 2007; 28(2):85-89.

MAA Yussac , A Cahyadi, AC Putri, AS Dewi, A Khomaini, S Bardosono, **E Suartha**. [Obesity prevalence in children 4-6-year age and its relation with nutrition intake]. *Maj Kedok Indon* 2007; 57(2): 47-53.

D Gautrin, H Ghezzi, C Infante-Rivard, M Magnan, J L'Archevêque R.T, **E Suartha**, JL Malo. Long-term outcome of a prospective cohort of apprentices exposed to high-molecular-weight agents. *Am. J. Respir. Crit. Care Med.* 10.1164/rccm.200707-991OC

JH Jacobs, T Meijster, E Meijer, **E Suartha**, D Heederik. Wheat allergen exposure and the prevalence of work-related sensitization and allergy. *Accepted for publication in Allergy, the European Journal of Allergy and Clinical Immunology*

**E Suartha**, E Meijer, D Heederik, H Ghezzi, JL Malo, D Gautrin. The generalizability of the Dutch diagnostic model for laboratory animal allergens sensitization in Canadian apprentices. *Provisionally accepted for publication in the Journal of Clinical Epidemiology*

**E Suartha**, D Heederik, S Kennedy, H Ghezzi, JL Malo, D Gautrin. Risks for the development of various outcomes related to occupational allergies: An application of the asthma-specific job exposure matrix. *In revision for the Journal of Occupational and Environmental Medicine.*

**E Suartha**, Y.Vergouwe, KGM Moons, D.E. Grobbee, J de Monchy, D Heederik, E Meijer. Development and validation of a diagnostic model for detection of sensitization to wheat allergens: *Prediction study in bakery workers. Submitted*

E Meijer, **E Suartha**, J de Monchy, F van Rooy, J Rooijackers, T Meijster, J Jacobs, E van Otterloo, J Spithoven, V Zaat, DE Grobbee, D Heederik. Diagnostic research in occupational respiratory allergy: *Results from a nationwide surveillance among bakery workers. Submitted*

**E Suarthana**, JL Malo, D Heederik, H Ghezzi, J L'Archevêque RT, D Gautrin. Which tools best predict incidence of work-related sensitization and symptoms? *Predictors of occupational allergy*. *Submitted*

**E. Suarthana**, E. Meijer, D.E. Grobbee, D. Heederik. Predicting occupational diseases. *Submitted*

# ACKNOWLEDGEMENTS

First of all, I would like to thank God more for making this happen.

*Alhamdulillahirobil al'amin... all praise is due to Allah subhanahu wa ta'ala...*

Just a short flash back, my story started in April 2002, when a group of professors from the Julius Center for Health Sciences and Primary Care, of UMC Utrecht, The Netherlands came to Jakarta for a one week course on clinical epidemiology. This course was an eye opener for me and showed that epidemiology is an interesting branch in medicine and is not all about statistics! I interviewed one of the professors about life as a clinical researcher and I thank him, who turned out to be my promotor, for inspiration.

In October 2002, I specifically asked the program co-ordinator at the NIHES if they had a research project for my Master's thesis in occupational medicine. That led me to a collaborative project between IRAS-Environmental Epidemiology division and Julius Center for Health Sciences and Primary Care, UMC Utrecht. From that time on, I have been working with Prof. Dr. Dick Heederik and Prof. Dr. Diederick Grobbee, who are now my promotors and and Dr. Evert Meijer my co-promotor. I am grateful that I had the chance to learn a lot from these excellent persons, professionally and personally. Thank you for everything: for the opportunity, for the encouragement, for being a teacher and supporter all at the same time.

To everyone helping me with the bakery project: Birgit Stuurman, Eef van Otterloo, Frits van Rooy, Isabella Oosting, Jack Spithoven, Jos Rooijackers, Jose Jacobs, Joris van Loenhout, Lieke Gerris, Mischa Zengeni, Tim Meijster, and Vanessa Zaat, thank you for the team work! Special thanks go to Tim and Jack, my paranimfen. Tim, thank you for helping me from the very beginning: finding me a room, showing me the beauty of the Netherlands, for all the dinners with Marjoke (and Wiebe), and yet being a very informative and constructive co-worker in the project! Jack, I have always been amazed by the speed with which you have been running those assays! You are the only person I know who could handle the hardest situation with laughter. Keep on being strong and cheerful!

At this occasion I would also like to thank all who contributed to my thesis, especially Dr. Yvonne Vergouwe, who has been my great tutor in statistical modelling. My analytical skills have also been improving thanks to the "Prediction Club" at Julius Center, lead by Prof. Dr. Carl Moons. The group keeps on growing in numbers and quality. I thank Corne, Jolanda, Kristel, Diane, Peter, Lidewij, Martijn, and Teus (the "old" members); Geert, Joris, Roelof, Mireille, Sjoerd, and Linda (the new members); and most of all, Carl, for sharing your knowledge every other Monday. Hope all of us will become "accurate and valid predictors"!

I would also like to thank Dr. Denyse Gautrin, Dr. Jean-Luc Malo and Dr. Heberto Ghezso for giving me a chance to do a 1-year fellowship at the Asthma in the Workplace Centre, Montreal, Canada. It is funny because before the fellowship I only knew you from text books and papers and I never thought that I would have this rare opportunity to work in close contact with you. *Merci beaucoup*, it's been wonderful to work with you! My appreciation also goes to Sylvie, Roberto, Lucero, Samah, Khadija, Christina, Sandra, Sebastien, and the members of the *axe de recherche en santé respiratoire, Hôpital du Sacré-Coeur de Montréal*, who made my stay a fruitful one.

To Dr. Mark Nieuwenhuijsen from Imperial College-NHLI (London), Dr. David Fishwick from HSL (Sheffield), Dr. Susan Kennedy from UBC (Vancouver), Dr. Kathleen Kreiss from NIOSH (Morgantown), Sally Brown from the EIS program (Atlanta), Dr. Manon Kluijtmans from Julius Center (Utrecht), Prof. Dr. Bambang Sutrisna from FKMUI (Jakarta), Prof. dr. Sudigdo Sastroasmoro from FKUI (Jakarta), and Prof. dr. Saleha Sungkar from YPIDI (Jakarta), it has been a great pleasure to knowing you in person. I hope our collaboration will continue in the future.

To my seniors and colleagues at the Community Medicine Department, especially Prof. dr. Bastaman Basuki, Prof. Dr. Corrie Wawolumaya, dr. Zarni Amri, Dr. Muchtaruddin Mansyur dr. Novi Arifiani and dr. Retno Asti Werdhani, thank you for your support and back up while I was away (for such a long period!). To dr. Nitra Rifky, thank you for giving me the very chance to enter the whole system. To my head of department, dr. Setyawati Budiningsih, thank you for your trust and being very supportive and open minded...

Last but not least, I'd like to thank all my friends and the greatest people who helped me through the good and bad times, and make these few years as one blink of an eye. To my office mates (Mieke, Mariella, Wouter, Frank, Lot, Maaïke, and Esmeralda) and my housemates (Ebru, Celal, Bilgen, Daniel, Jessica, Maxi, Scato, Sandra, Marja and Corine), thank you for cheering up my days. Especially Shuwen, Ilka, and Sarah, thank you for being more than good friends! To Indonesian muslim communities (PPMR, PPME Rotterdam, PPME Den Haag, Bina Da'wah Utrecht, and Syiar Montreal), especially to the family of Bpk. Endro Harsono, Bpk. Rifai Hasan, Bpk. Ridwan, Ibu Taty, Ibu Fatima, Mbak Wilkaziana, Mbak Laila, Bpk. Razak, Bpk. Pourchez, Bpk. Donni, Bpk. Bachtiar, Bpk. John Kasijo, Ibu Dewi, Mbak Nita and Mas Yusuf, Mbak Masinah and Mas Franklin, Risda and Supardi, Eha and Andi, thank you for making me feel like being at home and being my family abroad. To my friends during my master course: Hamid (and Gisso and Kiana), Made, Wangge, Isabelle, Haris, Novita, Henny, Vica, Jamaiyah, Kiad, Esther, Elisa, Gloria, Regie, Monica, Natalia, and Daniela, thank you for the genuine friendship and always being great supporters. To Ingrid, Evelyn, Giene, and Annina, thank you for taking care of my paper work and arranging countless appointments! To Hapidudin, everyone at IRAS EEPI, Jelena, Anjoeka, Rocio, Angeles, Fatima and Youssef, Mbak Hesti and Anya, Mbak Anna, Mbak Imelda and Nico, Ratih, Liesye, Teddy, Ari, Jessica, Fleur, Marije, Aicha, Nike, Ika, Vicka, Mbak Hermina and Rene, Atiet, Ratna, Fenida, Kadek, Suryaningsih, Soraya, Eka and everyone in my long list who I do not name one by one, thank you for your valuable contribution to my achievements!

I finally want to express my utmost gratitude to my family for giving me freedom to do what I believe in. With out your prayers and sincere love, I will never become of what I am. I love you!!! *Untuk Bapak, Mama, Santhy, Sherley, Media dan Clara terima kasih atas dukungan, doa, dan kasih sayang kalian. Tanpa kalian, hari ini tidak akan menjadi kenyataan...*

**The Road not Taken**  
*Two roads diverged in a yellow wood,  
And sorry I could not travel both....  
Two roads diverged in a wood, and I  
I took the one less travelled by,  
And that has made all the difference.*  
**Robert Frost, 1920**

# CURRICULUM VITAE

Eva Suarhana was born on April 24<sup>th</sup>, 1977, in Jakarta, Indonesia. She graduated from senior high school in 1995 in Denpasar. In 2001 she obtained her medical degree at the Faculty of Medicine, University of Indonesia, in Jakarta. After graduating she worked as an assistant lecturer at the Family Medicine Division, Community Medicine Department, Faculty of Medicine, University of Indonesia. In 2002 she was awarded a scholarship from the Netherlands Institute for Health Sciences, Rotterdam, to do a Master of Science in Clinical Epidemiology. During this period, she conducted a collaborative research project in occupational health between Institute for Risk Assessment Sciences (IRAS) Environmental Epidemiology Division and Julius Center for Health Sciences and Primary Care, UMC Utrecht. After completion of her master in 2003, she returned to her department in Jakarta and was assigned in the Epidemiology and Biostatistics Division for 1.5 years. In December 2004 she started working as a PhD-student at IRAS and Julius Center (supervised by Prof. Dr. Dick Heederik, Prof. Dr. Diederick Grobbee, and Dr. Evert Meijer). Her main project was diagnostic research in occupational respiratory allergy among Dutch bakery workers. Her master thesis entitled "Diagnostic model for sensitization in workers exposed to occupational high molecular weight allergens" eventually became part of her PhD thesis. In 2006 she was awarded a 1-year fellowship at the Asthma in the Workplace Center, Hôpital du Sacré-Coeur de Montréal, Montréal, Canada. She worked under supervision of Dr. Denyse Gauthier on the Apprentices Cohort Study as described in this thesis. In February 2007 she was appointed as lecturer at the Community Medicine Department, Faculty of Medicine, University of Indonesia. Since January 2004 she has been working as an editor for the journal of the Indonesian Medical Doctor Association.

