

# Adjusting wheal size measures to correct atopy misclassification

Hongmei Zhang<sup>1,\*</sup>  
 Wilfried Karmaus<sup>1,\*</sup>  
 Jianjun Gan<sup>2</sup>  
 Weichao Bao<sup>1,\*</sup>  
 Yan D Zhao<sup>3</sup>  
 Dewi Rahardja<sup>3</sup>  
 John W Holloway<sup>5</sup>  
 Martha Scott<sup>4,5</sup>  
 Syed Hasan Arshad<sup>4,5</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, The University of South Carolina, Columbia, SC, USA; <sup>2</sup>GlaxoSmithKline, Research Triangle Park, NC, USA; <sup>3</sup>Department of Clinical Science and Simmons Cancer Center, UT Southwestern Medical Center, Dallas, TX, USA; <sup>4</sup>The David Hide Asthma and Allergy Research Center, St Mary's Hospital, Newport, Isle of Wight, UK; <sup>5</sup>School of Medicine, University of Southampton, Southampton, UK

\*These authors contributed equally to this article

**Purpose:** Skin prick testing (SPT) is fundamental to the practice of clinical allergy identifying relevant allergens and predicting the clinical expression of disease. Wheal sizes on SPT are used to identify atopic cases, and the cut-off value for a positive test is commonly set at 3 mm. However, the measured wheal sizes do not solely reflect the magnitude of skin reaction to allergens, but also skin reactivity (reflected in the size of histamine reaction) and other random or non-random factors. We sought to estimate wheal sizes exclusively due to skin response to allergens and propose gender-specific cutoff points of atopy.

**Methods:** We developed a Bayesian method to adjust observed wheal sizes by excluding histamine and other factor effects, based on which revised cutoff points are proposed for males and females, respectively. The method is then applied to and intensively evaluated using a study population aged 18, at a location on the Isle of Wight in the United Kingdom. To evaluate the proposed approach, two sample *t*-tests for population means and proportion tests are applied.

**Results:** Four common aeroallergens, house dust mite (HDM), grass pollen, dog dander, and *alternaria* are considered in the study. Based on 3 mm cutoff, males tend to be more atopic than females (*P*-values are between 0.00087 and 0.062). After applying the proposed methods to adjust wheal sizes, our findings suggest that misclassifications of atopy occur more often in males. Revised allergen-specific cutoff values are proposed for each gender.

**Conclusion:** To reduce the gender discrepancy, we may have two potentially convenient solutions. One way is to apply allergen-specific and gender-specific cutoff values following the proposed method. Alternatively, we can revise the concentration of allergens in the SPT solutions but keep the cutoff values unchanged, which may be more convenient to clinicians.

**Keywords:** SPT, atopy, Bayesian method, joint modeling, misclassification

## Introduction

Atopy is regarded as an inherited predisposition for diseases such as eczema, asthma, or rhinitis. Skin prick testing (SPT) to allergens is commonly used to identify allergic sensitization or atopy. Although some previous studies have suggested different cutoff values,<sup>1,2</sup> in clinical practice, a 3 millimeter (mm) cutoff wheal size on SPT is the criterion generally used to define a positive response and, therefore, sensitization. An atopic status is defined as the presence of sensitization to one or more allergens using this cutoff.

The present work is motivated by a discrepancy observed in our birth cohort between atopy and atopy-related diseases such as eczema. The cohort was established between 1989 and 1990 on the Isle of Wight (IOW) in the United Kingdom to prospectively study the natural history of allergic disorders. Skin prick testing was performed on most participants attending the research center to a standard battery of

Correspondence: Hongmei Zhang  
 Department of Epidemiology  
 and Biostatistics, The University  
 of South Carolina, Columbia,  
 South Carolina, 29208 USA  
 Tel +1 803-777-3823  
 Fax +1 803-777-2524  
 Email hzhang@sc.edu

common allergens (ALK, Horsholm, Denmark), including aeroallergens (house dust mite, cat dander, dog dander, *Alternaria*, *Cladosporium herbarium*, grass pollen mix, and tree pollen mix) and food allergens (cows' milk, soya, hens' egg, peanut, and cod). In this work, we focus on individuals aged 18, which includes 405 males and 445 females. Table 1 provides sensitization prevalence for four allergens, house dust mite (HDM), grass pollen, dog dander, and *Alternaria*, together with percentages of eczema among sensitized cases. As shown in the table, with atopic status determined by the 3 mm cutoff, significantly more males are sensitized than females based on two-sided two-sample proportion tests (significance level is set at 0.01 after adjusting for multiple testing). However, among the atopic children, the proportion of males who developed eczema tends to be lower than (but not statistically significant based on the same type of tests) that of females, although more males are atopic. For instance, about 35% of males are sensitized to HDM, which is significantly higher than that of females (23.37%). However, among the atopic males, about 14% developed eczema, which is lower than that of females (21.15%). We also examined the other two allergic diseases, asthma and rhinitis. Comparable patterns are observed (results not shown). Although for allergen *Alternaria*, higher proportions of atopic males developed asthma and/or rhinitis, the differences are statistically insignificant. These observations (more atopic males but fewer with allergic diseases) conflict with the well-established positive association between atopy and atopy related diseases such as eczema and asthma.<sup>3-5</sup> These observed discrepancies made us wonder whether atopic status was misclassified.

To examine the existence of atopy misclassification, we first compared total immunoglobulin E (IgE) between gender in our cohort. The result shows insignificant difference between males and females (the average total IgE for males is 296.0 ng/ml vs 287.4 ng/ml for females;  $P$ -value = 0.23 from two-sided two sample  $t$ -test for population means), which

agrees with findings in other studies.<sup>6</sup> In addition, results from another cohort<sup>7</sup> along with findings from other studies examining the agreement of atopic status determined by SPT and by specific IgEs reached the same conclusion<sup>8</sup>. That is, insignificant differences of atopy prevalence exist between gender based on specific IgEs but significant differences are found based on SPT. Specific IgEs and total IgE are antibody classes regarded as an important factor in the pathogenesis of allergic diseases and higher IgE measures indicate higher probabilities of allergic sensitization. This implies that we ought to expect insignificant gender difference in atopy in general populations. This conflicts with our findings on SPT testing results and consequently indicates the possible existence of misclassifications from SPT.

Since skin reactivity plays a role in the determination of wheal sizes and this reactivity, reflected in histamine wheal size, varies between gender at any given age,<sup>1,9-11</sup> cutoff points determined based on SPT wheal sizes without adjustment for these factors may not correctly identify allergic sensitization or determine atopic status. Thus to reduce the possibility of misclassification, two venues may be taken: adjusting the wheal sizes or proposing revised cutoff values that are gender-specific. Both directions aim to correct misclassifications, specifically non-differential misclassifications (misclassifications independent from disease or exposure status). Non-differential misclassifications can cause misleading inferences if left unchecked. In linear or logistic regressions, such errors may lead to biased estimates of coefficients.<sup>12-14</sup> In many situations, the misclassifications are actually caused by mismeasured continuous variables.<sup>15</sup> This can be the situation of SPT wheal size measures in the sense that they are not exclusively a result of allergen reaction; skin reactivity also contributes to the size of a wheal.

In this article, through a Bayesian hierarchical joint modeling, we first infer wheal sizes in response exclusively to allergens, and then propose gender specific cutoff values for sensitization. The utilization of the Bayesian method was

**Table 1** Comparison between gender of atopy prevalence and percentage of atopic cases having eczema based on 3 mm cutoff (405 males and 445 females)

Allergen	Sensitization prevalence (%)			% of atopic cases with eczema		
	Males	Females	$P$ -value	Males	Females	$P$ -value
HDM	35.06	23.37	0.00087	14.08 (142)	21.15 (104)	0.073
Grass	26.73	21.34	0.033	13.89 (108)	18.95 (97)	0.16
Dog dander	12.35	8.99	0.056	26.00 (50)	25.00 (40)	0.46
<i>Alternaria</i>	9.63	6.74	0.062	17.95 (39)	20.00 (30)	0.41
All above	46.53	34.61	0.00013	12.23 (188)	20.13 (154)	0.023

**Notes:** The numbers of positive sensitization are included in the parentheses. Two-sided two sample proportion tests are used to test gender differences between percentages for each allergen. For tests in columns 4 and 7, respectively, multiple testing adjusted significance level is 0.0125 ( $0.05/5 = 0.01$ ) using the Bonferroni method.  $P$ -values are listed to show the significance of percentage differences between gender.

motivated by the hierarchical structure between atopy and atopy related diseases. Due to the similar discrepancy patterns observed in asthma, eczema, and rhinitis, throughout this article, we use eczema to demonstrate the method. In the Discussion section, we briefly summarize results from the other two allergic diseases. We focus on four common aeroallergens, HDM, grass pollen, dog dander, and *Alternaria*. These four allergens are well represented in the IOW cohort data. Other allergens are not considered due to the sparsity of positive SPT reactions or their cross-reactivity with these four allergens. We expect the work has the potential to resolve the disagreement noted above, which will consequently improve the diagnosis and management of allergic diseases.

## Material and methods

We start this section by presenting the modeling of observed wheal sizes and that of the association between eczema and wheal sizes in reaction to allergen (hereafter, true wheal sizes). From the discussion above, the atopy prevalence among males is expected to agree with atopy prevalence among females in general populations, and there should not exist gender differentiated association between true wheal sizes and the risk of eczema.

## Model construction

In the following, we discuss the modeling separately for observed wheal sizes, true wheal sizes, and the risk of eczema. In the next section, they are linked through a joint modeling process.

### The observed wheal size

Let  $O_{ij}$  denote the observed wheal size of person  $i$  with gender  $j$  ( $j = 1$  for females) in reaction to an allergen. The observed wheal size is a mixture of response to the allergen, skin reactivity and other unknown random factors. We use  $T_{ij}$  to denote the latent (unknown) true wheal size. The difference between the expected value of  $O_{ij}$  and  $T_{ij}$ , denoted as  $F_{ij}$ , is modeled as a function of histamine effects (which reflects skin reactivity) and possible interaction effects between the allergen and histamine. We formulate them as follows:

$$\begin{aligned} O_{ij} &= T_{ij} + F_{ij} + \epsilon_{ij} \\ F_{ij} &= \alpha_i + \beta_{ij} H_{ij} + \beta_{2ij} T_{ij} \times H_{ij} \end{aligned} \quad (1)$$

where  $\alpha_j$  denotes an overall gender effect on the observed wheal sizes and a constraint  $\alpha_1 + \alpha_2 = 0$  is applied to avoid singularity,  $\beta_{ij}$  is a random slope for the contribution of histamine  $H_{ij}$ , and  $\beta_{2ij}$  is treated as being random as well indicating an interaction effect between allergen and histamine. The

interaction is denoted as  $T_{ij} \times H_{ij}$  in model (1). The random coefficients are assumed to be normally distributed with  $\beta_{1ij} \sim N(0, \sigma_{\beta_1}^2)$  and  $\beta_{2ij} \sim N(0, \sigma_{\beta_2}^2)$ . Note that gender-specific variances are assumed in the distributions of  $\beta_{ij}$  and  $\beta_{2ij}$ . These flexible assumptions allow gender-specific effects of histamine and its interaction with  $T_{ij}$ , which were motivated by the findings from our cohort (at age 18, the average histamine wheal size of males is 5.46 cm and of females is 5.09 cm;  $P$ -value = 0.00004 based on two-sided two sample  $t$ -tests for population means). Finally,  $\epsilon_{ij}$  explains unknown random effects on  $O_{ij}$  and is assumed to be normally distributed with mean zero and variance  $\sigma_1^2$ .

### The true wheal size

The latent variable  $T_{ij}$  in (1) represents wheal sizes in reaction to an allergen (the true wheal size). We model  $T_{ij}$  as

$$T_{ij} = \gamma + \delta_{ij} \quad (2)$$

in which  $\gamma$  denotes what we expect on wheal sizes for a general population in reaction to an allergen, regardless of gender. The second term  $\delta_{ij}$  is for random errors. It represents possible differentiated reaction to the allergen at an individual level. We assume  $\delta_{ij}$  is half-normally distributed, that is,

$$\delta_{ij} \sim \text{Half} - N(0, \sigma_2^2),$$

where zero and  $\sigma_2^2$  denote the mean and variance in the corresponding normal distribution function. The density of the half-normal distribution defined above is in a shape of a half bell-curve starting at zero. This distribution function is utilized to reflect the fact that most subjects are non-atopic. Besides the half normal distribution, other skew distributions can be possibly applied, for instance, the skew normal distribution.<sup>16,17</sup>

The message conveyed by model (2) is that the wheal sizes of males and females are expected a priori to follow the same distribution. However, if  $F_{ij}$  and  $\epsilon_{ij}$  in (1) are not sufficient to explain gender discrepancy, the skew-distributed random error  $\delta_{ij}$  can still possibly differentiate males' and females' true wheal sizes. In this case, data for different genders are likely to be generated from different parts of a half-normal distribution.

### The association between eczema and atopy

Since atopy is considered as a risk factor of eczema, modeling the association between true wheal sizes ( $T_{ij}$ ) and eczema seems a reasonable instrument in the process of identifying misclassifications. A logistic regression given below is considered:

$$\text{logit}(P(Y_{ij} = 1)|T_{ij}) = \eta_0 + \tau_j + \eta_1 T_{ij}, \quad (3)$$

where  $Y_{ij}$  takes values 1 or 0 denoting the status of eczema. Equation (3) evaluates gender effect  $\tau_j$  (assuming  $\tau_1 + \tau_2 = 0$  to avoid singularity) and the effect of true wheal size ( $\eta_1$ ) to the odds of eczema.

## Statistical analysis

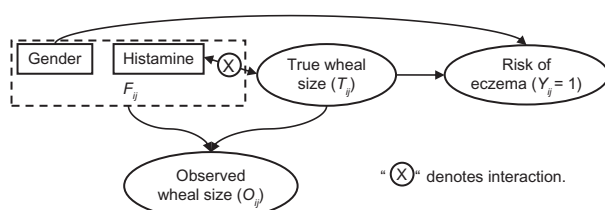
We consider a method of joint modeling to infer the true wheal size. Joint modeling is appropriate for data sharing features in common. It has been used in analysis of combined longitudinal and survival data,<sup>18–20</sup> random effects data,<sup>21</sup> and mark-recapture data.<sup>22</sup> The joint model in our analysis is composed of two parts: the modeling of observed wheal sizes and the modeling of association between the true wheal size and the risk of eczema. The models presented in (1) to (3) are linked to each other with  $T_{ij}$  being the joint. We include a brief structure of the joint model in Figure 1.

Let  $f(\cdot)$  denote a generic density function. Equations (1) to (3) induce a joint density of  $Y_{ij}$ ,  $O_{ij}$ , and  $T_{ij}$  in a hierarchical structure, which is

$$\begin{aligned} f(Y_{ij}, O_{ij}, T_{ij}) &= P(Y_{ij} = 1 | T_{ij}, O_{ij}) f(O_{ij} | T_{ij}) f(T_{ij}) \\ &= P(Y_{ij} = 1 | T_{ij}) f(O_{ij} | T_{ij}) f(T_{ij}) \end{aligned} \quad (4)$$

Here we suppress the dependence on the unknown parameters for simplicity. The last equality is due to the assumption that the observed wheal size does not provide any additional information at the presence of true wheal size (non-differentiable measurement errors). The advantage of a joint analysis is that more sources of information can be incorporated, and thus the inference on true wheal sizes is expected to be more accurate. The joint model presented in (4) allows us to infer  $T_{ij}$  based on information from two sources, the disease status  $Y_{ij}$  and observed wheal size  $O_{ij}$ . Inferences of  $T_{ij}$  will be further used to facilitate the revision of the 3 mm cutoff point.

Let  $\theta = \{\alpha_j, \sigma_{\beta_j}^2, \sigma_{\beta_2}^2, \sigma_1^2, \sigma_2^2, \gamma, \eta_0, \eta_1, \tau_j\}$  denote a collection of parameters. The hierarchical structure shown



**Figure 1** Conceptual structure of the joint model.

**Note:** Squares represent constant (fixed) and ovals represent stochastic variables.

in (4) motivated us to utilize the Bayesian method to infer the parameters and true wheal sizes  $T_{ij}$ . Here we briefly discuss the steps necessary to draw Bayesian inferences. To be fully Bayesian, we assign prior distributions to each parameter. Prior distributions of coefficients including  $\alpha_j$ ,  $\gamma$ ,  $\eta_0$ ,  $\eta_1$  and  $\tau_j$  are selected as vague normal distributions with mean zero and large variances. Prior distributions of variance components including  $\sigma_{\beta_j}^2$ ,  $\sigma_{\beta_2}^2$ ,  $\sigma_1^2$  and  $\sigma_2^2$  are assumed to follow inverse gamma distributions with scale and shape parameters being 0.5 and 0.0005, respectively.<sup>23</sup> Once prior distributions are specified, the joint posterior distribution of the parameters can be formulated using equations (1) to (4) in a hierarchical way. Next, we use the Markov Chain Monte Carlo (MCMC) approach, specifically, the Gibbs sampler to draw samples for each parameter and  $T_{ij}$  from the joint posterior distribution. The convergence of the MCMC chains is examined using the method proposed in previous studies.<sup>24–26</sup> The program is coded in WinBUGS.<sup>27</sup> A detailed graphical structure with all parameters included are given in the Appendix (Figure 3) along with the corresponding WinBUGS program (Figure 4).

## Defining adjusting factor $F_j$ and cutoff value $C_j^o$

The proposed approach draws inferences on the true wheal sizes  $T_{ij}$ . In practice, clinicians or epidemiologists may be more interested in handy cutoff values, which is practically more meaningful. Based on posterior inferences of  $T_{ij}$  and  $F_{ij}$ , we propose revised cutoff points  $C_j^o$  applied to  $O_{ij}$  for males and females, respectively. The revised cutoff value is defined as  $C_j^o = 3 - \hat{F}_j$  according to (1),  $j = 1, 2$ , where the 3 mm is the current cutoff value and  $\hat{F}_j$  denotes an estimate of adjusting factor  $F_j$ . The adjusting factor can be estimated by taking the means of posterior estimates of  $F_{ij}$  for each gender; recall  $F_{ij}$  in (1) represents an adjustment to each individually observed wheal size. It can also be estimated as the difference between  $O_j$  and  $\hat{T}_j$  for gender  $j$ , where  $O_j$  is the mean of observed wheal sizes  $O_{ij}$ ,  $\hat{T}_j$  is the estimate of mean true wheal size for gender  $j$  and is calculated as the sample mean of the inferred true wheal sizes  $T_{ij}$ .

## Examining the proposed method.

To evaluate the insights brought in by  $T_{ij}$  and  $C_j^o$ , we examine if the discrepancy is reduced or eliminated between males' and females' prevalence in atopy and in eczema among atopic cases. We apply two-sided two sample proportion tests to test the difference between genders of atopy prevalence and of eczema prevalence among atopic cases. We then compare the results from  $T_{ij}$  and  $C_j^o$  to those from  $O_{ij}$  and the 3 mm cutoff



value. We also evaluate the difference of  $T_j$  between genders, and compare it with the difference of  $O_j$  between genders. For this purpose, two-sided two sample  $t$ -tests applied to population means are applied. For each type of test, the Bonferroni approach is used for multiple testing corrections. The experiment-wise significance level is set at 0.05.

## Results

### Inferences of $T_{ij}$ and $C_j^o$

Using the proposed method, we infer  $T_{ij}$  and the adjusting factors for each of the four allergens (HDM, grass pollen, dog dander, and *Alternaria*). Table 2 includes descriptive statistics (mean and standard deviation) of  $O_{ij}$  and  $T_{ij}$  for each gender, denoted as  $O_j$  and  $\hat{T}_j$  in the table, respectively. The revised cutoff values  $C_j^o$  are presented in the last column of the table. As shown in the table, overall the results are consistent across different allergens for each gender. For females,  $\hat{T}_j$  agrees with  $O_j$ , and  $C_j^o$  agrees with the commonly used 3 mm cutoff point. However, this is not the case for males. For males,  $\hat{T}_j$ 's are all smaller than  $O_j$ 's and larger cutoff values are suggested.

By comparing the differences of  $O_j$  between genders and the differences of  $\hat{T}_j$  between genders, we can see that the adjusting process clearly has the ability to reduce the disagreement of wheal sizes between males and females. In

particular, as indicated by the  $P$ -values, the disagreement of wheal sizes between gender based on  $T_{ij}$  is eliminated for all allergens except for HDM. The significant differences between genders for HDM implies that besides histamine reactivity effect and its interaction with HDM, other unknown but non-random factors may also contribute to the formation of discrepancy. Further studies are needed to identify those possible factors. On the other hand, as indicated by the reduction of  $P$ -values, although gender discrepancy is not eliminated for HDM, the probability of observing such difference between males and females under the null hypothesis is increased compared to that based on  $O_{ij}$ . This implies that on average the dissimilarity between males' and females' wheal sizes is reduced based on the inferred true wheal sizes.

The inferences of  $T_{ij}$  discussed so far are promising in that gender discrepancy in wheal sizes is significantly reduced. However, the reduction of disagreement in wheal sizes does not necessarily lead to a discrepancy reduction in the prevalence of sensitization to an allergen. This is examined in the following section.

### Prevalence comparison based on $T_{ij}$

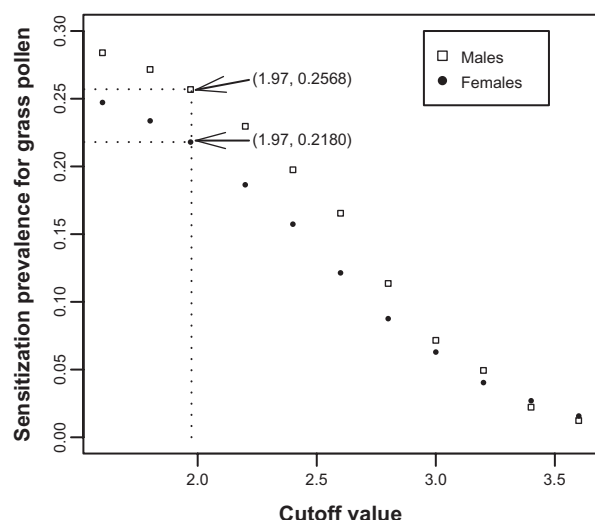
To compare sensitization prevalence between gender using  $T_{ij}$  (true wheal size), a cutoff value  $C_j^t$  applied to  $T_{ij}$  is needed. We use inferences related to grass pollen to demonstrate the  $C_j^t$  selection process. Figure 2 plots the prevalence of sensitization based on  $T_{ij}$  versus different cutoff values. The difference of prevalence between genders decreases as the cutoff value increases, which implies the importance of choosing an appropriate cutoff value. As indicated in Table 2, the revised cutoff values for females (the last column of Table 2) in general agrees with the commonly applied 3 mm cutoff point. Our paired tests further indicate that  $T_{i1}$  agrees with  $O_{i1}$  in all the four allergens (all  $P$ -values  $> 0.05$ . For this grass pollen example, the  $P$ -value is 0.79). These results imply that for females the sensitization prevalence based on  $O_{i1}$  is expected to agree with the prevalence based on  $T_{i1}$ . As discussed below, this finding is then utilized to determine an atopy cutoff value for  $T_{ij}$  applied to both genders, since there is no significant difference of  $T_{ij}$  between genders.

The prevalence of atopy for females based on  $O_{ij}$  is 21.34% as in Table 1. Given the agreement between  $O_{i1}$  and  $T_{i1}$ , we apply this prevalence to  $T_{ij}$  resulting in a cutoff value  $C_j^t = 1.97$  mm. Using this cutoff value, the sensitization prevalence for males is 25.68% and for females is 21.80% (different from 21.34% due to rounding errors), which are insignificantly different from each other ( $P$ -value = 0.09). The same proce-

**Table 2** Summary of observed and inferred wheal sizes and proposed cutoff values (The unit of each variable is in millimeters. Standard deviations are in the parentheses)

Gender	$O_j$	$\hat{T}_j$	$\hat{F}_j$	$C_j^o$
<b>HDM</b>				
Males	2.08 (2.78)	1.66 (0.98)	-0.43	3.43
Females	1.36 (2.37)	1.48 (0.97)	0.10	2.90
$P$ -value	0.000065	0.0061		
<b>Grass</b>				
Males	1.80 (3.03)	1.33 (0.89)	-0.51	3.51
Females	1.23 (2.36)	1.25 (0.86)	0.01	2.99
$P$ -value	0.0023	0.14		
<b>Dog dander</b>				
Males	0.77 (1.54)	0.59 (0.36)	-0.18	3.18
Females	0.55 (1.27)	0.57 (0.35)	0.02	2.98
$P$ -value	0.021	0.31		
<b>Alternaria</b>				
Males	0.61 (1.66)	0.39 (0.26)	-0.22	3.22
Females	0.39 (1.31)	0.38 (0.25)	0.01	2.99
$P$ -value	0.036	0.42		

**Notes:**  $j = 1, 2$  denotes two genders.  $O_j$ : the average of observed wheal sizes.  $\hat{T}_j$ : the average of inferred true wheal sizes.  $\hat{F}_j$ : estimate of the adjusting factor.  $C_j^o$  is the revised cutoff point. Two F sample  $t$ -tests were performed to test the differences of means between gender. In columns 2 and 3, respectively, multiple testing adjusted significance level is 0.0125 ( $0.05/4 = 0.0125$ ) using the Bonferroni method.  $P$ -values are listed to show the significance of mean difference of wheal sizes between genders.



**Figure 2** The relationship between sensitization prevalence of wheat and cutoff value based on  $T_{ij}$

**Table 3** Cutoff values and sensitization prevalence comparison between gender based on  $T_{ij}$

Allergen	Cutoff value (mm)	Sensitization prevalence (%)		
		Males	Females	P-value
HDM	2.4	28.89	22.92	0.023
Grass	1.97	25.68	21.80	0.09
Dog dander	1.28	8.40	9.21	0.34
<i>Alternaria</i>	0.99	6.17	6.97	0.32

**Notes:** Two-sided two sample proportion tests are used to test gender differences between percentage for each allergen. Multiple testing adjusted significance level is 0.0125 (0.05/4) using the Bonferroni method. P-values are listed to show the significance of prevalence differences between gender.

**Table 4** Eczema cases with positive sensitization based on  $T_{ij}$

Allergen	% of atopic cases with eczema		
	Males	Females	P-value
HDM	16.24 (117)	22.55 (102)	0.12
Grass	13.72 (102)	19.59 (97)	0.13
Dog dander	41.18 (34)	29.26 (41)	0.14
<i>Alternaria</i>	16.00 (25)	19.35 (31)	0.37

**Notes:** The numbers of positive sensitization are included in the parentheses. Two-sided two sample proportion tests are used to test gender differences between percentage for each allergen. Multiple testing adjusted significance level is 0.0125 (0.05/4) using the Bonferroni method. P-values are listed to show the significance of percentage differences between gender.

**Table 5** Comparison of prevalence between gender based on  $C_j^o$  of sensitization and percentage of atopic cases having eczema

Allergen	Sensitization prevalence (%)			Prevalence of eczema among atopic cases (%)		
	Males	Females	P-value	Males	Females	P-value
HDM	32.35	23.37	0.0017	15.27	21.15	0.12
Grass	24.75	21.35	0.12	14.00	18.95	0.18
Dog dander	10.62	8.99	0.21	27.91	25.00	0.38
<i>Alternaria</i>	8.89	6.74	0.12	19.44	20.00	0.47

**Notes:** Two-sided two sample proportion tests are used to test gender differences between percentages for each allergen. For tests in columns 4 and 7, respectively, multiple testing adjusted significance level is 0.0125 (0.05/4 = 0.0125) using the Bonferroni method. P-values are listed to show the significance of percentage differences between gender.

procedure is repeated for the remaining three allergens. The cutoff values  $C_j^t$  and the comparison of sensitization prevalence between genders are summarized in Table 3. After adjusting for multiple testing using the Bonferroni method (multiple testing adjusted significance level = 0.0125), the differences of atopy prevalences between genders are no longer statistically significant for any single allergen. Utilizing information summarized in Table 1, we then identified the individuals, especially males, who are misclassified into the atopic group by applying the 3 mm cutoff to the observed wheal sizes  $O_{ij}$ . Specifically, among 142 males originally classified into the HDM sensitization group, 25 (17.61%) are misclassified; results of misclassification for grass pollen, dog dander, and *Alternaria* are 4 (3.70%) out of 108, 16 (32%) out of 50, and 14 (35.90%) out of 39, respectively.

So far our focus is on the prevalence of atopy. We now turn to the comparison of eczema prevalence among atopic cases. The results are given in Table 4. Comparing the prevalence differences between genders, we can see that the results of insignificant differences drawn from  $O_{ij}$  (see Table 1) are kept for  $T_{ij}$ . This finding, coupled with the findings on  $T_{ij}$ -based atopy prevalence, demonstrates the applicability of the proposed adjusting process. This process has the potential to correct the bias in wheal size measuring and resolve the conflict between atopy prevalence and findings on eczema prevalence and IgE measures.

In an early section, we proposed cutoff values  $C_j^o$  applied to  $O_{ij}$ . Even with the promising gain from inferring  $T_{ij}$  to clinicians and epidemiologists, it is possibly more convenient to use  $C_j^o$ . In the next section we examine if we can reach the same conclusion on the elimination of gender discrepancy by using  $C_j^o$ .

## Prevalence comparison based on $C_j^o$

We apply the revised cutoff value  $C_j^o$  given in Table 2 to the observed wheal sizes to infer the prevalence of sensitization and that of eczema among atopic cases. The results are presented in Table 5. Compared to the results given in Table 1,

**Table 6** Comparison of prevalence between gender of atopy and percentage of atopic cases having eczema based on different cutoffs on all allergens

Allergen	Sensitization prevalence (%)			Prevalence of eczema among atopic cases (%)		
	Males	Females	P-value	Males	Females	P-value
3 mm (males and females)	47.89	35.36	0.00011	11.92	20.38	0.00010
3.5 mm (males), 3 mm (females)	45.91	35.36	0.00089	11.89	20.38	0.00023

**Notes:** Two-sided two sample proportion tests are used to test gender differences between percentages for each allergen. For tests in columns 4 and 7, respectively, multiple testing adjusted significance level is 0.0125 ( $0.05/4 = 0.0125$ ) using the Bonferroni method. P-values are listed to show the significance of percentage differences between gender.

the updated prevalence differences between genders for each allergen are less significant, although for HDM the significant gender discrepancy still exists. The prevalence of eczema among atopic cases is not significantly different between males and females for each individual allergen, the same trend as in Table 1.

By using although the gender discrepancy cannot be completely eliminated for all allergens and the results are not as promising as those based on  $T_{ij}$  (Tables 3 and 4), the revised cutoff value  $C_j^o$  does have the potential to decrease the significance of gender discrepancy in sensitization prevalence.

## Discussion

Motivated by the inconsistency between the wheal size-based atopy prevalence in men and women and the results of IgE measures, we developed a Bayesian method to estimate true wheal sizes and proposed gender-specific sensitization cutoff values for different allergens. The disease outcome considered in this work is eczema. The allergens considered are common aeroallergens including house dust mite, grass pollen, dog dander, and *Alternaria*.

Based on inferred true wheal sizes ( $T_{ij}$ ), the differences in sensitization prevalence between males and females are statistically insignificant with respect to each individual allergen, and the chance that sensitized males developing eczema is comparable to that of sensitized females. This indicates a significant improvement compared to the results based on observed wheal sizes with the 3 mm cutoff. On the other hand, if we apply the revised cutoff values  $C_j^o$  to observed wheal sizes (other than directly using inferred  $T_{ij}$ ), our results showed limited improvement. It thus seems more reasonable to adjust the observed wheal sizes than to revise the 3 mm cutoff value for each allergen and gender. The findings can be further assessed by calculating the sensitivity and specificity with the help of gold standard such as radioallergosorbent test proposed in early studies.<sup>28,29</sup> We also applied the method to two other allergic diseases, asthma and rhinitis. Similar results are drawn for allergens HDM, grass pollen, and dog

dander. For *Alternaria*, results from the method suggest adjusting wheal sizes is unnecessary.

On the other hand, since utilizing revised cutoff values ( $C_j^o$ ) does reduce the differences between males and females, although the reduction is not significant in some situations, one may wonder the possibility to propose cutoff values separately for males and females but apply to all allergens. We further investigated this possibility. Based on the results given in Table 2, we used 3.5 mm cutoff for males and 3 mm cutoff for females. These cutoff values are applied to all allergens. However, as indicated by the results in Table 6, very limited reduction in the statistical significance is observed. Similar results are obtained when using other cutoff values different from 3.5 mm. This finding indicates that just revising cutoff points applied to all allergens may not at all solve the prevalence discrepancy between males and females. Instead, we may have to deal with each individual allergen.

## Conclusion

To correct misclassified atopic cases caused by gender discrepancy, we can utilize the proposed statistical methods to adjust wheal size measures. Besides adjusting observed wheal sizes, we found that using allergen-specific cutoff values different for men and women ( $C_j^o$ ) will also reduce the occurrence of misclassifications.

Clinicians may prefer the same cutoff value applied to all allergens. To achieve this goal and keep misclassification reduced, a laboratory-related alternative solution may be possible. Specifically, based on the inferred cutoff values  $C_j^o$ , one can adjust the concentration of allergens used in the prick test solutions for men and women, respectively. The results from this work may assist the determination of allergen concentration. This approach makes it possible to reduce the gender discrepancy in atopy without revising cutoff values. Our work shows that future clinical, mechanistic, and epidemiological studies are needed to optimize the skin prick test to make it agree with results from specific IgE levels.

## Acknowledgment

This work was supported by the National Heart, Lung, and Blood Institute at the National Institute of Health (5R01 HL082925, PI: Arshad SH; 1R03HL095429, MPI: Zhang H).

## Disclosure

The authors report no conflict of interest in this work.

## References

- Gergen P, Turkeltaub P, Kovar M. The prevalence of allergic skin test reactivity to eight common aeroallergens in the US population: results from the second national health and nutrition examination survey. *J Allergy Clin Immunol*. 1987;80:669–679.
- Chinn S, Jarvis D, Luczynska C, et al. Measuring atopy in a multi-centre epidemiological study. *European Journal of Epidemiology*. 1996;12:155–162.
- Arshad S, Karmaus W, Matthews S, et al. Association of allergy related symptoms with sensitization to common allergens in an adult european population. *J Inves Allergology Clin Immunol*. 2001;11:94–102.
- Sly P, Boner A, Björkstén B, et al.
- Gergen P, Arbes S Jr, Calatroni A, et al. Total ige levels and asthma prevalence in the us population: results from the national health and nutrition examination survey 2005–2006. *Journal of Allergy and Clinical Immunology*. 2009;124:447–453. doi:DOI: 10.1016/j.jaci.2009.06.011.
- Shoormasti R, Pourpak Z, Eshraghian M, et al. The study of total ige reference range in healthy adults in tehran, iran. *Iranian J Public Health*. 2010;39:32–36.
- Van Eerdewegh P, Little R, Dupuis J, et al. Association of the adam33 gene with asthma and bronchial hyperresponsiveness. *Nature*. 2002;418:426–430.
- Bousquet P, Castelli C, Daures J, et al. Assessment of allergen sensitization in a general population-based survey (european community respiratory health survey i). *Ann Epidemiol*. 2010;20:797–803.
- Barbee R, Lebowitz M, Thompson H, et al. Immediate skintest reactivity in a general population sample. *Ann Intern Med*. 1976;84:129–133.
- Peat J, Woolcock A. Sensitivity to common allergens: relation to respiratory symptoms and bronchial hyper-responsiveness in children from three different climatic areas of australia. *Clin Exp Allergy*. 1991;21:573–581.
- Meinert R, Frischer T, Karmaus W, et al. Influence of skin prick test criteria on estimation of prevalence and incidence of allergic sensitization in children. *Allergy*. 1994;49:526–532.
- Gustafson P. *Measurement error and misclassification in statistics and epidemiology: impacts and bayesian adjustments*. Chapman and Hall/CRC Press, Boca Raton, FL; 2003.
- Bross I. Misclassification in 2x2 tables. *Biometrics*. 1954;10:478–486.
- Goldberg J. The effects of misclassification on the bias in the difference between two proportions and the relative odds in the fourfold table. *Journal of the American Statistical Association*. 1975;70:561–567.
- Natarajan L. Regression calibration for dichotomized mismeasured predictors. *The International Journal of Biostatistics*. 2009;5: article 12.
- Azzalini A. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*. 1985;12:171–178.
- Azzalini A. Further results on a class of distributions which includes the normal ones. *Statistica (Bologna)*. 1986;46:199–208.
- Ding J, Wang J. Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data. *Biometrics*. 2008;64(2):546–556.
- Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics (Oxford)*. 2000;1(4):465–480.
- Tsiatis A, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*. 2004;14(3):809–834.
- Bigelow J, Dunson D. Bayesian Semiparametric joint models for functional predictors. *Journal of the American Statistical Association*. 2009;104(485):26–36.
- Barker RJ. Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics*. 1997;53:666–677.
- Kelsall J, Wakefield J. Discussion of “Bayesian models for spatially correlated disease and exposure data” by best et al. *Bayesian Statistics*. 1999;6:151.
- Gelman A, Rubin BD. Inference from iterative simulation using multiple sequences. *Statistical Science*. 1992;7:457–511.
- Gelman A, Rubin DB. A single series from the Gibbs sampler provides a false sense of security. In: Bernardo JM, Berger JO, Dawid AP, et al, editors, *Bayesian statistics 4 proceedings of the fourth valencia international meeting*. Clarendon Press: Oxford University Press; 625–631.
- Gelman A, Carlin JB, Stern HS, et al. *Bayesian Data Analysis*. Chapman and Hall/CRC; 2003.
- Spiegelhalter D, Thomas A, Best N, et al. *WinBUGS user manual*. MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK, 2003.
- Dreborg S. Skin tests used for epidemiological studies. *Allergy*. 1989;44:52–59.
- Haahela T. Skin tests used for epidemiologic studies. *Allergy*. 1993;48:76–80.



## Appendix

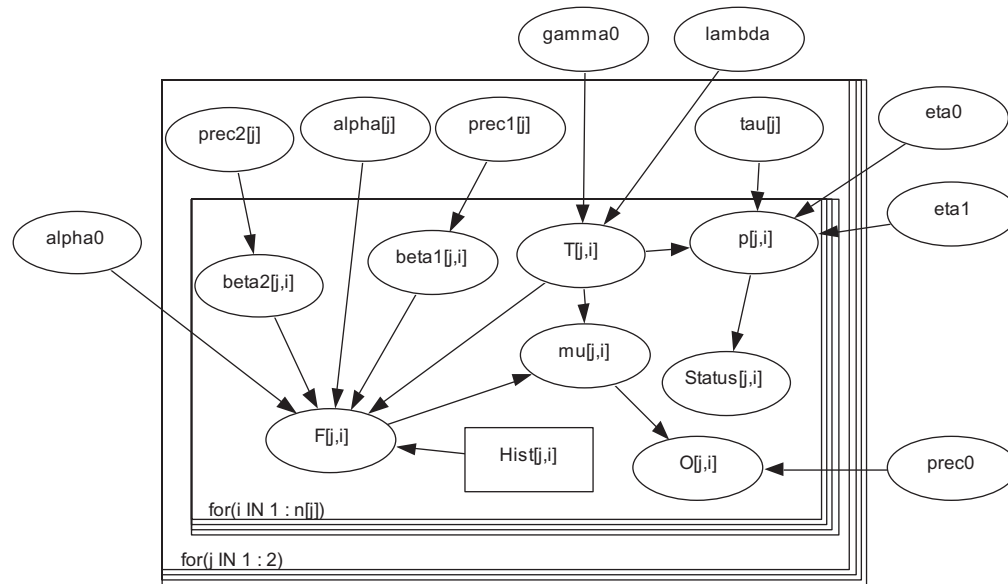
### The model structure

The structure in Figure 1 shows the formulation of the Bayesian joint model including models (1) to (3) along with the specified prior distributions. It is drawn using DoodleBUGS.<sup>29</sup> Squares represent constant (fixed), ovals represent stochastic or unknown variables, hollow (thicker) arrows are for logical functions (such as definitions or identities), and single arrows are for stochastic dependencies (distributions involved). There are two

panels in the figure. The smaller panel indicated by “for (i IN 1:n[j])” is for parameters and random variables related to each individual, the ovals between the bigger panel and the smaller one are for the parameters exclusively related to gender, and the ovals outside the two panels are for common parameters.

### The WinBUGS codes

The codes below are consistent with the structure given in Figure 1.



**Appendix Figure I** The detailed structure of the Bayesian joint model.

```

model
{
  for (j in 1:2) #j gender
  {
    for (i in 1: n[j])# in the observation
    {
      # logistic regression modeling the association between true wheal size
      # and disease outcome (status)
      status[j,i]~dbin(p[g,i], 1)
      logit(p[j,i])<-eta0+tau[j]+eta1*T[j,i]

      # linear regression evaluating factor effects contributed to the observed
      # wheal sizes
      O[j,i]~dnorm(mu[j,i], precO)
      F[j,i]<-alpha0+alpha[j]+beta1[j,i]*His t[j,i]+beta2[j,i]*Hist[j,i]*T[j,i]
      mu[j,i]<- T[j,i]+F[j,i]

      # half-normal distribution describing the distribution of true wheal size
      T[j,i]~djl.dnorm.trunc(gamma0,lambda,gamma0,1000)

      # prior distributions of the coefficients
      beta1[j,i]~dnorm(0, prec1[j])
      beta2[j,i]~dnorm(0, prec2[j])
    }
    # prior distribution for the precision parameters (inverse of variance) in the prior
    # distributions of beta1 and beta2.
    prec1[j]~dgamma(0.5, 0.5)
    prec2[j]~dgamma(0.5, 0.5)
  }
  # prior distributions for the coefficients in the logistic regression
  eta0~dnorm(0,0.01)
  eta1~dnorm(0,0.01)
  tau[1]~dnorm(0,0.01)
  tau[2]<--1*delta[1]

  # prior distributions for the precision parameter precO (inverse of the variance) in the
  # distribution of O_ij, overall effect alpha0, and gender effects alpha_j, j=1,2.
  precO~dgamma(0.5, 0.5)
  alpha0~dnorm(0,0.01)
  alpha[1]~dnorm(0,0.01)
  alpha[2]<--1*alpha[1]

  # prior distributions for the parameters in the half-normal distribution
  gamma0~dgamma(0.5, 0.5)
  lambda~dgamma(1,0.001)
}

```

**Appendix Figure 2** The WinBUGS program corresponding model structure in Appendix Figure 1.

## International Journal of General Medicine

### Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas.

Submit your manuscript here: <http://www.dovepress.com/international-journal-of-general-medicine-journal>

A key focus is the elucidation of disease processes and management protocols resulting in improved outcomes for the patient. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress