Finite Mixture Models in Assessing Anti-thyroglobulin Antibody Positivity as a Marker of Chronic Thyroiditis

Eiji Nakashima*1, Yoshinori Fujii*2, Misa Imaizumi*3 and Kiyoto Ashizawa*3

*1 Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732-0815, Japan
*2 Faculty of Education and Culture, Miyazaki University, 1-I Gakuen Kihanadai Nishi, Miyazaki 889-2155, Japan
*3 Department of Clinical Studies (Nagasaki) Radiation Effects Research Foundation, 1-8-6 Nakagawa, Nagasaki 850-0013, Japan
e-mail: nakashima@rerf.or.jp

Positivity of anti-thyroglobulin antibody (TgAb) is one of the markers of chronic thyroiditis (Hashimoto disease). From 2000 to 2003, a thyroid disease prevalence study was conducted at the Radiation Effects Research Foundation, in Hiroshima and Nagasaki. Utilizing the study’s results, we show that via EM algorithm log-transformed TgAb level is compatible with a two-component mixture normal distribution, with the smaller normal distribution corresponding to the TgAb negative group but the larger distribution not necessarily corresponding to the TgAb positive group. A subject is determined to be TgAb positive if TgAb level is greater than a given cutoff. We compared the cutoff values from population-based methods and the laboratory method. The population-based methods consist of a simple method, a receiver operating characteristic (ROC) curve method, and a minimum misclassification rate (MMR) method. The simple method is used to determine positivity from only TgAb negative populations. Since the ROC curve and MMR methods are valid only when TgAb positivity and negativity are known but the simple method is valid only when TgAb negativity is known, the simple method was deemed useful for determining the cutoff in our data. In comparison with the simple, population-based method, we show that the cutoff from the laboratory method is appropriate and that the TgAb positive rates from various methods are approximately equal. With the two-component mixture normal distribution in TgAb level, our simple population-based method for determination of cutoff is another more practical example of handling the clinical measurement than the method given in Thompson et al. (Applied Statistics 1998).

Key words: EM algorithm; misclassification; ROC; chronic thyroiditis; Hashimoto disease.

Received August 2007. Revised October 2007. Accepted November 2007.
1. Introduction

Antithyroglobulin antibody (TgAb) is a marker for chronic thyroiditis (Hashimoto disease), a common autoimmune disease characterized by lymphocytic infiltration of the thyroid gland. TgAb positivity is closely associated with a histological diagnosis of chronic thyroiditis (Kasagi et al. 1996). The positive rate of TgAb increases with age and, in the United States, ranges from 10% – 25% in women and 5% – 10% in men (Hollowell et al. 2002); TgAb concentration is not affected by treatments for hypothyroidism or lifestyle factors.

TgAb positivity is usually determined by a cutoff-point, that is, a subject is diagnosed as positive when the TgAb level lies outside the laboratory’s reference range. A binary logistic regression is made to estimate the population’s positive rate. TgAb is one of the biomarkers for which it is recommended that a cutoff-point be determined at each testing facility, taking into account the laboratory conditions and characteristics of subjects in the study population. Clinicians sometimes hesitate to apply the standard laboratory cutoff-point provided by the manufacturer (laboratory cutoff) to patients instead of the cutoff-point from the study population itself (population-based cutoff), although sometimes it is impractical to avoid using the laboratory cutoff. In this regard, we have compared the laboratory cutoff with the population-based cutoff for TgAb level.

We report that the log of TgAb levels shows a bimodal distribution. The 2-hr plasma glucose level (Omar et al. 1994; Thompson et al. 1998) is bimodally distributed, as is forced expiratory volume (Fujii and Nakano 2002). The hemoglobin A1c (HbA1c) level is also suspected to be bimodal since it reflects long term glucose levels.

In this paper, we log-transformed the measurements and, using the EM algorithm, divided the log of TgAb level distribution into two components with a mixing probability. In Thompson et al. (1998), the latent population-mixing variable was equal to the disease-indicator variable, the expectation of which corresponded to the prevalence. This result is probably because the diagnostic criterion of diabetes mellitus given as an example in Thompson et al. (1998) is based solely on the serum glucose level of subjects under certain conditions. In this paper, however, we treat cases in which the diagnostic criteria are essentially multiple. TgAb positivity is only one marker for chronic thyroiditis and does not necessarily define the condition. Therefore, in addition to a population mixing variable, we defined TgAb positivity as an indicator variable derived from the mixing distribution using three methods – the simple method (which uses only the TgAb negative population), the receiver operating characteristic (ROC) curve method, and the minimizing misclassification rate (MMR) method. We estimated the TgAb positive rate by comparing study population data with laboratory standards.

2. Data

The Adult Health Study (AHS) is a clinical program that was established in 1958 by the Radiation Effects Research Foundation (RERF) (formerly the Atomic Bomb Casualty Commission).
sion) to investigate the late effects of atomic-bomb survivors (Yokoro 1991). The subjects in the present study were 4,091 AHS cohort members (1,352 men and 2,739 women) who underwent biennial health examinations at RERF and who agreed to participate in thyroid disease screening between March 2000 and February 2003. All provided informed consent.

Blood samples were collected from subjects in Hiroshima or Nagasaki, and the serum was processed in the Nagasaki laboratory (serum of Hiroshima samples was frozen before being sent to Nagasaki). Serum samples were diluted as necessary, and the TgAb level was measured by enzyme linked immunosorbent assay (Medical & Biological Laboratory Co., Ltd, Nagoya, Japan; Kuroda et al. (2002)).

The cutoff point for TgAb positivity (determined according to the manufacturer's instructions) was 10 IU/ml, which was the laboratory standard determined by 120 measurements of healthy employees for whom the thyroid tests determined by the particle agglutination (PA) method were negative (Kuroda et al. 2002). Sex and age were not taken into account. The cutoff point was obtained as a mean +2 SD after transforming TgAb level into the log of the value, TgAb level plus one (to avoid log 0), which is a method for normalizing transformation indicated in the International Federation of Clinical Chemistry (IFCC) expert panel recommendation (Solberg 1987).

We used the Dosimetry System 2002 (Young and Bennett, 2006) to estimate the thyroid radiation doses. We excluded from the study those with unknown radiation doses and those who were exposed in utero, which left 3,185 participants. The dose distribution is shown in Imaizumi et al. (2006). The mean age at exposure was 15.4 yrs (range, 0-40), and the mean age at examination was 71.3 yrs (range, 55-97).

3. Statistical Methods

Since TgAb level distribution consist of TgAb positive and negative groups, TgAb level distribution would have two components. Therefore, we described the statistical method for 2-component mixture distribution for TgAb levels. In the application, we made goodness of fit chi-square test for a 2-component distribution. We based our statistical analysis on an EM algorithm (Dempster et al. 1977) for finite mixture distributions (McLachlan and Peel, 2001). The finite mixture model is a typical example of the EM algorithm (Louis 1982). Following Thompson et al. (1998), we formulated the EM algorithm for a 2-component mixture model and estimated the parameters for the 2-component populations and the mixing probability. Using the components and the mixing probability, we described three methods for estimating the TgAb positivity cutoff point and the positive rate. We calculated the standard errors for the cutoff point and the positive rate by the bootstrap method.
3.1 Log-likelihood

We assumed that the variable $Y$ is a bimodal directly observed random variable with two normal components with covariate $x$ for a subject. Let us assume a latent binary random variable $Z$, which is an unobserved variable that indicates the subject belongs to population pop(1) when $Z = 1$ and to population pop(0) when $Z = 0$. In our data, population pop(1) is mostly TgAb positive while population pop(0) is TgAb negative. An observation $y$ is sampled from pop(1) with probability $\pi$ or from pop(0) with probability $1 - \pi$. The data consist of size $n$. The log-likelihood for the observed data is a finite mixture log-likelihood,

$$l_{\text{obs}} = \sum_{i=1}^{n} \ln\{(1 - \pi_i)f_0(y_i) + \pi_if_1(y_i)\},$$

where $f_0(y) = \Pr(Y = y \mid Z = 0)$ and $f_1(y) = \Pr(Y = y \mid Z = 1)$ are the normal densities of pop(0) and pop(1), respectively, $\pi_i = \Pr(Z_i = 1)$, i.e., the $i$-th subject is from pop(1)) is a mixing probability of the two populations and $\ln(\cdot)$ is the natural log. Since variable $Z$ is missing, the estimation problem is a missing data problem, and we can employ the EM algorithm. If the variable $Z$ is observed, we have the joint density $\{(1 - \pi)\pi f_0(y)\}^{1-Z}\{\pi f_1(y)\}^{Z}$ for $(Z, y)$. Then we obtain the complete data log-likelihood (Thompson et al. 1998),

$$l_c = \sum_{i=1}^{n} \{(1 - Z_i)\ln(1 - \pi_i) + Z_i\ln\pi_i\} + \sum_{i=1}^{n} (1 - Z_i)\ln f_0(y_i) + \sum_{i=1}^{n} Z_i\ln f_1(y_i)$$

If the mixing probability and the two component densities are parameterized separately, this complete data log-likelihood can be maximized in each of three summations separately.

We denote the covariate column vectors for the pop(0), pop(1) and the mixing probability as $x_{0i}$, $x_{1i}$, and $x_{2i}$, respectively, for the $i$-th subject. Specifically we assume normal component distributions for pop(0) and pop(1), $f_0(y_i) = \phi((y_i - \mu_0)/\sigma_0)/\sigma_0$ with $\mu_0 = x_{0i}^T\beta_0$, $f_1(y_i) = \phi((y_i - \mu_1)/\sigma_1)/\sigma_1$ with $\mu_1 = x_{1i}^T\beta_1$, respectively, where $\phi(\cdot)$ is the standard normal density and $\beta_0$ and $\beta_1$ are regression parameter columns for the populations, and assume a logistic model for $\pi_i$, ie, $\text{logit}(\pi_i) = x_{2i}^T\beta_2$, where $\beta_2$ is a parameter column vector for the mixing probability of the two populations. Since pop(1) is possibly a TgAb positive population, we can naturally assume $\mu_0 < \mu_1$ for identifiability as well. The first summation of the complete data log-likelihood is a binary logistic log-likelihood and the second and the third are normal log-likelihoods with weight $1 - Z_i$ and $Z_i$, respectively.

3.2 Estimation of parameters via the EM algorithm

Dempster et al. (1977) showed that the incomplete data log-likelihood could be maximized by iteratively maximizing the conditional expectation of the complete data log-likelihood given the incomplete data. That is, three components of summation in the complete data log-likelihood are maximized in the EM algorithm after replacing $Z_i$ with the conditional mixing probability, $p_i = E(Z_i \mid Y_i = y_i) = \Pr(Z_i = 1 \mid Y_i = y_i)$, which can be written in terms of $\pi_i$, $f_0(y_i)$, and $f_1(y_i)$

$$p_i = E(Z_i \mid Y_i = y_i) = \Pr(Z_i = 1 \mid Y_i = y_i),$$

which can be written in terms of $\pi_i$, $f_0(y_i)$, and $f_1(y_i)$.
using Bayes Theorem. In the iteration, to avoid convergence to a local maximum, we set the starting values as $\pi_i^{(0)} = 1$ if $y_i > 1$ for log TgAb level and $\pi_i^{(0)} = 0$ otherwise with $\beta_0^{(0)} = 0$ and $\beta_1^{(0)} = 0$. We decided the convergence was obtained when the successive relative difference of the observed log-likelihood was less than $10^{-6}$. The EM algorithm has two steps, E and M steps.

**E-step:** The conditional mixing probability that observation $i$ comes from pop(1), $p_i$, is estimated at the $k$-th iteration as

$$p_i^{(k)} = \pi_i^{(k)} f_1^{(k)}(y_i)/\{(1 - \pi_i^{(k)}) f_0^{(k)}(y_i) + \pi_i^{(k)} f_1^{(k)}(y_i)\}. \quad (3)$$

**M-step:** The distribution of the two components pop(0) and pop(1) and the logistic mixing probability are generalized linear models and can be easily fit:

(a) Use $y_i$ as the response data in normal regression with covariate column vector $x_{0i}$ with weight $1 - p_i^{(k)}$ to estimate $\mu_{0i}^{(k+1)} = x_{0i}^T \beta_0^{(k+1)}$ and $\sigma_{0i}^{(k+1)}$. Calculate $f_0^{(k+1)}(y_i)$ using these estimates.

(b) Use $y_i$ as the response data in normal regression with covariate column vector $x_{1i}$ with weight $p_i^{(k)}$ to estimate $\mu_{1i}^{(k+1)} = x_{1i}^T \beta_1^{(k+1)}$ and $\sigma_{1i}^{(k+1)}$. Calculate $f_1^{(k+1)}(y_i)$ using these estimates.

(c) Use $p_i^{(k)}$ as the response in logistic regression with covariate column vector $x_{2i}$ to estimate the linear predictor $\eta_{i}^{(k+1)} = x_{2i}^T \beta_2^{(k+1)}$ and calculate $\pi_i^{(k+1)} = 1/(1 + \exp(-\eta_{i}^{(k+1)}))$.

The EM algorithm iterates the E and M steps. At convergence the final estimates are denoted as $\hat{\beta}_0, \hat{\sigma}_0, \hat{\beta}_1, \hat{\sigma}_1, \hat{\beta}_2$ and $\hat{p}_i, i = 1, \ldots, n$. The EM algorithm has the property that, in the iteration, the observed log-likelihood is non decreasing, usually converges to a maximum, but needs a much larger number of iterations than other methods, for example, the Newton-Raphson method (Dempster et al. 1977). The standard errors are obtained from Louis (1982), as exemplified in Thompson et al. (1998). The best model is selected for the three sets of mean parameters using the Bayesian Information Criteria (Schwarz 1978), BIC.

### 3.3 Covariance matrix

The variance covariance matrix of the parameter estimates is obtained by inverting the observed information matrix,

$$I(\theta) = -\partial^2 l_{obs}/\partial \theta \partial \theta^T, \quad (4)$$

where $\theta = (\beta_0^T, \sigma_0, \beta_1, \sigma_1)$. Now, we put $I_{i}^{(c)}(\theta; y_i, Z_i) = (1 - Z_i) \ln(1 - \pi_i) + Z_i \ln(\pi_i) + (1 - Z_i) \ln f_0(y_i) + Z_i \ln f_1(y_i)$, which is the $i$-th component of the complete data log-likelihood. From Louis (1982), the observed information matrix is approximated by

$$I(\hat{\theta}) \cong \sum_i \hat{h}_i \hat{h}_i^T, \quad (5)$$

where $\hat{h}_i = \partial I_{i}^{(c)}(\hat{\theta}; y_i, \hat{p}_i)/\partial \theta$. The derivative $\hat{h}_i = (\hat{h}_{0i}, \hat{h}_{1i}, \hat{h}_{2i})^T$ can be written as,

$$\hat{h}_{0i} = \left[\begin{array}{c} (1 - \hat{p}_i)(y_i - \hat{\mu}_0) / \sigma_0^2 \\ (1 - \hat{p}_i)(y_i - \hat{\mu}_0)^2 / \sigma_0^4 \end{array}\right].$$

Jpn J Biomet Vol. 28, No. 2, 2007
The estimate of variance-covariance matrix for $\theta$ can be obtained by inverting the information matrix $I(\hat{\theta})$. The standard errors from the information matrix are compared with the bootstrap standard errors.

### 3.4 Population-based Cutoff Point and TgAb positive Rate

We show the population-based methods for determining the cutoff point and the positive rate. Let $V_i$ be the TgAb positive indicator for the $i$-th subject, where we define $V_i = 1$ if $y_i > c(i)$ and $V_i = 0$ otherwise with some cutoff point $c(i)$ of log TgAb level for the $i$-th subject.

To determine the cutoff point, we use three methods. The first one is simple. Since pop(0) is the TgAb negative population, the cutoff point is determined as

$$
c(i) = \tilde{\mu}_0 + 2\tilde{\sigma}_0,
$$

the upper bound of the normal range of the pop(0). With the AHS data, this method seems to be the only choice since pop(0) is truly TgAb negative but pop(1) can include both TgAb positive and negative subjects.

If population indicator $Z$ is a true TgAb positive indicator, Thompson et al. (1998) and Rosenthal et al. (1985) define the cutoff point by minimizing the misclassification rate (MMR), $Pr(V_i = 1, Z = 0) + Pr(V_i = 0, Z = 1) = (1 - \pi_i)\int_{c(i)}^{\infty} f_0(u)du + \pi_i\int_{-\infty}^{c(i)} f_1(u)du$. The cutoff point is determined as a larger solution of a quadratic equation, $(1 - \pi_i)f_0(c(i)) = \pi_i f_1(c(i))$, which minimizes the misclassification rate in terms of $c(i)$. The MMR cutoff depends on the mixing probability. McNeil et al. (1975) discussed methods for the selection of a cutoff point, one of which was the ROC curve method. If the population indicator $Z$ is a true TgAb positive indicator, the ROC curve method can also be applied to our data. The cutoff point is obtained by maximizing the sensitivity plus the specificity, $Pr(V_i = 0 | Z = 0) = Pr(V_i = 1 | Z = 1) = \int_{-\infty}^{c(i)} f_0(u)du + \int_{c(i)}^{\infty} f_1(u)du$. The cutoff point $c(i)$ is determined by solving the quadratic equation $f_0(c(i)) = f_1(c(i))$. The ROC curve cutoff does not depend on the mixing probability.

Note that, if $Pr(Z = 1) = 0.5$, the ROC curve and MMR methods are equivalent.

The logistic regression for $V_i$ can be used to estimate the positive rate for various methods. The standard errors are obtained by the bootstrap method for the cutoff point and prevalence.

### 4. Application

First we examined the log mixture normality of TgAb levels of 3185 atomic bomb survivors. Following the IFCC expert panel recommendation, we transformed the TgAb level as follows.
To avoid the log of 0, we added constant $c$ to the TgAb level. We assumed two cases, $c = 1$ and $c = 0.1$. The Box-Cox transformation (Box and Cox 1964) of the shifted TgAb level is defined as,

$$W^{(\lambda)} = \begin{cases} \ln(W) & \text{if } \lambda = 0 \\ \frac{W^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \end{cases},$$

(7)

where $W$ is the TgAb level plus $c$. We applied models with all covariates for the means in mixing probability and two normal components, and examined the relative goodness of fit in terms of BIC. We searched the best $\lambda$ by varying $\lambda$ from $-1$ to 1 by 0.01 increments. When $c = 1$, the BICs were 17532.9 for $\lambda = 0.00$ and 17443.7 for the best $\lambda = -0.14$. When $c = 0.1$, the BICs were 17260.8 for $\lambda = 0.00$ and 17254.6 for the best $\lambda = -0.02$. Since $c = 0.1$ gave a much better fit, we chose $c = 0.1$ and $\lambda = 0.00$, which indicated log-transformation for the value, TgAb level $+0.1$, or a modified logarithmic transformation of TgAb level.

Then we put $Y$ as the log of TgAb level plus 0.1 in base 10. Fig. 1 shows the $Y$ distribution for the subjects and the fitted density by the intercept models using the EM algorithm. In regression analysis, the explanatory variables were city (0 for Hiroshima, 1 for Nagasaki), sex (0 for male, 1 for female), age at examination coded as (age at examination $- 60$)/10, and radiation dose (Sv), $d$. The full model for linear predictors of the three mean models, the pop(0) population mean $\mu_0$, the pop(1) population mean $\mu_1$, and the mixing probability $\pi$, are two-factor interaction models of city, sex, age at examination, and radiation dose. For the mixture normal distribution, we selected the model from among hierarchical models that, if they included an interaction, also
Table 1. Log TgAb level model selection from the full model using BIC (n = 3185)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std Error*</th>
<th>Z-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population: pop(0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.546</td>
<td>0.013 (0.013)</td>
<td>40.69</td>
</tr>
<tr>
<td>City</td>
<td>−0.046</td>
<td>0.013 (0.012)</td>
<td>−3.472</td>
</tr>
<tr>
<td>Sex</td>
<td>−0.047</td>
<td>0.013 (0.013)</td>
<td>−3.662</td>
</tr>
<tr>
<td>(Age at Exam. – 60)/10</td>
<td>0.026</td>
<td>0.0078 (0.0074)</td>
<td>3.261</td>
</tr>
<tr>
<td>Sigma</td>
<td>0.226</td>
<td>0.0055 (0.0061)</td>
<td>41.18</td>
</tr>
<tr>
<td>Population: pop(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.160</td>
<td>0.080 (0.084)</td>
<td>14.43</td>
</tr>
<tr>
<td>City</td>
<td>−0.325</td>
<td>0.075 (0.073)</td>
<td>−4.304</td>
</tr>
<tr>
<td>Sex</td>
<td>0.362</td>
<td>0.082 (0.088)</td>
<td>4.434</td>
</tr>
<tr>
<td>Sigma</td>
<td>1.137</td>
<td>0.023 (0.026)</td>
<td>207.57</td>
</tr>
<tr>
<td>Mixing probability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.999</td>
<td>0.110 (0.107)</td>
<td>−9.061</td>
</tr>
<tr>
<td>City</td>
<td>0.366</td>
<td>0.104 (0.104)</td>
<td>3.467</td>
</tr>
<tr>
<td>Sex</td>
<td>0.496</td>
<td>0.112 (0.107)</td>
<td>4.448</td>
</tr>
</tbody>
</table>

*: The values in the parentheses are bootstrap standard errors with 1000 re-samplings.

included the main effects. We used GAUSS software for all computations.

We selected the best model by BIC model selection criteria from the full model for the two component densities and mixing probability (Table 1). In the pop(0) component, city, sex, and age at examination were included in the model although parameter estimates were small. In the pop(1) component, city and sex were included in the model. The mixing probability model included city and sex. The best mixing probability did not include age. The bootstrap standard errors were similar to the information matrix standard errors by Louis (1982), where the number of bootstrap replications was 1,000 for all bootstrap simulations. The mean of pop(1) is about twice the mean of pop(0) and the SD of pop(1) is about 5 times the SD of pop(0). The covariate effects are relatively small compared with the intercept in both pop(0) and pop(1). Following Thompson et al. (1998), we performed the goodness of fit test. For the best model, the number of parameters was 12 and the number of intervals was 63. The goodness of fit chi-square statistic was 57.60 with 50 degrees of freedom with \( P = 0.215 \), which implies that the distribution of the shifted TgAb level was a mixture of log-normal distributions with two components.

In the laboratory standards, we used 120 thyroid disease-free subjects to determine the cutoff point. In this method, covariates such as sex and age are not taken into account. The Akaike Information Criterion (AIC) (Akaike 1973, 1974) best linear logistic prevalence model using a laboratory cutoff point equal to 1.0 in \( Y \) is given in the formula \( \logit(\text{prevalence}) = −1.578(0.083) + 0.589(0.096) \) sex, where the values in the parentheses are standard errors. The estimates for positive rate using laboratory standards were 0.171 (95 % Wald CI, 0.149 to 0.195) for men and 0.271 (95 % Wald CI, 0.253 to 0.290) for women.

To compare our population-based method with the laboratory method, we used intercept
Application of Finite Mixture Models

Table 2. Log TgAb level normal mixture analysis with intercept models for two components using BIC (n = 3185)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std Error*</th>
<th>Z-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population: pop(0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.531</td>
<td>0.0062</td>
<td>85.13</td>
</tr>
<tr>
<td>Sigma</td>
<td>0.230</td>
<td>0.0056</td>
<td>40.96</td>
</tr>
<tr>
<td>Population: pop(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.294</td>
<td>0.039</td>
<td>33.17</td>
</tr>
<tr>
<td>Sigma</td>
<td>1.167</td>
<td>0.024</td>
<td>208.06</td>
</tr>
<tr>
<td>Mixing probability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.915</td>
<td>0.092</td>
<td>−9.908</td>
</tr>
<tr>
<td>Sex</td>
<td>0.554</td>
<td>0.107</td>
<td>5.184</td>
</tr>
</tbody>
</table>

*: The values in the parentheses are bootstrap standard errors with 1000 re-samplings.

models with no covariate for the two components means and selected a model for the linear logistic mixing probability model using BIC criteria. Since we wanted to search for the true model for the mixing probability but not to predict the prevalence using the mixing probability, we used BIC criteria not AIC criteria. Table 2 shows the results. The selected best linear logistic mixing probability model was 1 + sex, where 1 stands for intercept. The cutoff points using the first simple method, formula (6), was 0.991 with bootstrap S.E. = 0.0148 for both sexes, which resulted in cutoff points close to 1.0, the laboratory standard. The cutoff points in the ROC curve method were 0.951 (bootstrap SE = 0.012) for both sexes, and the cutoff points in the MMR method were 1.051 (bootstrap SE = 0.0203) for men and 0.993 (bootstrap SE = 0.0174) for women. The positive rate estimates for \( V_i \) using the best linear logistic regression models in the three methods (simple, ROC curve, and MMR) were similar since the cutoff points were approximately the same. Those values were 0.172 (95% bootstrap CI, 0.149 to 0.195) for men and 0.273 (95% bootstrap CI, 0.254 to 0.292) for women for the simple method, 0.185 (95% bootstrap CI, 0.159 to 0.211) for men and 0.284 (95% bootstrap CI, 0.265 to 0.304) for women for the ROC curve method, and 0.159 (95% bootstrap CI, 0.131 to 0.187) and 0.272 (95% bootstrap CI, 0.250 to 0.294) for the MMR method. The results are shown in Fig. 2.

5. Discussion

The mean of population pop(0) increases with age at examination (Table 1), though the parameter estimate is quite small and the population pop(1) and the mixing probability do not depend on age, which might be caused by the fact that the study population is an older population, i.e., age at examination greater than or equal to 55. If the variable \( Z \) is a TgAb positive indicator although this is not correct, our data suggest that the positive rates are 0.27 for Hiroshima men, 0.38 for Hiroshima women, 0.35 for Nagasaki men, and 0.47 for Nagasaki women.
Fig. 2. Rates of TgAb positivity and the 95% confidence bands by sex for four methods. The circle represents the logistic regression prevalence from laboratory standards with Wald CI, the square represents our simple method, the triangle represents the ROC curve method, and the upside-down triangle represents the minimizing misclassification rate (MMR) method. For the latter three methods, the positive rates were calculated by logistic regression for each cutoff point, and the 95% CIs were calculated by the bootstrap method.

We separated the TgAb level distribution into two populations, pop(0) and pop(1), with a mixing probability. Although the population pop(1), $Z = 1$, was not truly a TgAb positive population, the population pop(1) was suspected to be equal to the chronic thyroiditis population. This is merely a hypothesis because the true diagnostic criteria for chronic thyroiditis (Japan Thyroid Association) are 1st positive for TgAb, 2nd positive for anti-thyroid peroxidase antibody (TPO-Ab), or 3rd lymphocytic infiltration of thyroid gland confirmed with cytological examination. In our data (Imaizumi et al. 2006), 761 subjects were TgAb positive, 427 were TPO-Ab positive, and 290 subjects were positive for both. Thus 67.9% of the TPO-Ab positive subjects were also TgAb positive, and 38.1% of TgAb positive subjects were also TPO-Ab positive. This indicates that TgAb is a major marker for chronic thyroiditis, as stated in the introduction, and TgAb negative people can be TPO-Ab positive.

We showed that log-transformed TgAb level is two-component mixture normal, in which the smaller normal distribution corresponds to the TgAb negative group but the larger distribution does not necessarily correspond to the TgAb positive group. We compared several population-based methods with the laboratory method. Since the ROC curve and MMR methods are valid in determination of cutoff point only when TgAb positivity and negativity are known but the simple method (formula (6)) is valid only when TgAb negativity is known, the simple method was useful for determining the cutoff in our data. A laboratory reference value for TgAb positive was justified using the several population-based methods. Typically, diabetes mellitus is reflected merely by serum glucose level, but chronic thyroiditis is reflected by both TgAb level and TPO-
Application of Finite Mixture Models

Ab level. Therefore, the use of two-component mixture normal distribution in TgAb level and the simple, population-based method for cutoff would be another typical example of better handling of the clinical measurements than with the method given in Thompson et al. (1998).

Acknowledgements

We would like to thank the referee for the constructive comments. We thank Drs. Kazuo Neriishi in the Hiroshima clinical laboratory, and Kei Nakachi in the radiation biology and molecular epidemiology laboratory of RERF. RERF is a private, non-profit foundation funded by the Japanese Ministry of Health Labour and Welfare and the U.S. Department of Energy, with the latter funding provided through the National Academy of Sciences. This publication was supported by RERF research protocol #2-99.

REFERENCES


Japan Thyroid Association. Guideline for the diagnosis of chronic thyroditis (Hashimoto disease).
In *Guideline for the diagnosis of thyroid disease* (the 7th draft).
http://thyroid.umin.ac.jp/flame.html


