ASSESSMENT OF MISCLASSIFICATION IN A BINARY RESPONSE: RECOVERING INFORMATION ON CLINICALLY SIGNIFICANT CATARACT PREVALENCE FROM CATARACT SURGERY DATA IN ATOMIC-BOMB SURVIVORS

Eiji Nakashima*, Yoshinori Fujii**, Kazuo Neriishi*** and Atsushi Minamoto****

Cataract surgery results when a patient decides to undergo lens surgery following a diagnosis of a clinically significant cataract (CSC). Because the presence of a CSC is generally latent and unobserved, a person might not receive cataract surgery even if the person has a CSC. This misclassification needs to be adjusted in the statistical analysis of CSC so as to reduce the bias in the parameter estimation. Following Magder and Hughes (1997) and using the cataract surgery data on atomic-bomb survivors at the Radiation Effects Research Foundation, we used this method for estimating the prevalence of CSC in a linear logistic dose response model taking account of the sensitivity and/or specificity of the decision for lens surgery. The estimated sensitivity was 0.385 (95% CI: 0.268, 0.517) and the estimated specificity was perfect. The odds ratio estimate for the radiation dose response changed from 1.39 (95% CI: 1.24, 1.55) to 1.58 (95% CI: 1.26, 1.98) when allowing for the imperfect sensitivity. A large sample simulation study with a continuous covariate was conducted, assuming either imperfect sensitivity or imperfect specificity, to investigate the performance of the method. Results indicated that the parameter estimates are almost correct. We calculated the asymptotic relative efficiency (ARE) for a simple logistic regression slope estimate and showed that the ARE depends only on the values of slope and intercept parameters.

Key words and phrases: Atomic bomb, cataract, EM algorithm, misclassification, sensitivity, specificity.

1. Introduction

A cataract is an irreversible lens disorder caused by diabetes mellitus and other diseases, exposure to ultraviolet light or ionizing radiation, and lenticular injuries. From normal lenses, the opacity grade tends to increase with age, often with ambiguous lens opacity diagnosis observed and, in severe cases, blindness can result. When severe opacity causes vision impairment and the individual is...
greatly inconvenienced in his or her activities of daily life, lens cataract surgery is recommended. Such a clinically significant cataract and the patient’s decision to undergo lens surgery result in a cataract operation. Clinically significant cataracts are often latent and may not be observed unless a physician or an ophthalmologist is consulted. Thus, the estimation of the prevalence of clinically significant cataracts from diagnosed severe cataracts is important in epidemiology. The majority of clinically significant cataracts are cortical cataracts as shown, for example, in Nakashima et al. (2006).

The level of opacity at which a patient decides to undergo recommended surgery may be greater than that at which a physician diagnoses the need for surgery. The conditional probability that a patient decides on lens surgery given that the subject truly has clinically significant cataract symptoms is the “sensitivity”, and the conditional probability that a subject does not decide on cataract surgery, given that the subject has no clinically significant cataract symptoms, is called the “specificity”. Accordingly, the false negative and false positive rates, respectively, are one minus the sensitivity and one minus the specificity. For a situation where a decision is made to perform cataract surgery, the specificity may be assumed to be one, because the subject does not need a lens operation in the absence of any sign of a clinically significant cataract. Based on this consideration, our hypothesis is that, whereas the specificity of operated cataract is 1.0 (perfect specificity), the sensitivity is imperfect (less than 1.0). Under imperfect sensitivity and/or specificity, estimating the parameters of the disease prevalence and sensitivity and/or specificity is an important statistical problem.

Because the problem of estimating sensitivity, specificity, and the true response probability is a missing data problem (Magder and Hughes (1997)), the EM algorithm (Dempster et al. (1977)) may be applied. Buonaccorsi (2010) has discussed this type of missing data problem as a measurement error problem. Magder and Hughes (1997) have developed a method for estimating the regression parameters, sensitivity, and specificity using the EM algorithm. In their model, the parameters are identifiable under the condition that the sensitivity and specificity are constant, and the model includes a continuous covariate. Bandeen-Roche et al. (1997) and Chung et al. (2006) considered a more general latent class regression model and Bandeen-Roche et al. (1997) discussed the local identifiability of parameters in the context of latent class regression. Neuhaus (1999) considered bias when true sensitivity and specificity are less than one but are nevertheless ignored, and showed that the regression parameter is exaggerated when adjusting the sensitivity and/or specificity compared to the cases that the sensitivity and/or specificity are ignored.

Neriishi et al. (2007) analyzed the operated cataract data among atomic-bomb survivors using a logistic regression model and reported a significant radiation dose response. In the current report, using the method by Magder and Hughes (1997), we demonstrated how to estimate the prevalence of an unobserved clinically significant cataract taking into account the sensitivity and/or specificity with the operated cataract data, and we reanalyzed the data in terms
of the radiation dose response. In the application of the dose-response analysis, we test the hypothesis that the specificity is perfect and the sensitivity is imperfect in the decision of whether or not to perform cataract surgery in an atomic-bomb survivor.

In Section 2, we review the method by Magder and Hughes (1997) for estimating the prevalence of an unobserved clinically significant cataract taking into account the sensitivity and/or specificity in the decision of lens surgery, and derive a likelihood ratio test for imperfect sensitivity and/or specificity from Self and Liang (1987). In Section 3, we describe the data by Neriishi et al. (2007) on operated cataract data from the Adult Health Study cohort of the Radiation Effects Research Foundation, and analyze the data allowing for imperfect sensitivity and/or specificity in the decision for cataract surgery. In Section 4, we report a simulation study using a simple logistic regression when there is imperfect sensitivity or specificity. In Section 5, we make some comments regarding the validity of our results.

2. Statistical methods

2.1. Probabilities of clinically significant cataract and operated cataract

A clinically significant cataract (CSC), and the patient’s decision to undergo lens surgery, together constitute a case of operated cataract (OC). Let $T$ be a latent, unobserved indicator representing the CSC response of an individual depending on the individual’s covariates, and let $Y$ be an indicator variable representing the OC outcome resulting from $T$ with decision uncertainty (uncertainty in the decision processes leading up to surgery—a person with CSC might not seek medical help, or a physician might not recommend surgery in a case of CSC due to the patient’s health condition). We assume that a decision to undergo lens surgery is based solely on the extent of lens opacity and do not allow for variation in the decision-making process within or between doctors.

Assuming that the CSC response depends on covariate vector $x$ that includes a continuous covariate, and that the sensitivity and specificity are constant and do not depend on any covariates, for the OC response probability, $\Pr(Y = 1 \mid x) = \Pr(Y = 1, T = 1 \mid x) + \Pr(Y = 1, T = 0 \mid x)$, we obtain

$$
\Pr(Y = 1 \mid x) = \Pr(Y = 1 \mid T = 1) \Pr(T = 1 \mid x) + \{1 - \Pr(Y = 0 \mid T = 0)\} \Pr(T = 0 \mid x).
$$

We write the sensitivity $\Pr(Y = 1 \mid T = 1) = \rho$, the specificity $\Pr(Y = 0 \mid T = 0) = \tau$, and the CSC probability $\Pr(T = 1 \mid x) = \pi$, and we assume logistic models for all of these probabilities. The observed OC probability can then be expressed as $P = \Pr(Y = 1 \mid x) = (\rho + \tau - 1)\pi + (1 - \pi)$, for which we model $\logit(\pi) = x^T \beta$, $\logit(\rho) = \gamma$, and $\logit(\tau) = \delta$, where $\beta$, $\gamma$, and $\delta$ are the regression parameter vectors. We denote the collection of regression parameters as $\theta = (\beta^T, \gamma^T, \delta^T)^T$. 
2.2. Log-likelihood

We consider a model for $(T, Y)$ where $T$ is an unobserved latent CSC binary variable and $Y = y$ is an observed OC response. From the fact that $\Pr(T, Y = y) = \Pr(T) \Pr(Y = y \mid T)$, we have a joint probability,

\begin{equation}
\Pr(T, Y = y) = \Pr(T) \Pr(Y = y \mid T = 1)^T \Pr(Y = y \mid T = 0)^{1-T},
\end{equation}

where $\Pr(T)$ is a binary model for latent variable $T$ depending on covariates, and $\Pr(Y = y \mid T = t)$ is a generalized linear model for $t = 0, 1$. When $\Pr(Y = y \mid T = t)$ for $t = 0, 1$ are binary models not depending on covariates, model (2.2) results in the model given by Magder and Hughes (1997). In the present case we put $\Pr(Y = y \mid T = 1) = \rho y (1 - \rho) 1 - y$ and $\Pr(Y = y \mid T = 0) = (1 - \tau) y \tau 1 - y$, so the marginal model for the observed response $\Pr(Y = y) = \pi \rho y (1 - \rho) + (1 - \pi)(1 - \tau) y \tau 1 - y$ is a finite mixture binary model with $P = (\rho + \tau - 1)\pi + (1 - \tau)$. When the sensitivity and specificity are both known fixed constants, the model for $P$ is a generalized linear model under the condition $\rho + \tau > 1$ (Neuhaus (1999)). This can be seen from the fact that $\rho + \tau > 1$ indicates $P$ is monotone in $\pi$, in which case $P$ is also monotone in the linear predictor of $\pi$.

When we observe OC data $\{y_i\}$ of sample size $n$ with unknown sensitivity $\rho$ and unknown specificity $\tau$, the observed data log-likelihood can be written as a binary log-likelihood,

\begin{equation}
l_{obs} = \sum_{i=1}^{n} \{y_i \ln(P_i) + (1 - y_i) \ln(1 - P_i)\},
\end{equation}

where $P_i = (\rho + \tau - 1)\pi_i + (1 - \tau)$. Because it is difficult to fit this incomplete observed likelihood, we instead use formula (2.2) including the missing variable $T$. This is a missing data problem, so we can employ the EM algorithm (Magder and Hughes (1997), Bandeen-Roche et al. (1997)). By replacing three probability components in formula (2.2) with corresponding binary formula, respectively, the complete data log-likelihood can be written as

\begin{equation}
l_c = \sum_{i=1}^{n} \{T_i \ln(\pi_i) + (1 - T_i) \ln(1 - \pi_i)\}
+ \sum_{i=1}^{n} T_i \{y_i \ln(\rho) + (1 - y_i) \ln(1 - \rho)\}
+ \sum_{i=1}^{n} (1 - T_i) \{(1 - y_i) \ln(\tau) + y_i \ln(1 - \tau)\}.
\end{equation}

If the true response probability $\pi_i$, sensitivity $\rho$, and specificity $\tau$ in the above summations are parameterized separately, maximizing $l_c$ amounts to maximizing the three summations separately, which are binary log-likelihoods with the latter two having weights $T_i$ and $1 - T_i$, respectively. The parameters are identifiable.
from the complete data log-likelihood (2.4) when the sensitivity and specificity are constant and the covariate vector $x$ includes a continuous covariate (Magder and Hughes (1997)).

### 2.3. Estimation using the EM-algorithm

Dempster et al. (1977) showed that the incomplete data log-likelihood can be maximized by iteratively maximizing the conditional expectation of the complete data log-likelihood given the incomplete data. In our case, because the complete data log-likelihood is linear in unobserved latent response variables, the expectation is easily calculated. From the Bayes theorem or Magder and Hughes (1997), the conditional expectation 

$$E(T_i | Y_i = y_i) = \frac{\pi_i \Pr(Y_i = y_i | T_i = 1)}{\pi_i \Pr(Y_i = y_i | T_i = 1) + (1 - \pi_i) \Pr(Y_i = y_i | T_i = 0)}.$$  

(2.5)

When the specificity is perfect ($\tau = 1$), we have $E(T_i | Y_i = 1) = 1$ and $E(T_i | Y_i = 0) = \pi_i/(1 - \pi_i \rho)$. When sensitivity is perfect ($\rho = 1$), we have $E(T_i | Y_i = 1) = \pi_i/(1 - \tau + \pi_i \tau)$ and $E(T_i | Y_i = 0) = 0$. When the sensitivity and specificity are both equal to one, $E(T_i | Y_i) = Y_i$.

Suppose the initial estimate of sensitivity, specificity and response probability are $\rho^{(0)}$, $\tau^{(0)}$ and $\pi^{(0)}$, respectively, with $\rho^{(0)} + \tau^{(0)} > 1$. The $(k + 1)$th step estimates are calculated from the $k$-step estimates in the following two steps.

**Expectation step (E step):** Calculate the estimated conditional true response probability given the observed data $y_i$, from formula (2.5),

$$p^{(k+1)}_i = \left\{ \begin{array}{ll} \frac{\pi_i^{(k)} \rho^{(k)}}{\pi_i^{(k)} \rho^{(k)} + (1 - \pi_i^{(k)}) (1 - \tau^{(k)})} & \text{if } y_i = 1 \\ \frac{\pi_i^{(k)} (1 - \rho^{(k)})}{\pi_i^{(k)} (1 - \rho^{(k)}) + (1 - \pi_i^{(k)}) \tau^{(k)}} & \text{if } y_i = 0 \end{array} \right.$$

(2.6)

**Maximization step (M step):** This step consists of three separate logistic regressions,

1. Estimate the parameter $\beta^{(k+1)}$ using a logistic regression for the conditional expectation $p_i^{(k)}$ on covariate $x$ and calculate $\pi_i^{(k+1)} = \text{logit}^{-1}(x_i^T \beta^{(k+1)})$.

2. Estimate the sensitivity parameter $\gamma^{(k+1)}$ using a logistic regression for the data $y_i$ with weight $p_i^{(k)}$ and calculate the sensitivity by $\rho^{(k+1)} = \text{logit}^{-1}(\gamma^{(k+1)})$. If the sensitivity is set to one, this part is skipped.

3. Estimate the specificity parameter $\delta^{(k+1)}$ using a logistic regression for the data $1 - y_i$ with weight $1 - p_i^{(k)}$, and calculate the specificity by $\tau^{(k+1)} = \text{logit}^{-1}(\delta^{(k+1)})$. If the specificity is set to one, this part is skipped.

The E step and M step are iterated until convergence of the observed log-likelihood. Convergence is guaranteed because the observed log likelihood in the EM iteration always increases. We define our convergence criterion to be that the relative difference of the successive observed log-likelihoods is less than
10^{-8}. We denote the parameter estimates at convergence as $\hat{\theta} = (\beta^T, \gamma^T, \delta^T)^T$, the estimated true probability, sensitivity, and specificity as $\hat{\pi}_i = \logit^{-1}(x_i^T \hat{\beta})$, $\hat{\rho} = \logit^{-1}(\hat{\gamma})$, and $\hat{\tau} = \logit^{-1}(\hat{\delta})$, respectively, and the estimated conditional true response as $\hat{p}_i$.

2.4. Covariance matrix

The variance-covariance matrix of the parameter estimates $\hat{\theta}$ is obtained as the inverse of the observed information matrix,

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

where $\theta = (\beta^T, \gamma^T, \delta^T)^T$. We put $l_c(i)(\theta; y_i, T_i) = T_i \ln(\pi_i) + (1 - T_i) \ln(1 - \pi_i) + T_i \{y_i \ln(\rho) + (1 - y_i) \ln(1 - \rho)\} + (1 - T_i) \{(1 - y_i) \ln(\tau) + y_i \ln(1 - \tau)\}$, which is the $i$th component of the complete data log-likelihood for $i = 1, \ldots, n$. From Louis (1982), the observed information matrix reduces to

$$I(\hat{\theta}) \cong \sum_{i=1}^{n} \hat{h}_i \hat{h}_i^T,$$

where $\hat{h}_i = \partial l_c(i)(\hat{\theta}; y_i, \hat{p}_i) / \partial \theta$. The derivative $\hat{h}_i = (\hat{h}_{1i}, \hat{h}_{2i}, \hat{h}_{3i})^T$ can be written as $\hat{h}_{1i} = (\hat{p}_i - \hat{\pi}_i)x_i$, $h_{2i} = \hat{p}_i(y_i - \hat{\rho})$, and $\hat{h}_{3i} = (1 - \hat{p}_i)(1 - y_i - \hat{\tau})$.

2.5. Likelihood ratio test for imperfect sensitivity and/or specificity

Because the null hypothesis $\rho = \tau = 1$ is on the boundary of the parameter space, we employ the likelihood ratio (LR) test under nonstandard conditions as described by Self and Liang (1987). Our situation corresponds to their Case 7. The $p$ value from the LR test for the null hypothesis of perfect sensitivity and specificity is

$$\text{Pr}(\hat{\chi}^2_0 \geq LR) + 0.5 \text{Pr}(\hat{\chi}^2_1 \geq LR) + c \text{Pr}(\hat{\chi}^2_2 \geq LR),$$

where $\chi^2_0$, $\chi^2_1$, and $\chi^2_2$ are independent chi-square variables with degrees of freedom 0, 1, and 2, respectively, and $c = \cos^{-1}\{J_{12}/\sqrt{J_{11}J_{22}}\} / (2\pi)$ with $J_{lm}$’s being $(l, m)$ entries of the information matrix of observed likelihood $J(\rho, \tau) = E(-\partial^2 l_{\text{obs}} / \partial \theta_1 \partial \theta_1^T)$ under the null hypothesis of $\rho = \tau = 1$ with $\theta_1 = (\rho, \tau)^T$. The entries are estimated as $\hat{J}_{11} = \sum_{i=1}^{n} \hat{\pi}_i / (1 - \hat{\pi}_i)$, $\hat{J}_{12} = -n$ and $\hat{J}_{22} = \sum_{i=1}^{n} (1 - \hat{\pi}_i) / \hat{\pi}_i$, where $n$ is the sample size and $\hat{\pi}_i$’s are the fitted values from the logistic regression under $\rho = \tau = 1$.

3. Application

We applied the present method to cataract surgery data of atomic-bomb survivors. Four models taking account of sensitivity and/or specificity in the decision of lens surgery are considered and applied to the data.
3.1. Data

At the Radiation Effects Research Foundation (RERF, formerly the Atomic Bomb Casualty Commission, ABCC), Hiroshima and Nagasaki, Japan, biennial health examinations of atomic-bomb survivors have been conducted since July 1, 1958. Originally, about 20,000 subjects, which constitute RERF’s Adult Health Study (AHS) cohort, participated in these series of examinations (Beebe et al. (1960), Yokoro (1991)). Radiation dose estimates have been estimated for most of the subjects, using the atomic-bomb radiation dosimetry system 2002, or DS02 (Young and Bennett (2006)). After truncation at 4 Gray (Gy), gamma and neutron DS02 radiation doses were adjusted for log-normal dosimetric measurement error with a 35% coefficient of variation (Pierce et al. (1990)). For use in statistical analysis, the radiation dose to the eye (eye dose) in Gy was calculated as the gamma eye dose plus ten times the neutron eye dose after this measurement error adjustment, assuming a neutron relative biological effectiveness of 10. This adjusted eye dose was used in regression analysis and the adjustment reduces radiation effect estimation bias due to radiation dosimetric error.

Using this cohort, an operated (or surgical) cataract prevalence study was conducted in the period between June 2000 and September 2002, about 55 yrs after the atomic bombings. During this period, excluding members with unknown exposed radiation doses and members that were exposed in utero, a total of 3,761 subjects were interviewed regarding whether or not they had undergone lens surgery (Neriishi et al. (2007)) and whether they had had lens opacity diagnoses. A total of 479 subjects were found to have had operated cataracts and 1,882 subjects were recorded as having a lens opacity diagnoses. Overall, 12.7% of the subjects had an operated cataract, and 50.0% had a lens opacity diagnosis, of which 25.5% had had an operated cataract. The radiation dose and age at radiation exposure are continuous covariates. The mean radiation dose was 0.494 Gy (range: 0–5.14 Gy). The mean age at radiation exposure was 16.6 yrs (range: 0–46 yrs) and the mean age at examination was 72.4 yrs (range: 55—101 yrs), resulting in a mean difference between examination and exposure of 55.8 yrs.

3.2. Model

With constant sensitivity and/or specificity, we assume a linear logistic model for $\pi$, the prevalence of CSC, in terms of covariates, city, sex, age at exposure to radiation, diabetes mellitus indicator and radiation dose:

$$\logit(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 City + \beta_2 Sex + \beta_3 Age \text{EXP} + \beta_4 DM + \beta_5 Dose,$$

where the $City$ and $Sex$ variables are Nagasaki and female indicators, respectively, $Age \text{EXP} = (\text{age at exposure (yrs)} - 10)/10$, $DM$ is the diabetes mellitus indicator, $Dose$ is the eye dose of radiation (Gy) and the $\beta$’s are the regression parameters.

Four models were assumed for the OC response probability (prevalence), $P$. The first is the naïve logistic model assuming $\rho = \tau = 1$, i.e., $P = \pi$, with a logistic model for $P$. This logistic model, although not realistic, can be fit
using a usual logistic regression program. The second is a sensitivity model that
estimates $\pi$ and the sensitivity $\rho$ assuming the specificity is set to one ($\tau = 1$).
The third is a specificity model that estimates $\pi$ and specificity $\tau$ assuming the
sensitivity is set to one ($\rho = 1$). The fourth model is an un-restricted model that
estimates $\pi$, the sensitivity $\rho$, and the specificity $\tau$. We set initial values for the
EM algorithm to $\rho^{(0)} = 0.3$ for the sensitivity, $\tau^{(0)} = 0.8$ for the specificity and
$\pi_i^{(0)} = 0.5$ for the true response probability.

We compared the four models using Akaike’s Information Criterion (AIC) (Akaike (1973, 1974)) and the Baysian Information Criterion (BIC) (Schwarz (1978)). AIC is the deviance of the fit plus two times the number of parameters
in the model; BIC is the deviance of the fit plus the logarithm of the number of
subjects times the number of parameters in the model, where the deviances are
calculated from the observed data log-likelihood converged in the EM iteration.
The smaller the AIC value the better is the fit for purpose of prediction and the
smaller the BIC the closer is the model to the true model. All computations were
made using GAUSS software (1998). Tests and confidence intervals were based
on the Wald statistic unless otherwise specified.

3.3. Results

First, we compared the four models: the naïve logistic regression, the sen-
itivity, the specificity, and the un-restricted models. Assuming sensitivity and
specificity to be constant and a model with main effects only for the true response
probability, the deviances, AIC and BIC values for the four models were shown
in Table 1. BIC indicates little difference between the naïve logistic model and
the sensitivity model. The sensitivity model is the best model in terms of both
AIC and BIC criteria. However, in general, for predicting disease prevalence, the
AIC criterion should be used. With the specificity and un-restricted models, the
parameter estimate for the specificity was close to 1.0, i.e., specificity was near
perfect. The difference in deviance between the sensitivity and the un-restricted
models is 0.14, with one degree of freedom, which is small enough to treat the
specificity as one. The LR test statistic under nonstandard conditions (Self and
Liang (1987)) derived from the deviances above for the null hypothesis of perfect
sensitivity and specificity was 8.44, and, from formula (2.9), it was statistically
significant ($p = 0.008$), indicating imperfect sensitivity and/or specificity.

Table 2 shows the parameter estimates of the naïve logistic regression using
the main effects model, with the assumption that sensitivity and specificity are
both one. The radiation dose effect adjusted for city, sex, age at exposure and
diabetes mellitus was 0.328 (odds ratio (OR) = 1.39; 95% CI: 1.24, 1.55) and
statistically significant ($p < 0.001$). Table 3 shows the parameter estimates of
the true response probability using the main effects model and the constant sen-
itivity estimated using the sensitivity model. The radiation dose effect adjusted
for city, sex, age at exposure and diabetes mellitus was 0.457 (OR = 1.58; 95%
CI: 1.26, 1.98) and statistically significant ($p < 0.001$). All parameter estimates
in Table 2 from the naïve logistic regression are attenuated compared with the
estimates from the sensitivity model in Table 3. The sensitivity parameter esti-
Table 1. Deviance, AIC and BIC values of four models for the cataract surgery data ($n = 3761$).

<table>
<thead>
<tr>
<th>Model</th>
<th>Deviance</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve model</td>
<td>2638.33</td>
<td>2650.33</td>
<td>2687.72</td>
</tr>
<tr>
<td>Sensitivity model</td>
<td>2630.03</td>
<td>2644.03</td>
<td>2687.66</td>
</tr>
<tr>
<td>Specificity model</td>
<td>2638.33</td>
<td>2652.33</td>
<td>2695.96</td>
</tr>
<tr>
<td>Un-restricted model</td>
<td>2629.89</td>
<td>2645.89</td>
<td>2695.75</td>
</tr>
</tbody>
</table>

Table 2. Results of naïve logistic regression analysis (3761 subjects, 479 cases).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. error</th>
<th>Odds ratio</th>
<th>2 sided p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−2.980</td>
<td>0.134</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>City (Naga/Hiro)</td>
<td>−0.0579</td>
<td>0.107</td>
<td>0.94</td>
<td>0.586</td>
</tr>
<tr>
<td>Sex (Female/Male)</td>
<td>0.126</td>
<td>0.116</td>
<td>1.13</td>
<td>0.281</td>
</tr>
<tr>
<td>AgeEXP</td>
<td>0.815</td>
<td>0.062</td>
<td>2.26</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DM</td>
<td>0.558</td>
<td>0.128</td>
<td>1.75</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dose</td>
<td>0.328</td>
<td>0.057</td>
<td>1.39</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AgeEXP = (age at exposure − 10)/10; Deviance = 2638.33; AIC = 2650.33.

Table 3. Logistic regression analysis using the sensitivity model (3761 subjects, 479 cases).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. error*</th>
<th>Odds ratio</th>
<th>2 sided p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−2.190</td>
<td>0.227 (0.185)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>City (Naga/Hiro)</td>
<td>−0.138</td>
<td>0.156 (0.154)</td>
<td>0.87</td>
<td>0.377</td>
</tr>
<tr>
<td>Sex (Female/Male)</td>
<td>0.229</td>
<td>0.167 (0.165)</td>
<td>1.26</td>
<td>0.171</td>
</tr>
<tr>
<td>AgeEXP</td>
<td>1.298</td>
<td>0.187 (0.117)</td>
<td>3.66</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DM</td>
<td>0.747</td>
<td>0.227 (0.208)</td>
<td>2.11</td>
<td>0.001</td>
</tr>
<tr>
<td>Dose</td>
<td>0.457</td>
<td>0.116 (0.100)</td>
<td>1.58</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AgeEXP = (age at exposure − 10)/10; Deviance = 2630.03; AIC = 2644.03; Sensitivity estimate = 0.385 (95% Wald CI: 0.268, 0.517); *: The values in parentheses are the S.E. of parameter estimates with sensitivity fixed at the maximum likelihood estimate 0.385.

Of the subjects who needed lens surgery, 61% did not actually undergo the operation under the current mathematical modeling assumptions. In our cohort of 3761 subjects with a mean age at examination of 72 yrs, the estimated number of clinically significant cataracts is 1244.

4. Simulation study

In order to ascertain the performance of the current method, we conducted several simulation studies. When both sensitivity and specificity are imperfect, a few of the simulations failed because the regression parameter estimates became huge. Therefore, only two settings were considered: sensitivity is perfect fixed and imperfect specificity and regression parameters were estimated, and specificity is perfect fixed and imperfect sensitivity and regression parameters were
Table 4. Simulation results with 1000 replications.

<table>
<thead>
<tr>
<th>Assumed sensitivity or specificity</th>
<th>Estimated regression parameters under fixed sensitivity or specificity</th>
<th>Estimated sensitivity, specificity and regression parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean $\hat{\beta}_0$ (SD)</td>
<td>Mean $\hat{\beta}_1$ (SD)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$\tau$</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>---</td>
<td>0.3</td>
<td>$-1.03$ (0.26)</td>
</tr>
<tr>
<td>---</td>
<td>0.4</td>
<td>$-1.01$ (0.21)</td>
</tr>
<tr>
<td>---</td>
<td>0.5</td>
<td>$-1.02$ (0.18)</td>
</tr>
<tr>
<td>---</td>
<td>0.6</td>
<td>$-1.01$ (0.15)</td>
</tr>
<tr>
<td>---</td>
<td>0.7</td>
<td>$-1.00$ (0.13)</td>
</tr>
<tr>
<td>---</td>
<td>0.8</td>
<td>$-1.01$ (0.11)</td>
</tr>
<tr>
<td>---</td>
<td>0.9</td>
<td>$-1.01$ (0.09)</td>
</tr>
<tr>
<td>---a)</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>0.9</td>
<td>---</td>
<td>$-1.00$ (0.08)</td>
</tr>
<tr>
<td>0.8</td>
<td>---</td>
<td>$-1.01$ (0.08)</td>
</tr>
<tr>
<td>0.7</td>
<td>---</td>
<td>$-1.00$ (0.09)</td>
</tr>
<tr>
<td>0.6</td>
<td>---</td>
<td>$-1.00$ (0.09)</td>
</tr>
<tr>
<td>0.5</td>
<td>---</td>
<td>$-1.00$ (0.11)</td>
</tr>
<tr>
<td>0.4</td>
<td>---</td>
<td>$-1.00$ (0.12)</td>
</tr>
<tr>
<td>0.3</td>
<td>---</td>
<td>$-1.00$ (0.14)</td>
</tr>
</tbody>
</table>

---: fixed perfect sensitivity $\rho$ or specificity $\tau$; mean and standard deviation (SD) of the estimates are calculated from 1000 replications for 2000 simulated data; The response probability is assumed as a simple logistic regression, $\text{logit}(\pi_i) = \beta_0 + \beta_1 x_i$ with $\beta_0 = -1$, $\beta_1 = 1$ and $x_i = (i/2000) - 3$ for $i = 1, \ldots, 2000$. Row a): $\rho = \tau = 1$ fixed indicates the results with ordinary logistic regression.
Figure 1. Asymptotic relative efficiency (ARE) calculated from the observed log-likelihood for the slope estimate of a simple logistic regression under a setting similar to that of Table 3. The response probability is assumed as a simple logistic regression with logit(\(\pi_i\)) = \(\beta_0 + \beta_1 x_i\), where \(x_i = 6(i/2000) - 3\) for \(i = 1, \ldots, 2000\), \(\beta_0 = -1.0\) (upper panels) or \(-2.0\) (lower panels) and \(\beta_1 = 0.5, 1.0,\) or \(2.0\). Known perfect sensitivity (right panels) or known perfect specificity (left panels) were assumed. The sensitivity (rho) and specificity (tau) range from 0.1 to 1.0 by increments of 0.1. The small circles around the lines of \((\beta_0, \beta_1) = (-1.0, 1.0)\) in the upper panels are calculated from Table 3.

The sensitivity \(\rho\) or specificity \(\tau\) ranged from 0.3 to 0.9 in increase of 0.1. The response probability was assumed to be logit(\(\pi_i\)) = \(\beta_0 + \beta_1 x_i\) with \(\beta_0 = -1,\) \(\beta_1 = 1\) and \(x_i = 6(i/2000) - 3\) for \(i = 1, \ldots, 2000\). A total of 2000 binary responses were generated having value one for \(U_i \leq \rho \pi_i + (1 - \tau)(1 - \pi_i)\) and zero for \(U_i > \rho \pi_i + (1 - \tau)(1 - \pi_i)\), where the \(U_i, i = 1, \ldots, 2000\), are independent, uniformly distributed random variables on the interval \((0, 1)\). One thousand independent simulations were conducted. The mean and standard deviation of the maximum likelihood estimates (MLEs) of sensitivity, specificity, and regression parameters were calculated.
Columns 3 and 4 of Table 4 show the MLEs for regression parameters under true, fixed sensitivity or specificity. The regression parameter estimates are almost correct, the biases of which are less than 5%. Columns 5 through 8 show MLEs for sensitivity or specificity and the regression parameters. The biases of the sensitivity or specificity estimates are mostly less than 5% and the biases of the regression parameter estimates are all less than 5%. The sensitivity, specificity and regression parameter estimates are quite good although in almost all cases the sensitivity or specificity are slightly overestimated and the regression parameter estimates are slightly exaggerated. Note that the true response probability can change from $1 - \hat{\tau}$ to $\hat{\rho}$. If $\hat{\rho}$ and $\hat{\tau}$ were overestimated, then the slope parameter would be exaggerated.

The SD’s of the regression parameter estimates are larger, when either the sensitivity or specificity is estimated, than those in the case where both sensitivity and specificity are fixed. The SD’s become larger when the true value of sensitivity or specificity becomes smaller. The SD of the sensitivity estimate is always larger than the SD’s of the specificity estimate in our simulation setting.

Figure 1 shows the asymptotic relative efficiency (ARE) for the present method assuming a simple logistic regression with either unknown sensitivity and known perfect specificity (left panel), or unknown specificity and known perfect sensitivity (right panel). The covariate is set to the same as in Table 4. The ARE was calculated from the observed log-likelihood by dividing the asymptotic variance of the slope estimate with known sensitivity (specificity) by that with unknown sensitivity (specificity). Two values were considered for the intercept parameter: $-2$ (upper panel) and $-1$ (lower panel). For values of the slope parameter equal to 0.5, 1.0 and 2.0, the AREs were approximately 10%, 30% to 40%, and 60% to 70%, respectively, irrelevant of the values of unknown sensitivity when specificity is known and perfect (left panel), and approximately 15%, 50% and 80%, respectively, irrelevant of the values of unknown specificity when sensitivity is known and perfect (right panel). The points in the upper panels around the lines for $(\beta_0, \beta_1) = (-1.0, 1.0)$ are the relative efficiencies calculated from Table 4; these closely approximate the AREs.

5. Discussion

Bandeen-Roche et al. (1997) stated the condition of local identifiability of the parameter and recommended to try a variety of starting values to assure that the estimate is the global maximizer of the observed log-likelihood (2.3), $l_{\text{obs}}$. In our simpler model, the parameters are identifiable because the model includes continuous covariates as discussed by Magder and Hughes (1997). Thus, the maximum likelihood estimate of the parameter vector is the global maximizer of the log-likelihood (2.3). Nevertheless, using our AHS data, we tried several starting values for an un-restricted model allowing for both sensitivity and specificity, and assured that our estimate of $\theta$ was indeed the global maximizer of the log-likelihood (2.3).

Whether or not the sensitivity and specificity are constant and do not depend
on covariates is important in our application. In our epidemiologic survey, random and blind decisions for cataract surgery conditional on CSC would guarantee constant sensitivity and specificity. However, in practice, it would be difficult to implement the randomness and blindness because physicians would recommend lens surgery upon taking into account the health condition of the patient. In the current analysis, if the sensitivity depends on covariates, the model would not be identifiable because the covariate effects can be incorporated into either sensitivity or true probability, since the observed probability is sensitivity times true probability, i.e., \( P = \rho \pi \). Thus the model is identifiable only with constant sensitivity. Furthermore, as described in Neuhaus (1999), if the sensitivity and specificity depend on covariates, the naïve logistic model parameters have bias away from the null. These were the rationales for assuming the constant sensitivity and specificity in the application.

McCarty et al. (2000) reported in an Australian study that certain demographic factors (such as age, gender, rural residence, occupation, employment status, health insurance status, and ethnicity) were unrelated to a decision to forgo cataract surgery. In other words, sensitivity does not depend on any covariates and can be assumed to be constant. In the analyses of atomic-bomb survivor data, however, radiation dose is a necessary covariate and related to the distance from the hypocenters of the atomic-bombings. If sensitivity is related to radiation dose, the dose response in the true response probability could become statistically non-significant. However, compared with non-operated severe cataracts (Nakashima et al. (2006)), the percentage of severe cataracts that had been surgically removed did not vary significantly with radiation dose (Neriishi et al. (2007)). Therefore, in the present application, we deemed it reasonable to describe the data using a model with constant sensitivity.

In development of lens opacity from diagnosed to operated cataract, for each subject we have \( Y = 1 \Rightarrow T = 1 \Rightarrow W = 1 \), where \( Y \) is an indicator of observed operated cataract (OC), which is a good surrogate for clinically significant cataract (CSC), \( T \) is an unobserved indicator of CSC, and \( W \) is an indicator of observed lens opacity diagnosis (LOD) that is often ambiguous and may even include false-positive non-opacity. Although we can consider LOD as a surrogate of CSC, clinically, OC is a better surrogate of CSC than LOD and, because Neriishi et al. (2007) treated OC as a surrogate of CSC, we focused on the OC as a surrogate of CSC. From the above relationship, one can easily show that a lower bound on the sensitivity is given by \( \rho = \Pr(Y = 1 \mid T = 1) \geq \Pr(Y = 1)/\Pr(W = 1) \). The right hand side of the inequality is 0.255 for our data. This lower bound on sensitivity does not contradict the result from our sensitivity analysis. Since humans have two eyes, patients do not necessarily undergo cataract surgery when only one of the two eyes has a CSC and loses visual acuity. Therefore, the reported sensitivity of 0.385, which expresses the proportion of people actually undergoing surgery among those who should be operated on, is reasonable.

One report from Finland (Hirvela et al. (1995)) showed that, when both sexes
were combined, the proportion of subjects above the age of 70 yrs, who required cataract surgery, had no lens, or had an artificial lens, was 30.3%, which is close to our estimate of the prevalence of CSC. Using the estimates from our analysis, the true prevalence $\pi$ is about 34.7% and observed prevalence $P$ is 11.9% for Hiroshima, males, with doses estimated to be 0 Gy, and no diabetes mellitus, who were 15 yrs or older at the time of the bombings (age at examination above 70 yrs and mean age at the time of bombing 22.0 yrs). The difference in CSC estimates between ours and that in Finland can be explained by the fact that Japan is located at a lower latitude (Sasaki et al. (2003)). In a Swedish study, Carlsson et al. (1996) reported that, above that age of 70 years, females tended to undergo surgery at a significantly higher rate. Our analysis suggests that female atomic-bomb survivors have a higher rate of surgery, but the gender difference is not statistically significant; this might reflect the fact that the overall examination rate is higher in the atomic-bomb survivors than in the general population.

Finally, the odds ratio estimates at 1 Gy of radiation dose obtained earlier by Neriishi et al. (2007) and our estimate here from the sensitivity model do not differ drastically, which indicates the model ignoring sensitivity and specificity, i.e., naïve model, gives roughly the correct estimate of the radiation dose effect. However, we have additionally shown that the sensitivity is imperfect and specificity is perfect in the decision of cataract surgery, which is an important finding in ophthalmic epidemiology.

Acknowledgements

We would like to thank the anonymous referees for their helpful comments and also Drs. John Cologne and Ravindra Khattree and Mr. Jeffrey Hart of RERF for their editorial advice. RERF, Hiroshima and Nagasaki, Japan, is a private, non-profit foundation funded by the Japanese Ministry of Health, Labour and Welfare (MHLW) and the U.S. Department of Energy (DOE), the latter in part through the National Academy of Sciences. This publication was supported by RERF Research Protocol RP 2-75.

REFERENCES


