Author's response to reviews

Title: Practical application of cure mixture model for long-term censored survivor data from a withdrawal clinical trial of patients with major depressive disorder

Authors:
Ichiro Arano (Ichiro.Arano@pfizer.com)
Tomoyuki Sugimoto (sugimoto@medstat.med.osaka-u.ac.jp)
Toshimitsu Hamasaki (hamasakt@medstat.med.osaka-u.ac.jp)
Yuko Ohno (ohno@sahs.med.osaka-u.ac.jp)

Version: 2 Date: 16 March 2010

Author's response to reviews: see over
Dear Editor, BMC Medical Research Methodology,

Please find the revised manuscript (MS: 3999481134769886), entitled “Practical application of cure mixture model to long-term censored survivor data from a withdrawal clinical trial of patients with major depressive disorder” (the title was revised, reflecting the contents of the manuscript), along with our comments to the reviewers.

We found the comments of you and the reviewers to be very useful for improving our manuscript. We tried to address the remaining points in the revised version of our manuscript. We have made all possible changes in compliance raised by the reviewers in the revised manuscript. Please find our detailed answers to the reviewers’ comments following.

The changes and responses we have made would make our revised manuscript to be acceptable for publication in BMC Medical Research Methodology.

We would thank you again for the consideration of our manuscript for publication in BMC Medical Research Methodology

Yours sincerely,

Toshimitsu Hamasaki, PhD
Department of Biomedical Statistics
Osaka University Graduate School of Medicine
2-2 Yamada-oka
Suita, Osaka 565-0871, Japan
Tel: +81-6-6879-3597  Fax: +81-6-6879-3598
E-mail: hamasakt@medstat.med.osaka-u.ac.jp
**Editorial Requests**

<table>
<thead>
<tr>
<th>Request #1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title?</strong> Please revise the title to make sure that it accurately reflects the contents of the manuscript, in particular, we feel that the statistical focus of the work should be mentioned.</td>
</tr>
<tr>
<td><strong>Response:</strong></td>
</tr>
<tr>
<td>We revised the title to reflect the contents of the manuscript as follows:</td>
</tr>
<tr>
<td>“Practical application of cure mixture model for long-term censored survivor data from a withdrawal clinical trial of patients with major depressive disorder”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Request #2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Competing interests?</strong> We note that Dr Arano is affiliated at Pfizer, please add a note to reflect this under the declaration of Competing interests.</td>
</tr>
<tr>
<td><strong>Response:</strong></td>
</tr>
<tr>
<td>We added a note to reflect our situation under the declaration of competing interests as follows:</td>
</tr>
<tr>
<td>“Tomoyuki Sugimoto and Yuko Ohno declare no competing interests. Ichiro Arano is an employee of Pfizer Japan, and Toshimitsu Hamasaki was an employee of Pfizer Japan between 1997 and 2004.”</td>
</tr>
</tbody>
</table>
Dear Referee #1,

Thank you very much for your constructive comments and valuable suggestions with regard to our manuscript. Specific responses to each of your comment are provided as follows.

**Comment #1**
For the use of cure model, it is usually required sufficiently long follow-up. For the sertraline randomized withdrawal study considered in this paper, the total follow-up is only about 120 days. In addition, there is a very late relapse time (nearly at 110 days) in the placebo group, and beyond this time point, there are only a few censored times. So my questions are: a.) is there any biological support for the “cure” in the study? b.) is the follow-up sufficiently long for identifying the cure? Some statistical tests may be conducted for checking whether the follow-up is sufficiently long (for example, Maller and Zhou, 1996).

**Response:**

The sertraline study discussed in the manuscript was to contribute the submission of the new drug application in Japan, so that the length of follow-up period (16 weeks) of the study was considered to be minimized at least to detect the drug effect compared with placebo, in order to reduce the duration of unnecessary exposure of the drug or placebo to patients.

However, in the application of Cox cure regression to real data, it is important to assess whether or not the length of follow-up is sufficient before the formal analysis because Cox cure regression in general requires long-term follow-up as you pointed out.

We performed the $q_n$-test for checking whether the follow-up is sufficiently long, using the idea of Maller and Zhou (1996). The result supported that the follow-up was sufficiently long for the sertraline randomized withdrawal study.

These were included in “discussion section” as follows on **Pages 18-19**:

“Cox cure regression (or survival cure CART) generally requires a long-term follow-up [10, 24]. The sertraline study discussed in this paper was intended to contribute to a new drug application in Japan, so the length of the follow-up period (16 weeks) in the study was minimized to merely detect the drug’s effect, in order to reduce the duration of unnecessary exposure of patients to the drug or placebo. However, in the application of Cox cure regression to real data, before the formal analysis, an assessment of whether or not the length of follow-up is sufficient would be useful for interpreting the result. To confirm this for the sertraline data, the $q_n$-test discussed by Maller and Zhou [24] was performed for the sertraline and placebo groups, constructed by the estimated cure rates and censoring distribution for this data. For the sertraline group, the observed value of 0.08547 of $q_n$ was between 94% and 96% critical points of the test, which supported that the length of follow-up for the sertraline group was acceptably minimal and the data had leveled off. On the other hand, for the placebo group, the observed value of 0.0085 of $q_n$ was much smaller
than the value of 0.068 for the 95% point, which did not support that the length of follow-up for the placebo group was sufficient and that the data had leveled off. According to the results of two $q_{nt}$-tests, we could conclude that the length of follow-up period in the sertraline study was sufficient at least to detect the drug’s effect compared with placebo.”

**Comment #2**
For the real data application, it is desirable that the authors can compare the results obtained from the Cox cure model with those of the standard Cox model without cure fraction, and show that the Cox cure model considered in this paper can provide a better fit to the data.

**Response:**

We provided more detailed discussion and finding on the application of Cox cure regression and standard Cox regression, as follows on **Pages 15-16**:

“The result of standard stepwise logistic and Cox regressions with the minimum value of AIC is shown in Table 4. Standard logistic and Cox regressions suggested the effects of treatment and gender, which were in agreement with the result of Cox cure regression. In addition to the baseline HAM-D score at DP and complication, the standard logistic regression indicated a weak effect on the number of episodes. On the other hand, standard Cox regression was not able to detect the effect of baseline HAM-D score at DP, but it indicated a weak effect on the number of episodes. This was a major discrepancy between Cox cure regression and standard Cox regression. Compared with the best subset of covariates of standard logistic and Cox regressions, Cox cure regression suggested the covariates that were more important for cured incidences but less important for the uncured survivals, and vice versa. The subset for Cox cure regression provided a smaller value of AIC than that of standard Cox regression. The standard Cox regression model is a special case of Cox cure regression with an infinitely small intercept in cured incidences [10]. The estimate and SE for the intercept were 1.571 and 2.870, respectively. These results support the use of Cox cure regression rather than standard Cox regression.”

**Comment #3**
The authors should give more interpretations for the tree results obtained using the cure survival CART, such as Figure 2. What is the main difference of the tree results compared to those based on variable selection of the Cox cure model?

**Response:**

We provided more detailed discussion and finding on the application of the cure survival CART, as follows on **Pages 16-17**:

“Next, the cure survival CART was used to identify groups of patients with differing prognoses, and to refine the model previously obtained by Cox cure regression with
meaningful interpretation. Four covariates (treatment, baseline HAM-D score at DP, gender and complication), which were identified by Cox cure regression, were selected to characterize the time to relapse for the sertraline data. In the application of the cure survival CART with exponential distribution, based on the average of 500 replications, the minimum cross-validated deviance residual was \( R^\alpha(\gamma_k) = 1.753 \) (SE = 0.083), and the corresponding tree \( T_k \) was quite large. The 1-SE rule was used to choose a simpler tree, shown in Figure 2, which provides the cross-validated estimate of \( R^\alpha(\gamma) = 1.791 \). Similarly, for the cure survival CART with Weibull distribution, the minimum cross-validated estimate based on the average of 500 replications is \( R^\alpha(\gamma_k) = 1.981 \) (SE = 0.229), and the corresponding tree \( T_k \) is shown in Figure 3. As seen in Figure 2, the cure survival CART with exponential distribution showed that the primary split was gender and the secondary was treatment for the subgroup of female patients. Further partitioning of the tree was based on the baseline HAM-D score at DP of 6 points for the subgroup of female patients assigned to the sertraline treatment group. On the other hand, as seen in Figure 3, the cure survival CART with Weibull distribution provided a simpler tree structure compared with that seen in the cure survival CART with exponential distribution; there was no partitioning of the tree based on the baseline HAM-D score at DP for the subgroup of female patients assigned to the sertraline treatment group. By comparison with the cross-validated estimates between exponential and Weibull distributions, we could find that the cure survival CART with exponential distribution provided a better fit than that with Weibull distribution.”

<table>
<thead>
<tr>
<th>Comment #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the cure survival CART, the exponential distribution is used for the event time of uncured subject, which is too restrictive. It will be more useful that the authors can extend the method to the usual semiparametric Cox cure model or at least the Weibull distribution for the event time of uncured subject.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response:</th>
</tr>
</thead>
<tbody>
<tr>
<td>In addition to the exponential distribution, we considered the Weibull distribution for the event time of uncured subject and described both the results on Page 16-17 (Please see Response to Comment #3).</td>
</tr>
</tbody>
</table>
Dear Referee #2,

Thank you very much for your constructive comments and valuable suggestions with regard to our manuscript. Specific responses to each of your comment are provided as follows.

**Major Compulsory Revisions:**

<table>
<thead>
<tr>
<th>Comment #1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The novelty of this work is to fit both the Cox cure model and cure survival CART for the data. Using both Cox cure model and cure survival CART is rarely seen in the literature and the authors made a good attempt to combine the two methods with a hope that the combination of the two methods will reveal more than using them individually for the data. Unfortunately, the work does not show any solid gain from the proposed approach. The results from the Cox cure model look good. But it is not clear how using cure survival CART will improve the results from Cox cure model. There is very little discussion on the results from cure survival CART in the discussion and conclusion sections.</td>
</tr>
</tbody>
</table>

**Response:**

We provided more descriptions of combination of use survival cure CART with Cox cure regression as follows on Page 17:

“From the results obtained by the two cure survival CARTs, refined Cox cure regression was reperformed; the model included treatment, the baseline HAM-D score at DP and the interaction between the treatment and the baseline HAM-D score at DP into cured incidence, and gender and complication into uncured survivals. The baseline HAM-D score at DP was categorized into two groups: baseline HAM-D score at DP >6 (=1) and baseline HAM-D score at DP <6 (=0).

The result of refined Cox cure regression is shown in Table 5. For cured incidence, a major difference between original and refined Cox cure regressions was a significant effect of the interaction between the treatment and the categorized baseline HAM-D score at DP. Refined Cox cure regression provided the negative estimates for the interaction, which indicated a higher cure rate for patients with a lower value of the score who received sertraline treatment, but a lower cure rate for patients with a higher value of the score who received the placebo. For uncured survival, there was no major difference between the original and refined Cox cure regressions. Also, as refined Cox cure regression provided a smaller AIC value than original Cox cure regression, refined Cox cure regression would lead to an improved fit. Figure 4 shows estimated curves for the time to relapse for each combination of the treatment and categorized baseline HAM-D score at DP, adjusted by gender and complication, using refined Cox regression. The important differences in cured incidence were observed between the sertraline and placebo groups with regard to the baseline HAM-D score.”
**Comment #2**

I also have concern over the application of Cox cure regression to the sertraline data starting on page 13. First of all, it is not clear where the three guidelines came from. It is not clear either how the AIC values for the semiparametric Cox’s cure regression are computed, and how the AIC values are combined with the guidelines to produce the results in Table 3. The AIC values should be given in the table too.

**Response:**

We provided more details of our aspect of how to select the variables in the application of the Cox cure regression to the sertraline data. However, the aspect may differ, depending on the scientific interest and the area of application. So that we provided some note on the aspect of variable selection in discussion section.

Also we provided more detailed descriptions of how the AIC values for the semiparametric Cox’s cure regression are computed, and how the AIC values are combined with the guidelines to produce the results in Table 3. The AIC and log-maximum likelihood values were included in the tables.

These are as follows:

**On Pages 14-15**

“In our situation, cured incidences in this model are more scientific and practical, as the objective of the study was to show how well the drug prevents an eventual episode of recurring illness (cured incidence), compared with placebo and, further, how other covariates influence a cured incidence. Thus, in selecting the covariates, we considered that, giving priority to cured incidences over uncured survivals, a minimal number or no covariate in uncured survivals may be appropriate, and then we suggested the following guidelines on variable selection in our situation: (1) treatment is always factored into cured incidence, (2) either a minimal number or no covariates are included into uncured survivals, (3) a covariate already included into cured incidence is not included into uncured survivals. Following these guidelines, the “best” subset of all possible combinations of covariates can be selected by a minimum value of Akaike’s information criterion (AIC) [20], given by $AIC = -2l_f(\hat{\theta}, \hat{\Lambda}_0) + 2(p + q)$; $(\hat{\theta}, \hat{\Lambda}_0)$ is $(\theta, \Lambda_0)$ that maximizes $l_f(\theta, \Lambda_0)$, discussed in the previous Methods section.”

**On Page 19**

“In variable selection for the fitting of Cox cure regression to the sertraline data, we gave priority to cured incidence over uncured survivals. However, if there was more scientific interest in when the illness may recur rather than in the eventual cure, giving priority to uncured survivals over cured incidence could be appropriate. There are several aspects of variable selection, depending on the applications of interest.”
**Comment #3**
The authors indicated that they tried to examine the results from the semiparametric Cox cure model by fit the data with cure survival CART based on a parametric Cox cure model with exponential baseline distribution. Unless the true baseline distribution for uncured patients is approximately exponential, the estimates from the semiparametric Cox cure model and those from the parametric Cox cure model are usually sensitive to the baseline distribution assumption. The authors did not consider the potential inconsistency between the two models when they considered both for their data

**Response:**

We added notes on

- the inconsistency between the two models when they considered both for their data (on Page 19), and
- the reason why we preferred the fitting of cure regression with the exponential distribution, rather than the straightforward fitting to the Cox cure regression model (on Page 8)

In addition to the exponential distribution, following to the comment by Referee #1, we considered Weibull distribution for the latency time of relapse.

**On Page 19**

“In the paper, we discussed the two methods of cure Cox regression and cure survival CART. As described in Method section, the former method is a semiparametric regression, but the latter method use a parametric cure regression. Although the cure survival CART output provided information in refining Cox cure regression leading to meaningful interpretations for the sertraline data, note that there is the potential inconsistency between the two regressions when they consider both for data as the estimates from the semiparametric Cox cure regression and those from the parametric Cox cure regression are usually sensitive to the baseline distribution assumption. Our future challenge is to develop the semiparametric cure survival CART with fewer amounts of computations.”

**On Page 8**

“With respect to the cost of calculations in estimating the parameters, rather than the straightforward use of the semiparametric Cox cure regression discussed in the previous section, the fitting of parametric cure regression is more reasonable, where the exponential distribution is assumed to be an underlying distribution for the latency time of relapse. However, since an assumption of exponential distribution is restrictive, the Weibull distribution is also considered. First, the model will be described and then the algorithm will be discussed.”

**Minor Essential Revisions**

**Comment #1**

Page 3, line 2: There should be a space between “facts” and “on”.

**Response:**

We corrected it.
**Comment #2**
Page 4, line 13: “he” should be “The”

**Response:**

We corrected it.

**Comment #3**
Page 9, line 7: There should be a pair of parentheses around 1-ch in the numerator. Otherwise the equation would not be right.

**Response:**

We corrected it.

**Comment #4**
Page 10, line 1: No definitions are given for R(l(h)) and R(r(h))

**Response:**

We provided the definitions of R(l(h)) and R(r(h)) in the manuscript as follows on Page 10:

“When growing trees, the improvement measure for a split \((\varepsilon)\) at node \(h\) into left and right daughter nodes, \(l(h)\) and \(r(h)\), respectively, is measured by

\[
R(\varepsilon, h) = R(h) - \{R(l(h)) + R(r(h))\},
\]

where

\[
R(h) = \frac{1}{n} \sum_{i \in \{x \in X_h\}} d_i(\hat{c}_h, \hat{\mu}_h, \hat{\rho}_h),
\]

represents the -2 times maximum log likelihood divided by the sample size \(n\) (deviance residual) obtained for the data of node \(h\) \((= h, l(h), r(h))\); \(\hat{c}_h\), \(\hat{\mu}_h\) and \(\hat{\rho}_h\) are the maximum likelihood estimates (MLEs) of \(c_h\), \(\mu_h\) and \(\rho_h\) obtained for node \(h\),

\[
d_i(\hat{c}_h, \hat{\mu}_h, \hat{\rho}_h) = -2 \left[ \Delta_i \left( \log \hat{\lambda}_h(T_i) + \log \hat{w}_h(T_i) \right) + \log \left[ \hat{c}_h + (1 - \hat{c}_h) \hat{S}_h(T_i) \right] \right],
\]

\[
\hat{\lambda}_h(T_i) = \lambda_h(T_i) |_{\rho_1 = \rho_h, \rho_2 = \rho_h}, \quad \hat{S}_h(T_i) = S_h(T_i) |_{c_h = \hat{c}_h, \mu_h = \mu_h, \rho_h = \rho_h}, \quad \text{and}
\]

\[
\hat{w}_h(T_i) = w_h(T_i) |_{c_h = \hat{c}_h, S_h(T_i) = \hat{S}_h(T_i)}.
\]
“The average cross-validated deviance residual over $V$ subsamples is

$$R_v^\gamma(\gamma) = \frac{1}{n} \sum_{v=1}^{V} \sum_{i \in L_v} \sum_{h \in T(v)} \sum_{(x_i, z_i) \in \mathcal{X}_v} d_i(C_i(h)^{(v)}(\gamma), \hat{\mu}_i^{(v)}(\gamma), \hat{\rho}_i^{(v)}(\gamma)).$$

Let $\gamma_k^*$ be the value of $\gamma$ that minimizes $R_v^\gamma(\gamma)$ of $\gamma = \sqrt{\gamma_k \gamma_{k+1}}$, $k = 0, 1, 2, \ldots$. Then a tree $T_k = T(\gamma_k^*)$ corresponding to $\gamma_k^*$ is selected.

We adopt $V=10$ because a smaller value of $V$ is preferred to $V=n$ in the application of CART [18]. While 10-fold cross-validation is a standard method for selecting tree size, it is subject to considerable variability. Therefore, in the application to this data, we performed 500 replications of 10-fold cross-validation and then determined $\gamma_k^*$ to minimize the average of 500 pairs of \{R_v^\gamma(\gamma), \gamma = \sqrt{\gamma_k \gamma_{k+1}}, k = 0, 1, 2, \ldots\}. We will be able to use the “1-SE (Standard Error)” rule [6] to choose a simpler tree, where such an SE is directly estimated by the variation of $R_v^\gamma(\gamma)$ with 500 replications in this application.”

**Comment #6**

Page 13, line 7: I don’t think this claim is appropriate for Cox’s cure model (and I don’t think ref[16] supports this claim either). Cox’s cure model itself will not lead to an overparameterization and it does not have identifiability issue if the tail of the baseline uncured distribution is properly handled as suggested in ref [16].

**Response:**

We are afraid that we refer the wrong paper: it is discussed not in ref [16] but in ref [10]. We provided more detailed description of the aspect of variable selection in the manuscript as follows on Page 14:

“Cox cure regression, including influential covariates for both cured incidences and uncured survivals, provides flexibility in model building. However, this approach may open discussions on the possibility of an overparameterization of models and the identifiability between the parameters of cured incidences and uncured survivals [10], although the parameters of the standard cure model are identifiable in some sense [19]. In our situation, cured incidences in this model are more scientific and practical, as the objective of the study was to show how well the drug prevents an eventual episode of recurring illness (cured incidence), compared with placebo and, further, how other covariates influence a cured incidence.”

**Comment #7**

Table 4 is not discussed anywhere in the paper

**Response:**

We provided more detailed discussion on the Cox cure regression and the standard logistic and Cox regressions, referring to Table 4.
Comment #8
Page 15, line 8 from bottom: Delete “what patients”

Response:

We deleted it.