

## A NOTE ON MISSPECIFIED TRANSFORMATION MODELS

Hattori, Satoshi  
The Biostatistics Center., Kurume University

<https://doi.org/10.5109/12601>

---

出版情報 : Bulletin of informatics and cybernetics. 38, pp.51-57, 2006-12. Research Association of Statistical Sciences

バージョン :

権利関係 :



# A NOTE ON MISSPECIFIED TRANSFORMATION MODELS

by

Satoshi HATTORI

---

*Reprinted from the Bulletin of Informatics and Cybernetics  
Research Association of Statistical Sciences, Vol.38*

---

FUKUOKA, JAPAN  
2006

# A NOTE ON MISSPECIFIED TRANSFORMATION MODELS

By

Satoshi HATTORI\*

## Abstract

The validity of the test of no treatment effect adjusting some covariates based on the misspecified transformation model, which includes the proportional odds model as a special case, is investigated. It is proven that if adjusted covariates are independent of treatment allocation, which is naturally satisfied in randomized clinical trials, and censoring mechanism is independent of the treatment allocation and covariates to adjust, the tests of no treatment effect based on the possibly misspecified transformation models are valid asymptotically. Results of a small simulation study are also demonstrated.

*Key Words and Phrases:* Covariate adjustment, Proportional odds model, Randomized clinical trial, Type 1 error.

## 1. Introduction

In medical researches, the primary interest is often comparison among two or several treatments. For example, in clinical trials to evaluate effectiveness of a chemotherapy for cancer patients, time to death or time to disease progression are often primary endpoints. The distributions of time to event are compared among two or several treatment groups. To this end, the logrank test is frequently used.

If there are some strong prognosis factors, imbalance of distributions of these factors may cause misleading results of the hypothesis test. Randomization methods, such as the stratified randomization and the minimization method, are frequently used to avoid imbalance of distributions of strong prognosis factors and factors accounted for at randomization are usually adjusted in analysis. It is required that the method to adjust covariates should be pre-specified in the protocols to avoid overfitting of the statistical models (Committee for Proprietary Medicinal Products (CPMP) 2003). That implies the fitted models may be misspecified. Then the impact of model misspecification on the validity of the hypothesis test is of great concern.

For the Cox proportional hazards models, Lin and Wei (1989) derived a robust variance estimate of the maximum partial likelihood estimator and showed that with it one can test the hypothesis of no treatment effect in randomized clinical trials even if the fitted Cox model does not hold. Related investigations were made by Kong and Slud (1997) and Dirienzo and Lagakos (2001a, 2001b). As stated, the Cox proportional hazards model with the robust variance estimate provides a valid test of no treatment effect even under misspecification. However, misspecification may cause serious loss

---

\* The Biostatistics Center., Kurume University, 67 Asahi-Machi, Kurume City, Fukuoka 830-0011, Japan. E-mail: hattori satoshi@med.kurume-u.ac.jp

of power (Kong and Slud 1997). One alternative is the proportional odds model and some inference procedures were developed (Murphy, Rossini and Van der Vaart 1997, Schefenstien, Tsiatis and Gilbert 1998, Yang and Prentice 1999). Cheng, Wei and Ying (1995) proposed an inference procedure for the transformation models, which contains the proportional odds model and the Cox proportional hazards model as special cases. Similar to the maximum partial likelihood estimation for the Cox proportional hazards model, their estimating function of the regression coefficients are free from the nuisance baseline survival function and can be solved easily by the standard Newton-Raphson method without complicated nonparametric estimation of the baseline survival function.

In this article, we study the validity of the test of no treatment effect based on possibly misspecified transformation models. We show that the variance estimate by Cheng, Wei and Ying (1995), which was derived under the assumption of correct model specification, is still a valid variance estimate of the limit of the estimator of the regression coefficient even if misspecified. In addition, it is proven that, under the null hypothesis of no treatment effect, the estimator of the regression coefficient corresponding to the treatment allocation consistently estimates 0 if the treatment allocation and the covariates to adjust are independent and the failure time and the censoring time are independent of the treatment allocation and the covariates. Therefore the test of no treatment effect based on the transformation models with estimation procedure by Cheng, Wei and Ying (1995) is valid even if misspecified in randomized clinical trials without any modification.

The organization of this article is as follows. In section 2, some notations are prepared for. In section 3, we summarized the inference procedure for the transformation model proposed by Cheng, Wei and Ying (1995), and some properties of the estimator in the presence of model misspecification. In section 4, some results of a simulation study is demonstrated. The proof is shown in the appendix.

## 2. Notation

Consider two treatment comparisons in randomized clinical trials. We test the null hypothesis of no treatment effect based on the transformation model, which will be explained in the next section. Let  $Z^{(1)}$  be a binary random variable, which represents the random treatment allocation. The allocation rate is denoted by  $q \in (0, 1)$ . Let  $T$  and  $C$  be the failure time and the censoring time, respectively. The failure time  $T$  is assumed to be right-censored by  $C$  and then only  $X = \min(T, C)$  and  $\Delta = I(T \leq C)$  can be observed. Let  $W$  be a  $p$ -dimensional covariate vector to adjust. Define  $(p+1)$ -dimensional vector  $Z$  as  $Z^T = (Z^{(1)}, W^T)$ . Let  $\{(X_i, \Delta_i, Z_i)\}_{i=1}^n$  be  $n$  i.i.d. copies of  $(X, \Delta, Z)$ .

Consider the null hypothesis  $H_0$  of no treatment effect that the failure time  $T$  does not depend upon  $Z^{(1)}$  conditional on covariate  $W$ . Suppose that  $Z^{(1)}$  and  $W$  are independent, that is naturally satisfied in randomized clinical trials. In addition, the censoring time  $C$  is assumed to be independent of  $T$  and to be independent of  $Z$ , that was assumed in Cheng, Wei and Ying(1995).

## 3. Validity of misspecified transformation models

### 3.1. Inference procedure

The transformation model provides a wide class of semiparametric regression models for

survival analysis which contain the Cox proportional hazards model and the proportional odds model as special cases. It is defined as

$$h_0(T) = -\beta_0^T Z + \varepsilon, \quad (1)$$

where  $h_0(t)$  is a completely unspecified strictly increasing function,  $\beta_0$  is a  $(p+1)$ -dimensional regression coefficient and  $\varepsilon$  is a random error term whose distribution function  $F$  is completely specified. The first element of  $\beta_0$ , which is corresponding to the treatment effect, is denoted by  $\beta_0^{(1)}$  and the  $p$ -dimensional vector of the 2nd to  $(p+1)$ -th components is denoted by  $\beta_0^{(2:p+1)}$ . If  $F$  is the standard extreme value distribution, the model is the Cox proportional hazards model, and if  $F$  is the standard logistic distribution, the model is regarded as the proportional odds model, which satisfies the relation,

$$\log\left(\frac{1 - S_Z(t)}{S_Z(t)}\right) = h_0(t) + \beta_0^T Z,$$

where  $S_Z(t)$  is the conditional survival function of  $T$  given  $Z$  (Cheng, Wei and Ying 1995). In the previous section, we assumed that  $C$  is independent of  $Z$ . Then let  $G$  be the common survival function of  $C$ . Cheng, Wei and Ying (1995) proposed an estimating function for the regression coefficients  $\beta$ ,

$$U_n(\beta) = \sum_{i=1}^n \sum_{j=1}^n Z_{ij} \left\{ \frac{\Delta_j I(X_i \geq X_j)}{\hat{G}^2(X_j)} - \xi(Z_{ij}^T \beta) \right\},$$

where  $\xi(s) = \int_{-\infty}^{\infty} \{1 - F(t+s)\} dF(t)$ ,  $Z_{ij} = Z_i - Z_j$ ,  $\hat{G}$  is the Kaplan-Meier estimator for  $G$ . Note that originally the estimating function involved the weight function, but in our argument the weight function is restricted to the constant one to ensure the existence of a unique solution of the limiting estimating equation under misspecification. Cheng, Wei and Ying (1995) showed that  $\hat{\beta}$ , which is a solution of the equation  $U_n(\beta) = 0$ , converges to the true value  $\beta_0$  almost surely and the distribution of  $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$  converges in distribution to a normal distribution with mean zero whose asymptotic variance-covariance matrix can be consistently estimated by the sandwich-type variance estimate. Note that these theoretical properties are based on the assumption of the correct model specification.

### 3.2. Validity under misspecification

In this subsection, we discuss asymptotic properties when the transformation model is misspecified. Here we consider a situation where the transformation model is applied but the true conditional distribution of the failure time given covariates can not be represented with the transformation models (1). Note that our following argument allows any misspecification; some important covariates may be omitted, the functional form of covariates may be misspecified, the distributional assumption of  $\varepsilon$  may be inappropriate and so on. It can be shown that there exists a unique solution of the limiting estimating equation of  $n^{-2}U_n(\beta)$ . Let the unique solution be denoted by  $\beta_*$ . It can be shown that  $\hat{\beta}$  converges to  $\beta_*$  almost surely. And using the argument similar to that by Cheng, Wei and Ying (1995), the sandwich-type variance estimate which derived under the correct

model specification is still valid as the variance estimate of  $n^{-\frac{1}{2}}(\hat{\beta} - \beta_*)$ . In general, the result is not useful in practice since  $\beta_*$  has no practical meaning and unknown. However it is shown that  $\beta_*^{(1)}$ , which is the first element of  $\beta_*$ , is 0 under the null hypothesis  $H_0$  of no treatment effect even if misspecified. That implies that the hypothesis test of no treatment effect based on  $\hat{\beta}$  is still valid even under misspecification in randomized clinical trials. The proof is provided in the appendix.

#### 4. Numerical study

We showed that the inference procedure by Cheng, Wei and Yin (1995) for the transformation model provides an asymptotically valid test of no treatment effect adjusting some covariates even if the fitted models are misspecified in randomized clinical trials. In this section, we present some results of a simulation study to investigate the small sample performance. We focused on the criterion validity such that even when the fitted models were misspecified, the type I errors kept their nominal size in small sample size situations. To this end, we generated 1000 datasets from the Cox proportional hazards models and the proportional odds model under the null hypothesis of no treatment effect. Details of the models to generate datasets will be stated below. We fitted the transformation models to these datasets and evaluated empirical sizes.

We intend two treatment comparisons in randomized clinical trials, where patients are randomly allocated to each treatment group with probability of 0.5. The dataset 1 was generated from the Cox proportional hazards model,  $T^{\frac{1}{2}} = 0 \times Z^{(1)} - 2 \times W_1 + \epsilon$ , where  $Z^{(1)}$  was a binary random variable with  $\text{prob}\{Z^{(1)} = 1\} = 0.5$ , which represents the treatment allocation,  $W_1$  was a covariate to adjust following the uniform distribution on  $[0, 1]$  independent of  $Z^{(1)}$  and  $\epsilon$  was a random error following the standard extreme value distribution. The dataset 2 was from the proportional odds model,  $T^{\frac{1}{2}} = 0 \times Z^{(1)} - 2 \times W_1 + \epsilon$ , where the setting of  $Z^{(1)}$  and  $W_1$  was same as that for dataset 1 and  $\epsilon$  was the standard logistic random error. By adding an additional covariate  $W_2$  to the models for the dataset 1 and 2, the dataset 3 and 4 were generated respectively, where  $W_2$  was a uniform random variable on  $[0, 1]$  independent of  $Z^{(1)}$  and  $W_1$ ; the dataset 3 and 4 were generated by  $T^{\frac{1}{2}} = 0 \times Z^{(1)} - 2 \times W_1 + 3 \times W_2 + \epsilon$  where  $\epsilon$ s were those for the proportional hazards and the proportional odds models, respectively.

For each dataset, the number of the subject was set 50 and 100 and 10% data was censored uniformly. For these datasets, we fitted the transformation models with the treatment allocation  $Z^{(1)}$  and the covariate  $W_1$ . Note that the covariate  $W_2$  were omitted. We evaluated the empirical size for the nominal two-sided 5% significance level. The results for the Cox proportional hazards and the proportional odds models were summarized in Table 1 and Table 2 respectively. Note that for the Cox proportional hazards models, the inference procedure by Cheng, Wei and Ying (1995) was used instead of the maximum partial likelihood method.

In all the numerical experiments, the empirical sizes were close to the nominal level even when the number of the subjects was rather small. Judging from our numerical study, the transformation models seems to provide an appropriate test of no treatment effect even if misspecified, in small sample size situations.

simulation data	n= 50	n=100
data 1 (PH with $W_1$ )=correct specification	6.0%	4.6%
data 2 (PO with $W_1$ )	5.7%	3.8%
data 3 (PH with $W_1$ and $W_2$ )	3.5%	3.4%
data 4 (PO with $W_1$ and $W_2$ )	5.7%	4.8%

Table 1: *Empirical sizes for transformation model (Cox proportional hazards model) with  $Z$  and  $W_1$  under misspecification*

$PH$ =Cox proportional hazards model,  $PO$ =proportional odds model,

simulation data	n= 50	n=100
data 1 (PH with $W_1$ )	7.2%	4.7%
data 2 (PO with $W_1$ )=correct specification	5.8%	3.9%
data 3 (PH with $W_1$ and $W_2$ )	3.6%	3.4%
data 4 (PO with $W_1$ and $W_2$ )	6.0%	4.8%

Table 2: *Empirical sizes for transformation model (proportional odds model) with  $Z$  and  $W_1$  under misspecification*

$PH$ =Cox proportional hazards model,  $PO$ =proportional odds model,

## 5. Appendix

Similar to the definitions of  $\beta_0^{(1)}$  and  $\beta_0^{(2:p+1)}$ ,  $\beta^{(1)}$  and  $\beta^{(2:p+1)}$  are defined for a  $(p+1)$ -dimensional vector  $\beta$ . It is assumed that each element of the covariate vector is bounded and the parameter space is compact.  $E(Z_{12}Z_{12}^T) > 0$  is assumed. In addition,  $\xi'(u) < 0$  is assumed. For the Cox proportional hazards model and the proportional odds model,  $F(t)$  is the cumulative distribution function of the extreme value distribution and the standard logistic distribution respectively. In each cases,  $\frac{dF(t)}{dt}$  is continuous and positive. Then this assumption is satisfied since  $\xi'(u) = -\int_{-\infty}^{\infty} f(t+u)dF(t)$ .

### Consistency and asymptotic normality

Let

$$\begin{aligned}\tilde{H}_n(\beta) &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left[ \beta^T Z_{ij} \frac{\Delta_j I(X_i \geq X_j)}{\hat{G}^2(X_j)} - \int_{-\infty}^{Z_{ij}^T \beta} \xi(u) du \right], \\ H_n(\beta) &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left[ \beta^T Z_{ij} \frac{\Delta_j I(X_i \geq X_j)}{G^2(X_j)} - \int_{-\infty}^{Z_{ij}^T \beta} \xi(u) du \right], \\ H(\beta) &= E \left[ \beta^T Z_{12} \frac{\Delta_2 I(X_1 \geq X_2)}{G^2(X_2)} - \int_{-\infty}^{Z_{12}^T \beta} \xi(u) du \right],\end{aligned}$$

and

$$U_{\infty}(\beta) = E \left[ Z_{12} \left\{ \frac{\Delta_2 I(X_1 \geq X_2)}{G^2(X_2)} \xi(Z_{12}^T \beta) \right\} \right].$$

Since  $\frac{\partial^2 H(\beta)}{\partial \beta^2} = -E\{Z_{12}Z_{12}^T \xi'(Z_{12}^T \beta)\} > 0$ ,  $H(\beta)$  is a convex function. Then there exists a unique minimizer, say  $\beta_*$ . By the standard law of large number for U-statistics and the compactness of the parameter space,

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \beta^T Z_{ij} \frac{\Delta_j I(X_i \geq X_j)}{G^2(X_j)} \rightarrow E\left\{ \beta^T Z_{12} \frac{\Delta_2 I(X_1 \geq X_2)}{G^2(X_2)} \right\}$$

almost surely uniformly in  $\beta$ . By lemma 22 of Nolan and Pollard (1987),  $\int_{-\infty}^{Z_{ij}^T \beta} \xi(u) du$  is Ecldian. Then by the same argument in Proposition A1. of Jin, Ying and Wei (2001),

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \int_{-\infty}^{Z_{ij}^T \beta} \xi(u) du \rightarrow E\left\{ \int_{-\infty}^{Z_{12}^T \beta} \xi(u) du \right\}$$

almost surely uniformly in  $\beta$ . Then  $H_n(\beta)$  converges to  $H(\beta)$  almost surely uniformly in  $\beta$ . Note that  $E[H_n(\beta)] = H(\beta)$  and  $\beta_*$  is a unique minimizer of  $H(\beta)$ . As defined in section 3.1,  $\hat{\beta}$  is a solution of  $U_n(\beta) = 0$  and maximizes  $\tilde{H}_n(\beta)$ . Then by Proposition A1. of Jin, Ying and Wei (2001) and the asymptotic equivalence between  $H_n(\beta)$  and  $\tilde{H}_n(\beta)$ ,  $\hat{\beta}$  converges to  $\beta_*$  almost surely as  $n \rightarrow \infty$ . Furthermore if  $\beta_*$  is an interior point of the parameter space, by expanding  $U_n(\hat{\beta}) - U_n(\beta_*)$  around  $\beta_*$ , the asymptotic normality can be shown in the same way as Cheng, Wei and Ying (1995).

**$\beta_*^{(1)} = 0$  under the null hypothesis  $H_0$**

Let  $W_{12} = W_1 - W_2$ .

In section 2, it is assumed that under  $H_0$ ,  $T$  is independent of  $Z^{(1)}$  conditional on  $W$  and  $C$  is independent of  $Z$ . Therefore  $X = \min(T, C)$  is independent of  $Z^{(1)}$  conditional on  $W$ .

$$\begin{aligned} & \text{the 1st element of } U_\infty(\beta_*) \\ &= E\left[E\{Z_{12}^{(1)} | W_1, W_2\} E\left\{ \frac{\Delta_2 I(X_1 \geq X_2)}{G^2(X_2)} | W_1, W_2 \right\}\right] - E\{Z_{12}^{(1)} \xi(Z_{12}^T \beta_*)\} \\ &= E(Z_{12}^{(1)}) E\left\{ \frac{\Delta_2 I(X_1 \geq X_2)}{G^2(X_2)} \right\} - E\{Z_{12}^{(1)} \xi(Z_{12}^T \beta_*)\} \end{aligned}$$

Note that the first equality holds since  $X_1$  and  $X_2$  are independent of  $Z_1^{(1)}$  and  $Z_2^{(1)}$  conditional on  $W_1$  and  $W_2$  under  $H_0$  and the second equality does since  $Z_1^{(1)}$  and  $Z_2^{(1)}$  are assumed to be independent of  $W_1$  and  $W_2$  in section 2. Furthermore

$$\begin{aligned} &= - \sum_{i=0,1} \sum_{j=0,1} P(Z_1^{(1)} = i, Z_2^{(1)} = j) E\left\{ Z_{12}^{(1)} \xi(Z_{12}^T \beta_*) \middle| Z_1^{(1)} = i, Z_2^{(1)} = j \right\} \\ &= -q(1-q) E\left[ \xi(\beta_*^{(1)} + W_{12}^T \beta_*^{(2:p)}) \right] - q(1-q) E\left[ -\xi(-\beta_*^{(1)} + W_{12}^T \beta_*^{(2:p)}) \right] \\ &= -2q(1-q) \beta_*^{(1)} E\left[ \xi'(x^+) \right] \end{aligned}$$

where  $x^+$  is a line segment between  $\beta_*^{(1)} + W_{12}^T \beta_*^{(2:p+1)}$  and  $-\beta_*^{(1)} + W_{12}^T \beta_*^{(2:p+1)}$ . Because  $\beta_*$  satisfies  $U_\infty(\beta_*) = 0$ ,  $-2\beta_*^{(1)} E\{\xi'(x^+)\} = 0$ . Then since  $E\{\xi'(x^+)\} \neq 0$  by the negativity of  $\xi'(x^+)$ ,  $\beta_*^{(1)} = 0$  is obtained.



### Acknowledgement

The author would like to thank the anonymous referee for careful reading and helpful comments.

### References

- Cheng, S. C., Wei, L. J. and Ying, Z. (1995). Analysis of transformation models with censored data. *Biometrika*, **82**, 835–845.
- Committee for Proprietary Medicinal Products (CPMP). (2003). Points to consider on adjustment for baseline covariates. *Statistics in Medicine*, **23**, 701–709.
- DiRienzo, A. G. and Lagakos, S. W. (2001a). Effects of model misspecification on tests of no randomized treatment effect arising from Cox’s proportional hazards model. *J. Roy. Statist. Soc. B*, **63**, 745–757.
- DiRienzo, A. G. and Lagakos, S. W. (2001b). Bias correction for score tests arising from misspecified proportional hazards regression models. *Biometrika*, **88**, 421–434.
- Jin, Z., Ying, Z. and Wei, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika*, **88**, 381–390.
- Kong, F. H. and Slud, E. (1997). Robust covariate-adjusted log-rank tests. *Biometrika*, **84**, 847–862.
- Lin, D. Y. and Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *J. Amer. Statist. Assoc.*, **84**, 1074–1078.
- Murphy, S. A., Rossini, A. J. and Van der Vaart, A. W. (1997). Maximum likelihood estimation in the proportional odds model. *J. Amer. Statist. Assoc.*, **92**, 968–976.
- Nolan, D. and Pollard, D. (1987). U-processes: rate of convergence. *Ann. Statist.*, **15**, 780–799.
- Scharfstein, D. O., Tsiatis, A. A. and Gilbert, P. B. (1998). Semiparametric efficient estimation in the generalized odds-rate class of regression models for right-censored time-to-event data. *The Lifetime Data Analysis*, **4**, 355–391.
- Yang, S. and Prentice, R. L. (1999). Semiparametric inference in the proportional odds regression model. *J. Amer. Statist. Assoc.*, **94**, 125–136.

*Received April 28, 2005*

*Revised November 16, 2005*