Isotonic proportional hazard model for breast cancer patients

Meme kanseri hastaları için izotonik orantılı tehlike modeli

Abstract

Aim: There are many survival analysis methods in oncological studies. Each model is used in different data structures. The isotonic proportional hazard model is a survival model among these methods. The use of isotonic models is proposed to evaluate the relationship between time to event outcome and monotonically increasing covariate in survival analysis. The aim of this study is to explain the theoretical properties and usage of isotonic models and demonstrate their application on an appropriate dataset.

Methods: Data on breast cancer patients treated at Seoul National University Hospital was used to make the application of isotonic models. In the modeling of recurrence risk, estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and proliferation marker (Ki-67) variables were considered as independent risk factors. The tumor size was included in the model as covariate. The isotonic proportional hazard model was used in modeling. Estimation of model coefficients was made by partial likelihood method.

Results: The tumor size had monotone increasing effect on recurrence events. It was determined that the risk of recurrence increased 1.008 times as the value of Ki-67 increased. It has been observed that ER and PR negativity increase the risk, while HER2 positivity increases the risk.

Conclusion: As a result, in cases where the covariate variable has a monotone increasing effect on the disease process, the isotonic regression model can be used by considering the relationships between the covariate and the event. Thus, the function that takes into account the monotone relationship between the covariate and the event will be included in the survival analysis.

Keywords: breast diseases, confounding factors, proportional hazards models, survival

Öz

Amaç: Onkolojik çalışmalarda çok sayıda sağkalım analiz yöntemi kullanılmaktadır. Farklı veri yapıları için farklı modeller tercih edilir. İzotonik orantılı tehlike modeli bu yöntemlerden birisidir. Olayın gerçekleştiği zamana kadar geçen süre ile monotonik olarak artan ortak değişken arasındaki ilişkiyi değerlendirmek için izotonik modellerin kullanılması önerilmektedir. Bu çalışmanın amacı, izotonik modellerin teorik özelliklerini ve kullanım alanlarını açıklamak ve uygun bir veri seti üzerinde uygulamasını göstermektir.

Yöntemler: İzotonik modellerin uygulanması için Seul Ulusal Üniversite Hastanesinde tedavi edilen meme kanseri hastalarına ait veriler kullanıldı. Tekrarlama riskinin modellenmesinde östrojen reseptör (ER), progesteron reseptör (PR), insan epidermal büyüme faktör reseptörü 2 (HER2) ve proliferasyon markerleri (Ki-67) değişkenleri bağımsız risk faktörleri olarak kabul edildi. Tümör boyutu, modele ortak değişken olarak dahil edildi. Modellemede izotonik orantılı tehlike modeli kullanıldı. Model katsayılarının tahmini kısmi olabilirlik yöntemi ile yapıldı.

Bulgular: Nüksetme üzerinde tümör boyutunun monoton artırıcı etkisi olduğu gözlendi. Ki-67 değeri arttıkça tekrarlama riskinin 1.008 kat arttığı belirlendi. ER ve PR negatifliğinin riski artırdığı, HER2 pozitifliğinin ise riski artırdığı gözlemlendi.

Sonuç: Ortak değişkenin hastalık sürecini monoton artırıcı etkisinin olduğu durumlarda, ortak değişken ve olay arasındaki ilişkiler dikkate alınarak izotonik regresyon modeli kullanılabilir. Böylece, ortak değişken ile olay arasındaki monoton ilişkiyi hesaba katan fonksiyon, hayatta kalma analizine dahil edilecektir.

Anahtar Sözcükler: karıştırıcı faktör, meme hastalıkları, orantısal risk modelleri, sağkalım

Ozge Pasin[1], Handan Ankarali[2]

[1] Department of Biostatistics and Medical Informatics, Faculty of Medicine, Bezmialem Vakif University

[2] Department of Biostatistics and Medical Informatics, Faculty of Medicine, Istanbul Medeniyet University

Received/*Geliş* : 01.06.2022
Accepted/*Kabul*: 08.09.2022

DOI: 10.21673/anadoluklin.1124700

Corresponding author/*Yazışma yazarı*

Handan Ankaralı

Istanbul Medeniyet University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, İstanbul, Türkiye
E-mail: handanankarali@gmail.com

ORCID

Özge Pasin: 0000-0001-6530-0942
Handan Ankaralı: 0000-0002-3613-0523

INTRODUCTION

The Cox regression model is widely used to examine factors associated with survival times. One of the assumptions required by this model is that the observations are independent of each other and the risk ratio does not change over time. This assumption is called the proportional risk assumption, and if it is not fulfilled, it refers to the Cox regression model with time-dependent variables (1,2). Isotonic models are proposed to be used in cases where covariates change the risk of hazard to monotone. Isotonic is used when the direction of the trend is strictly increasing, while monotonic could imply a trend that is either strictly increasing or strictly decreasing. Although there are no restrictive assumptions in the Isotonic regression model, covariates can be used monotonically when it reduces and increases the risk of hazard (3,5).

In the modeling process, some variables that are disease markers (for example, the CD4 marker level for HIV risk) are evaluated as covariate. In this case, the baseline or longitunal values of this variable can be added to the model as a function when modeling disease risk. By determining the population at high risk of infection, treatment and treatment timing can be optimized. In the isotonic regression model, the effect of the relevant variable is divided into time intervals so that the effect of the variable is fixed, and the time-related change points are determined and the optimal cut-off points are examined (3,6).

In this study, it is aimed to introduce the theoretical properties of isotonic models, to explain their usage areas and to show their application on a suitable data set.

METHODS

Isotonic models

In model definition, T represents failure time, C represents censored time, and z represents covariate in continuous structure. T and C are conditionally independent of z. It is defined as X=min(T, C) and Δ=I(T. Here I(.) is the indicator function. The observed data are defined as ((i=1,….n). Hazard model is specified as λ(t|Z)=. is unspesicified baseline hazard function and is monotone increasing function. If function is monotone, the estimator defined using the partial likelihood method is obtained with the help of the following equation.

$$\ell(\phi) = \prod_{i=1}^{n} \prod_{t \geq 0} \{ \frac{e^{\phi(Z_i)}}{\sum_{j=1}^{n} Y_j(t) e^{\phi(Z_i)}} \}^{dN_i(t)}$$

$Y_j\ (t)$ is the risk process of i. patient, $N_j\ (t)$ is the counting process.

Unlike the likelihood formulation of isotonic linear models, the partial likelihood of the isotonic proportional hazard model is obtained with integral terms dependent on time and covariate values. The estimator is constrained to be a piecewise constant so that the estimation is accurately estimated at specified intervals. With this restriction, a unique estimator is obtained for the observed z values. Calculating the constrained partial likelihood estimator of the isotonic proportional hazard model is a very difficult process. It has complex calculations. For l(ϕ), the iterative quadratic programming method is used to find the constrained estimator. The use of iterative convex minorant algorithms for problems in isotonic predictions reduces computational complexity. Yung et al (2018) used the pseudo-iterative convex minorant algorithm to overcome these calculations. It iteratively minimizes constrained negative pseudo partial likelihood to obtain partial likelihood estimators (4).

Data set and statistical analysis

As application data, information on 370 patients with stage 2 and 3 breast cancer who received chemotherapy treatment at Seoul National University Hospital between January 2002 and September 2008 were used (7). Ethical approval was not taken in this study. Because the data was open for users. Three patients were excluded from the study because there were many missing observations in their data. In gene expression studies, subtypes of breast cancer were defined according to the presence of estrogen receptor (ER). According to this; ER positive tumors have gene expression similar to luminal cells of the mammary glands. Some of the ER negative tumors are immunohistochemically (IHC) positive for human growth factor 2 receptor (cerbB2-HER2). HER2 negative tumors outside the luminal group show similar immune reactivity to normal gland cells of the mammary glands. Estrogen and progesterone (PR) receptors are specialized proteins found in certain cells in the body. In HER2-positive

Table 1. Cox and Isotonic proportional hazard model results

Variables	Cox proportional hazard model			Isotonic proportional hazard model	
	B	p	HR	B	HR
ER	-0.173	0.537	0.841	-0.148	0.861
PR	-0.446	0.136	0.640	-0.495	0.609
Ki-67	0.009	0.083	1.009	0.008	1.008
HER2	0.089	0.689	1.094	0.079	1.083
Initial tumor size	0.134	0.002	1.143		

ER: Estrogen receptor, PR: progesterone receptor, Ki-67: proliferation marker, HER2: human epidermal growth factor receptor 2, HR: hazard ratio

breast cancer, the grade and proliferation rate are generally high, and the hormone receptors are negative. ER, PR, Ki-67 and HER2 were included in the model as independent variables. Tumor size was included in the model as a covariate because it has a monotonically increasing effect on the hazard ratio. In the group included in the study, the tumor was larger, lymph node positivity, visceral and central nervous system metastases were more common. Accordingly, the expected survival time is also shorter. Ki-67 variant is an antibody that can show proliferating cells within the tumor. Tumors with a higher Ki-67 value tend to be more aggressive and metastases are more common in such cases (8,9). Outcome variable was accepted as recurrence time and status variable recurrence present or absent. The relation between the ER, PR, Ki-67 and HER2 and recurrence time was evaluated by using isotonic proportional hazard model after covariate effect was eleminated. The same relationships were also analyzed with the Cox proportional hazard model to compare the results. $p<0.05$ was accepted as the statistical significance level and calculations were made with the isoSurv package in the R program (10).

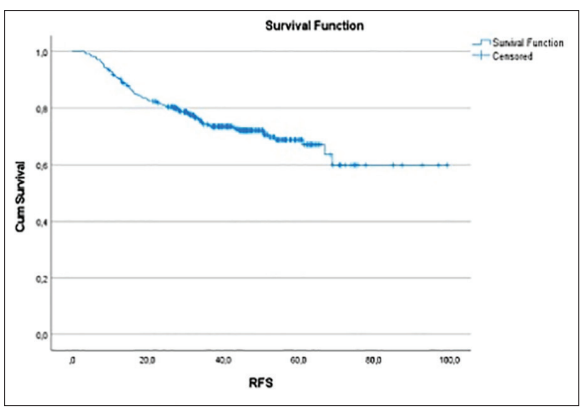

Figure 1. Kaplan Meier curves for recurrence

RESULTS

The mean survival time of 367 patients included in the study was 72.5 months. When the Kaplan Meier curve of the recurrence status of the patients is examined, it is observed that the cumulative recurrence rate between the 65th month and the 100th month is around 60%. (Figure 1).

48.8% (179 people) of the patients were ER positive, 34.9% (128 people) were PR positive and 30.5% (112 people) were HER2 positive. In addition, the median value of Ki-67 is 10 (over percent) (0-80) and the median tumor size is 4.5 cm (0.5-13.5).

When the effect of tumor size on the hazard risk of the patients was examined graphically, it was observed that the risk of recurrence of the disease increased monotonically as the tumor size increased (Figure 2). For this reason, the tumor was included in the model as covariate.

The results of the Isotonic proportional hazard model and the Cox proportional hazard model are summarized in Table 1. Tumor size was included as an independent risk factor in the Cox model. When the isotonic model results are evaluated, ER positivity has a 0.861-effect on hazard risk, while PR positivity has 0.609-effect. As the Ki-67 value increases, the patient's hazard risk increases 1.008 times. HER2 positivity increases the hazard risk 1.083 times. It has been observed that similar results are obtained in the Cox proportional hazard model (Table 1).

DISCUSSION AND CONCLUSION

Although the theoretical foundations of the isotonic regression model have been studied before, it is seen that it is rarely used in the analysis of survival data (11).

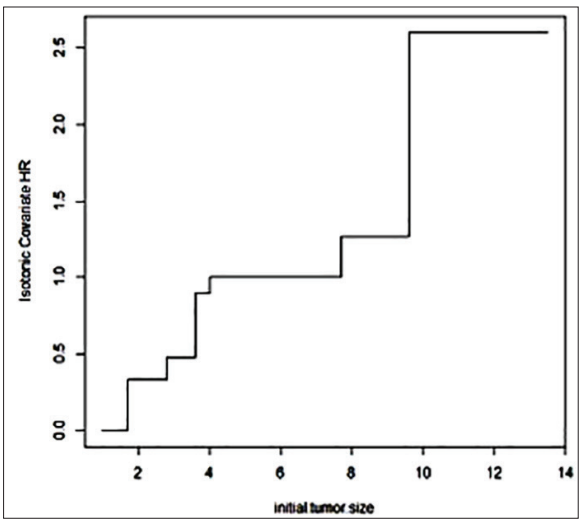

Figure 2. The monotone increasing relationship between tumor size and recurrence

In our country, the application of isotonic models on survival analysis has not been studied. Ancukiewicz et al. first applied isotonic regression to survival analysis in 2003. In their study, they examined the relationship between the continuous covariate and clinical event using isotonic regression analysis. In estimating the course of the disease, they modeled it as a monotone function of the continuous covariate variable. In their modeling, the isotonic model was used to describe the relationship of CD4 with HIV. In the literature related to the isotonic regression model, it is seen that both simulation studies have been carried out in order to develop the model from a theoretical perspective and applications have been made in the field of health. However, most of the studies in the field of health constitute dose-response studies. One reason it cannot be used often for survival analyzes is that the calculations can be quite complex in the censored data type. For this purpose, Yunro Chung facilitated the application of the method on survival analysis with the isoSurv package he realized in the R program in 2021, in order to implement the survival method of the method. We used the partial estimation of isotonic proportional hazard model studied by Chung and colleagues (4,10).

In this study, it is aimed to show the application of the isotonic model on survival analysis by stating this deficiency in the literature. By using the isotonic proportional model, the results are obtained by considering the monotonically increasing or decreasing effect of the covariate variable on the result variable. In our study, the main variables that may affect the risk of breast cancer were modeled and the cox and isotonic proportional hazard models were compared. As a result, it was determined that there was a similarity between the predicted values in the model between the two models. However, for another study or dataset, the result of the two models may differ considerably. Therefore, while evaluating which model should be used while creating the model, it is necessary to examine what kind of effect the covariate variable has on the outcome variable. After the examination, the appropriate model should be established and the results should be interpreted.

As a result, in cases where the covariate variable has a monotone increasing effect on the disease process, the isotonic regression model can be used by considering the relationships between the covariate and the event. Thus, the function that takes into account the monotone relationship between the covariate and the event will be included in the survival analysis. It should be evaluated whether there is a monotone effect of the covariate variable of the models that are rarely used for survival analysis, and it is recommended that these models should be used more in the presence of this relationship.

Conflict-of-interest and financial disclosure

The author declares that she has no conflict of interest to disclose. The author also declares that she did not receive any financial support for the study.

REFERENCES

1. Cox DR. Regression models and life-tables (with Discussion). J R Statist Soc B. 1972;34:187–220.
2. Sertkaya D, Nihal A, Sözer MT. Yaşam çözümlemesinde zamana bağlı açıklayıcı değişkenli Cox regresyon modeli. Ankara Üniversitesi Tıp Fakültesi Mecmuası. 2005;58(4):153-8.
3. Georgia S. The Isotonic regression framework (Estimating and testing under order restrictions). Thesis, LMU München, Mathematik, Informatik and Statistik Department, München; 2003.
4. Chung Y, Ivanova A, Hudgens MG, Fine JP. Partial likelihood estimation of isotonic proportional hazards mod-

els. Biometrika. 2018;105(1):133-48.
5. Ma Y, Lai Y, Lachin JM. Identifying change points in a covariate effect on time-to-event analysis with reduced isotonic regression. PLoS One. 2014;9(12):e113948.
6. Ancukiewicz M, Finkelstein DM, Schoenfeld DA. Modelling the relationship between continuous covariates and clinical events using isotonic regression. Stat Med. 2003;22(20):3151-9.
7. Keam B, Im SA, Park S, et al. Nomogram predicting clinical outcomes in breast cancer patients treated with neoadjuvant chemotherapy. J Cancer Res Clin Oncol. 2011;137:1301–08.
8. Abike F, Zengeroğlu S, Temizkan O, Payaslı A, Tapsız AL. Seröz ve Müsinöz Over Kanserleri İle Ki-67 İlişkisi. Turk Hij Den Biyol Derg. 2010;67(2):79-84.
9. Tekin L, Doğan E. Meme kanserlerinin alt tiplerine göre patolojik özelliklerinin değerlendirilmesi. FÜ Sağ Bil Tıp Derg. 2018;32(3):129-32.
10. Package 'isoSurv' isoSurv: Isotonic Regression on Survival Analysis. 2021.
11. Brunk HD, Barlow RE, Bartholomew DJ, Bremner JM. Statistical inference under order restrictions. (the theory and application of isotonic regression).1st ed. John Wiley & Sons 1972.