

Parametrik Sağlıkım Modellerinin Sağkalım Fonksiyonlarının Tanısal Grafikleri Yardımlıyla Karşılaştırılması

Comparison of the Parametric Survival Models Based on Diagnostic Plots of the Survival Functions

Esin AVCI,^a
Nural BEKİROĞLU^b

^aUygulamalı İstatistik AD,
Giresun Üniversitesi
Fen-Edebiyat Fakültesi,
Giresun

^bBiyoistatistik ve Bilişim AD,
Marmara Üniversitesi Tıp Fakültesi,
İstanbul

Geliş Tarihi/Received: 20.09.2013
Kabul Tarihi/Accepted: 24.01.2014

Yazışma Adresi/Correspondence:
Esin AVCI
Giresun Üniversitesi
Fen-Edebiyat Fakültesi,
Uygulamalı İstatistik AD, Giresun,
TÜRKİYE/TURKEY
esinavci@hotmail.com

ÖZET Amaç: Bu çalışmada, sağlıkım süresinin dağılımının belirlenmesinde kullanılan grafiksel yöntemlerden biri olan sağlıkım fonksiyonunu doğrusallaştıran tanısal grafiği, Meme kanserli hastaların sağlıkım süresinin dağılımının belirlenmesinde kullanılması ve en iyi uyum iyiliğini sağlayan parametrik sağlıkım modelinin belirlenmesinde kullanılması ve test istatistikleri yardımıyla desteklemesi amaçlanmıştır. **Gereç ve Yöntemler:** 458 meme kanseri teşhisi konan hastaların sağlıkım süresine etki eden prognostik faktörleri belirlemek için parametrik (Üstel, Weibull, Log-lojistik ve Log-normal) sağlıkım modelleri kullanılmıştır. Sağlıkım sürelerinin dağılımını belirlemek için grafiksel yöntem olarak sağlıkım fonksiyonlarını doğrusallaştıran tanısal grafiği, en iyi tanımlayan modeli belirlemek için test istatistiği olarak -2LogL ile Akaike Bilgi Kriteri (AIC) kullanılmıştır. **Bulgular:** Sağlıkım fonksiyonunu tanısal grafiğinde en iyi doğrusal görünüm, en yüksek belirleme katsayısına sahip ve en düşük -2LogL ile AIC değerini veren modelin Log-normal olduğu saptanmıştır. Log-normal sağlıkım modele göre meme kanseri sağlıkım süresine etki eden prognostik faktörler olarak; Tümör büyüklüğü, TNM evresi IIA, TNM evresi IIB, Östrojen reseptör durumu ve Aksiller lenf nodu istatistiksel olarak anlamlı bulunmuştur ($p<0,05$). **Sonuç:** Parametrik sağlıkım modelleri için hem grafiksel hem de test istatistiği yöntemi ile yapılan karşılaştırmada, meme kanseri verisi için en iyi ve uygun modelin Log-normal sağlıkım modeli olduğu görülmüştür.

Anahtar Kelimeler: AIC; sağlıkım fonksiyonunun tanısal grafiği; meme kanseri; parametrik sağlıkım modelleri; Log-normal sağlıkım modeli

ABSTRACT Objective: The objective of this study was to use diagnostic plots based on survival function, which is one of the graphical methods to define survival distribution of survival time, to determine the distribution of survival time of patients with breast cancer and to support statistical test that used to find the best fitted model. **Material and Methods:** The data obtained from 458 patients with breast cancer. Parametric (Exponential, Weibull, log-logistic and Log-normal) survival models were used to determine prognostic factors that affect survival time of patients. To describe the best fitted model of survival time, the test statistics -2LogL and AIC and diagnostic plots based on survival functions were used. **Results:** Plots based on the linearization of survival function showed the best linear fit and highest coefficient of determination for Log-normal survival model by ascertaining the smallest -2LogL and AIC value when comparing with other survival models. According to the Log-normal survival models, prognostic factors as tumor size, TNM stage IIA, IIB, estrogen receptor status and axillary nodal status were found statistically significant ($p<0.05$). **Conclusion:** Parametric survival models were compared with both statistical test and graphical methods. According to both methods Log-normal survival model was found to be the best fit model for breast cancer data.

Key Words: AIC; Diagnostic plots of survival function; breast cancer; parametric survival models; Log-normal survival model

Türkiye Klinikleri J Biostat 2014;6(1):33-41

Sağkalım en kısa tanımıyla yaşamda kalma süresidir. Daha genel bir ifade ile belirli bir başlangıç değerinden (doğum, tedavi başlangıcı veya işverenin iş vermesinden) belirli bir olguya ulaşmaya (ölüm, hastalığın

nüks etmesi veya işten çıkarılma) kadar geçen süre olarak tanımlanır. Sağkalım bir süreç olduğu için veriler negatif değer alamaz ayrıca çalışmada yer alan tüm birimler belirli bir olguya erişmemiş olabileceğinden sansürlü verilerden söz edilir. Sağkalım verilerinin analizi ve modellenmesi, 17. yy'da aktüerya ve demografinin gelişmesiyle şekillenmeye başlamıştır. Sağkalım analizi konusundaki ilk önemli gelişme, Kaplan ve Meier'in 1958 yılında yayımlanan "sağkalım eğrisinin tahmini" makalesiyle kendini göstermiştir.¹ Bu çalışmanın ardından, sağkalım süresine etki eden risk faktörlerinin belirlenmesi amacıyla yarı parametrik ve parametrik yöntemler geliştirilmiştir. Yarı parametrik modellerden olan Cox regresyon modeli orantılı hazard varsayımına oldukça duyarlıdır. Parametrik sağkalım modellerinde ise sağkalım süresine ait dağılımın belirlenmiş ve biliniyor olmasını gerekir. Bu varsayımın sağlanması halinde sağkalım süresine etki eden faktörlere ait parametrelerin tahmini gerçekleştirilebilir.² Parametrik sağkalım modeli, anahtar değerlerine daha yakın sonuçlar vermesi (daha küçük standart hata) nedeniyle Cox regresyon modeline göre daha avantajlıdır.³ Ancak parametrik sağkalım modellerinin kullanılmasında en önemli nokta doğru sağkalım dağılımının seçilmesidir.⁴

Grafiksel yöntemler, genel olarak hesaplama gerektirmemesi ve algılama kolaylığı gibi avantajlarından dolayı sıklıkla kullanılan bir yöntemdir. Parametrik sağkalım analizinde belirli bir sağkalım dağılımının doğru olup olmadığını görsel olarak ortaya koymak için grafiksel yöntemlerden yararlanılmaktadır. Bu yöntemler, sağkalım süresine ait olasılık yoğunluk, hazard fonksiyonu ve sağkalım fonksiyonunu ve sağkalım fonksiyonunu doğrusallaştıran tanısal grafikleridir. Sağkalım fonksiyonunu doğrusallaştıran tanısal grafikler, model doğru tanımlandığında doğrusal bir görünüm alması beklenir.

Bu çalışmada, 9 yıl izlemli meme kanserli hastaların modellenmesinde hem (-2logL) ve AIC test istatistikleri hem de Sağkalım fonksiyonunu doğrusallaştıran tanısal grafik yöntemi kullanarak en uygun parametrik sağkalım modelinin belirlenmesi amaçlanmış ve belirlenen modele göre prognostik faktörler tespit edilmiştir.

GEREÇ VE YÖNTEMLER

SAĞKALIM MODELLERİ

Sağlık alanında sağkalım modeli; ölümcül hastalığa yakalanan bir bireye tedavi uygulandıktan sonra, bireyin daha ne kadar yaşayabileceğini ya da hastalığın ne kadar süre sonra tekrarlanabileceğini tahmin etme amacı için geliştirilmiş bir yöntemdir. Sağkalım modeli sadece sağlık alanında değil, mekanik sistemlerinin bozulması, evliliğin boşanma ile sonuçlanması, işsizliğin sona erme süresi ve bir ürünün raf ömrünü tahmin etme gibi birçok değişik alanda da kullanılmaktadır. Sağkalım çalışmalarında yer alan tüm birimler belirli bir olguya erişmemiş olabileceğinden sansürlü veriler de söz konusu olmaktadır.

Sağkalım süresine etki eden prognostik (risk) faktörlerinin belirlenmesi amacıyla geliştirilen ve yarı parametrik modellerden olan Cox regresyon modeli orantılı hazard varsayımına oldukça duyarlıdır.⁵⁻⁷ Oransal hazard modelindeki temel varsayım; hazard oranlarının tüm zaman boyunca sabit olmasıdır. Cox regresyon modeli sağkalım süresinin belirli bir olasılıksal dağılıma sahip olmasını gerektirmediğinden esnek ve yaygın kullanım alanına sahiptir. Cox regresyonda, temel hazard fonksiyonu tanımlanmakta sadece bağımsız değişken(lerin) etkisi modellenmektedir, bu nedenle yarı parametrik bir modeldir.

Parametrik sağkalım modelinde ise, sağkalım süresine ait dağılımın belirlenmiş olmasından dolayı olasılık yoğunluk, hazard ve sağkalım fonksiyonları kesin olarak tanımlanmaktadır. Bu fonksiyonların tanımlanması, tahmin edilen sağkalım eğrisinin tüm veriler hakkında bilgi içermesini ve Bayes yöntemi ile parametre tahmini yapılmasına da olanak sağlamaktadır. Dağılımın belirlenmesi durumunda sağkalıma etki eden faktörlere ait parametrelerin tahmini gerçekleştirilebilir.² Parametrik sağkalım modeli, anahtar değerlerine daha yakın sonuçlar vermesi (daha küçük standart hata) bakımından yarı parametrik sağkalım modellerine göre daha avantajlıdır.³ Parametrik sağkalım modellerinin en önemli özelliği doğru sağkalım dağılımının seçilmesidir,⁴ çünkü ancak bu durumda

parametrik sağkalım modellerinin yarı parametrik sağkalım modeli olan Cox regresyon modeline göre daha etkili parametre tahminlerine sahip oldukları gösterilmiştir.⁸⁻¹⁰

PARAMETRİK SAĞKALIM MODELLERİ

Parametrik sağkalım modelinde, sağkalım süresine etki eden prognostik faktörlerin ($X = (X_1, X_2, \dots, X_p)$) modellenmesinde iki yaklaşım bulunmaktadır.

Birinci yaklaşım, sağkalım süresinin doğal logaritmasını alarak klasik doğrusal regresyon yaklaşımına benzer şekilde modellemektir.

İkinci yaklaşım ise, koşullu hazard fonksiyonunu, prognostik faktörlerin belirlenen bir fonksiyonu ile modellemektir.

Birinci yaklaşım ele alındığında;

$$\ln(t) = \mu + \beta'X + \sigma W \quad (1)$$

$\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ regresyon katsayı vektörü olup ve de hata terimidir. Hata teriminin dağılımı genelde, Log-normal sağkalım modelini veren standart normal dağılım, Weibull modelini veren uçdeğer (extreme value) ya da log-lojistik modelini veren lojistik dağılım gibi dağılımlardan birine sahiptir. Bu yaklaşım için kullanılan parametrik sağkalım modelleri hızlandırılmış başarısızlığa (ölüm gibi) ulaşmış modelleri (accelerated failure time) olarak da adlandırılmaktadır. Ancak bu yaklaşım hata terimi dağılımları için sıfır ortalama ve küçük varyansa sahip olma varsayımı içerdiğinden kullanım alanları oldukça sınırlıdır. Bu durumda ikinci

yaklaşımın kullanılması daha uygun olabilmektedir.

Prognostik faktörlerin sağkalım fonksiyonu üzerine olan etkilerini modellemek için kullanılan ikinci yaklaşımda koşullu hazard fonksiyonu modellenir.¹¹ Buna göre koşullu hazard fonksiyonu;

$$h(t/x) = h_0(t)\varphi(\beta'x) \quad (2)$$

şeklinde yazılır.¹²⁻¹⁴ Burada $\varphi(\beta'x)$ prognostik faktörlerin negatif olmayan bir fonksiyonudur. Birçok uygulamada $\varphi(\beta'x) = \exp(\beta'x)$ olarak seçilir. Parametrik sağkalım modeli için, sağkalım süresi hangi dağılıma uygunsa o dağılıma ait temel hazard fonksiyonu kullanılır.

En çok kullanılan parametrik sağkalım modellerinden Üstel, Weibull, Log-lojistik ve Log-normal dağılımları için temel hazard fonksiyonları ($h_0(t)$) ile olasılık yoğunluk fonksiyonları ($f(t)$) ve sağkalım fonksiyonları ($S(t)$) Tablo 1'de verilmiştir.

Bir veri seti için belirli bir dağılımın uygunluğunun test edilmesinde veya birçok parametrik sağkalım modeli arasından en iyi modeli değerlendirmede genel olarak görsel yöntem olan grafiksel yöntem ve test istatistiği yöntemi olan -2LogL ve AIC gibi uyum iyiliği testlerinden yararlanır.

GRAFİKSEL YÖNTEM

Grafiksel yöntem sağkalım fonksiyonunun tahminine dayanmaktadır. Parametrik modele uyumluluğu kontrol etmek için kullanılan görsel bir yöntemdir. Bu grafikler; sağkalım süresine ait olasılık yoğun-

TABLO 1: Parametrik dağılımlar için temel hazard fonksiyonları, olasılık yoğunluk fonksiyonları ve sağkalım fonksiyonları.

Dağılım	Parametre	$h_0(t)$	$f(t)$	$S(t)$
Üstel	$\lambda > 0$	λ	$\lambda \exp(-\lambda t)$	$\exp(-\lambda t)$
Weibull	$\lambda > 0$ $\alpha > 0$	$\lambda \alpha t^{\alpha-1}$	$\lambda \alpha t^{\alpha-1} \exp(-\lambda t^\alpha)$	$\exp(-\lambda t^\alpha)$
Log-lojistik	$\lambda > 0$ $\alpha > 0$	$\frac{(\alpha t^{\alpha-1} \lambda)}{(1 + \lambda t^\alpha)}$	$\frac{(\alpha t^{\alpha-1} \lambda)}{(1 + \lambda t^\alpha)^2}$	$\frac{1}{(1 + \lambda t^\alpha)}$
Log-normal	$\alpha > 0$	$f(t) / S(t)$	$\exp \left[\frac{-1/2 ((\ln t - \mu) / \sigma)^2}{t (2\pi)^{1/2} \sigma} \right]$	$1 - \Phi \left(\frac{(\ln t - \mu)}{\sigma} \right)$

μ : Ortalama, σ : varyans

Φ : standart normal bir değişkenin birikimli dağılım fonksiyonu

TABLO 2: Sağkalım dağılımlarına göre zaman ölçeği ve doğrusallaştırılmış sağkalım fonksiyonları.

Sağkalım Dağılımları	Doğrusallaştırılmış Sağkalım	
	Zaman Ölçeği	Fonksiyonu
Üstel	T	$-\log S(t)$
Weibull	Logt	$\log(-\log S(t))$
Log-lojistik	Logt	$\log(S(t)/(1-S(t)))$
Log-normal	Logt	$\Phi^{-1}(1-S(t))$

luk, hazard fonksiyonu ve sağkalım fonksiyonunu grafiği ile sağkalım fonksiyonunu doğrusallaştıran tanısal grafiklerinden oluşmaktadır.

Genel olarak sağkalım fonksiyonunun doğrusallaştırılmasına dayanan grafikler, sağkalım dağılımı hakkında bilgi vermektedir.¹⁵ Bu tür grafikler, varsayılan modelin doğruluğu varsayımı altında doğrusal bir görünüme sahip olması gerekir. Bu çalışmada ele alınan sağkalım dağılımları için beklenen yaklaşık doğrusallaştırılmış sağkalım fonksiyonları Tablo 2’de özetlenmiştir. Bu grafiklerde X ekseninde *Zaman Ölçeği* ve Y ekseninde *Doğrusallaştırılmış Sağkalım fonksiyonu* yer almaktadır. Doğrusallık gözle ve/veya X ile Y arasındaki uyumun göstergesi olan belirleme katsayısı (R^2) ile belirlenebilir.¹⁵

Burada t , $S(t)$ ve Φ sırasıyla sağkalım zamanı, sağkalım fonksiyonu ve standart normal bir değişkenin birikimli dağılım fonksiyonudur.

Sansürlü veriler için tanımlanan grafiksel yöntem parametrik modellerin karşılaştırılmasında en çok kullanılan yöntemlerdendir.¹⁶

Parametrik sağkalım modelinde uyum iyiliği Cox-Snell artıkları kullanılarak incelenmektedir. $\hat{S}(r)$, $S(r)$ ’nin Kaplan-Meier tahmin edicisidir. Cox-Snell artığı r_i ve Cox-Snell artığının sağkalım fonksiyonunun Kaplan-Meier tahmin edicisi $\hat{S}(r)$ olmak üzere r_i ’nin $-\log(\hat{S}(r))$ ’ye karşı grafiği bir eğim ve sıfır kesişim ile düz bir doğru ise parametrik sağkalım modelinin ilgilenilen veri kümesi için uygun olduğu sonucuna ulaşılır.¹⁷

TEST İSTATİSTİKLERİ

Parametrik sağkalım modellerinin uyumluluğunun test etmek ve en iyi uyumu veren modele karar

vermek için birçok test istatistikleri bulunmaktadır.³ Bunlardan en çok; Sapma İstatistiği ($-2\log L$), Akaike Bilgi Kriteri (Akaike’s information criterion, AIC) ve Bayesci Bilgi Kriteri (BIC) kullanılmaktadır. En küçük sapma istatistiğine, AIC veya BIC değerine sahip model en uygun model olarak kabul edilir. Bu çalışmada modelin uyum iyiliğini test edilmesi için yaygın olarak kullanılan $-2\log L$ ve Akaike Bilgi Kriteri (AIC)’nden yararlanılmıştır. En küçük değere sahip $-2\log L$ istatistiğine ya da AIC değerine sahip model, en iyi uyum sağlayan model olarak belirlenmiştir.¹²

Parametrik sağkalım modelleri için Akaike bilgi kriteri;

$$AIC = -2\log L + 2(p + 2 + k) \quad (3)$$

eşitliği ile verilir. Burada p modeldeki bilinmeyen parametrelerin sayısıdır. AIC değerleri hesaplanırken üstel model için $k=0$, Weibull, Log-lojistik ve Log-normal modelleri için $k=1$ olarak alınmaktadır.¹²

Analizler S-Plus 2000 istatistik programı kullanılarak yapılmıştır.

BULGULAR

Çalışmada kullanılan veriler, 1993-2002 yılları arasında İstanbul Üniversitesi Tıp Fakültesi Meme Cerrahi Kliniğinde tedavi gören 458 meme kanseri hastasından elde edilmiştir. Veriler sağdan sansürlü olup, sağkalım süreleri ay olarak hesaplanmıştır.

Meme kanseri olan hastalara ait prognostik faktörlerin dağılımı Tablo 3’te özetlenmiştir.

En uygun olan parametrik sağkalım modeline karar vermek için sağkalım süresinin dağılımı belirlenmelidir. Bu amaçla meme kanserli hastaların sağkalım süresine ait sağkalım fonksiyonlarını doğrusallaştıran tanısal grafiklerinden ve belirleme katsayısından (R^2) yararlanılmıştır. Üstel (a), Weibull (b), Log-lojistik(c) ve Log-normal (d) sağkalım modellerine ait tanısal grafikleri sırasıyla Şekil 1’de verilmiştir. Grafikler incelendiğinde düz bir doğru görünümüne ve en büyük R^2 değerine sahip dağılımın Log-normal olduğu görülmektedir.

Test istatistiklerinden yararlanarak, yukarıda grafiksel olarak belirlenen saptamanın doğruluğu

TABLO 3: Meme kanserli hastalara ilişkin değişkenlerin özellikleri.

Bağımsız Değişkenler	Hasta Sayısı	Yüzde	Ölen Hasta Sayısı	Sağkalan Hasta Sayısı
Yaş (\leq)	232	51	44	188
Yaş (>50)	226	49	25	201
Aile Hikayesi (Yok)	370	80	58	312
Aile Hikayesi (Meme kanseri)	55	12	7	48
Aile Hikayesi (Diğer Kanseri)	33	8	4	29
Menapoz Durumu (yok)	329	71	55	274
Menapoz Durumu (var)	129	29	14	115
Tümör Büyüklüğü (≤ 24)	243	53	24	219
Tümör Büyüklüğü (>24)	215	47	45	170
Histolojik Tip (Diğer)	95	21	21	74
Histolojik Tip (Duktal)	363	79	48	315
TNM Evresi (0)	123	26	18	105
TNM Evresi (IIA)	188	41	24	164
TNM Evresi (IIB)	147	33	27	120
Östrojen Reseptör Durumu (Negatif)	176	38	48	128
Östrojen Reseptör Durumu (pozitif)	282	62	21	261
Progesteron Reseptör Durumu (Negatif)	248	54	54	194
Progesteron Reseptör Durumu (pozitif)	210	46	15	195
Tedavi Türü (MKC)	133	29	23	110
Tedavi Türü (MRM)	325	71	46	279
Aksiller Lenf Nodu Tutulumu (Negatif)	238	52	52	186
Aksiller Lenf Nodu Tutulumu (Pozitif)	220	48	17	203

* MKC: Meme koruyucu cerrahi, MRM: Modifiye Radikal mastektomi.

ğunu ve meme kanseri hastaların sağkalım süresine etki eden prognostik faktörleri belirlemek için 10 adet açıklayıcı değişken analize alınmış ve 4 farklı parametrik sağkalım modeli uygulanarak $-2\log L$ ve AIC değerleri elde edilmiştir. Sonuçlar Tablo 4'te verilmiştir.

Tablo 4'te yer alan sonuçlar incelendiğinde sağkalım dağılımlarının $-2\log L$ ve AIC değerleri içinde en küçük değer Log-normal dağılıma ait olduğu görülmektedir.

Parametrik sağkalım modellerinin uyum iyiliğinin test edilmesinde Cox-Snell artıkları kullanılmaktadır. Meme kanseri verisi için incelenen her bir dağılıma ait Cox-Snell artıkları elde edilmiş ve Cox-Snell grafikleri Şekil 2'de verilmiştir.

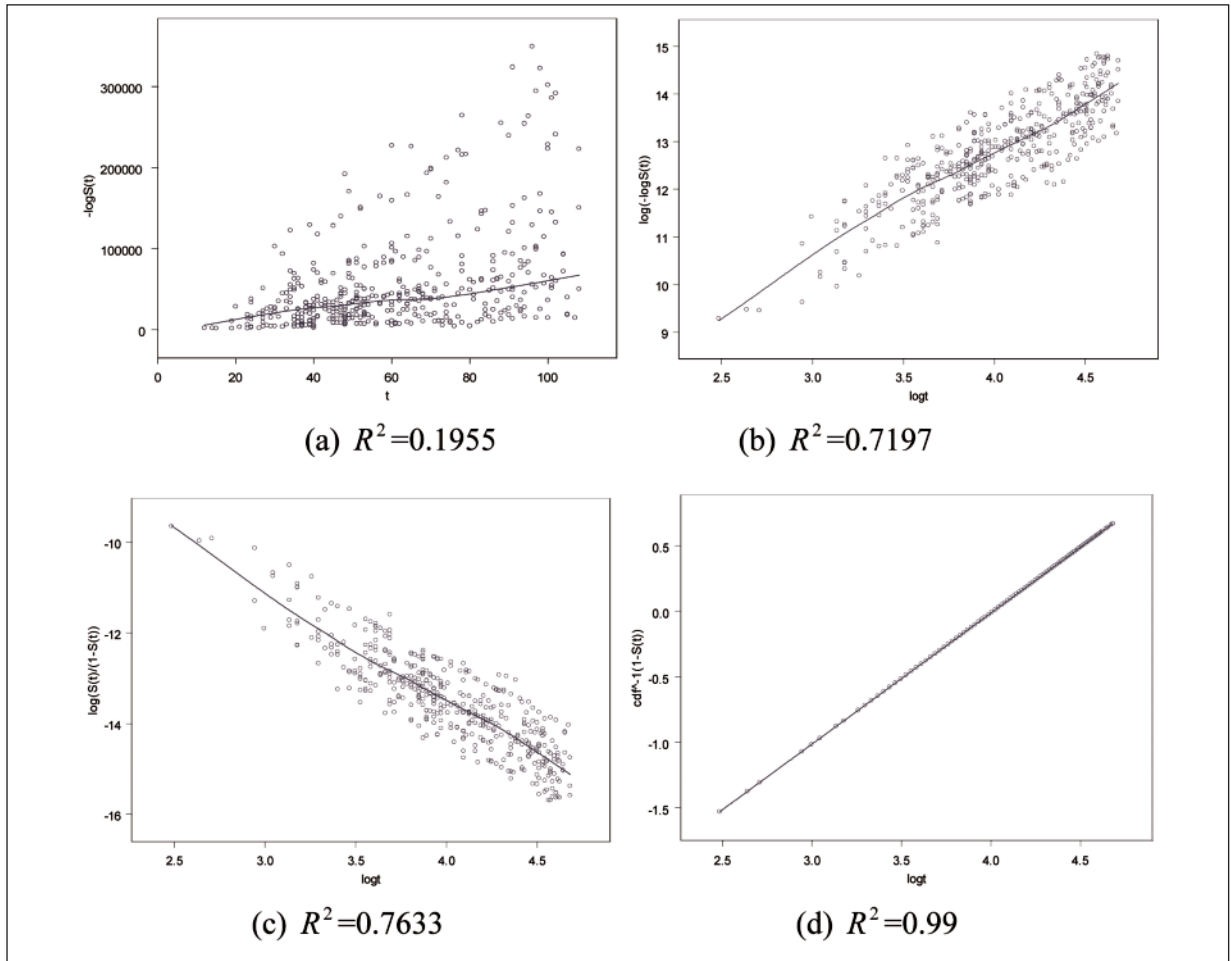
Cox-Snell artık grafiklerine göre Log-normal sağkalım modelinin veri kümesi için uygun olduğu söylenebilir.

Sağkalım süresine uygun dağılımın ($\mu=3.7147$, $\sigma=0.50728$) parametrelili Log-normal olduğu sap-

tandıktan sonra, etki eden prognostik faktörlere göre oluşturulan sağkalım modelinin sonuçları Tablo 5'te verilmiştir.

Tablo 5 incelendiğinde Tümör büyüklüğü ($p=0.001$), TNM evresi IIA ($p=0.003$), TNM evresi IIB ($p=0.003$), Östrojen reseptör durumu negatif olanların ($p=0.0075$) ve Aksiller lenf nodu pozitif olanların ($p<0.0001$) %95 güvenle istatistiksel olarak önemli olduğu ve sağkalım süresine etki ettiği belirlenmiştir.

Sağkalım analizinde prognostik faktörlerin düzey sayısı ikiden fazla olması durumunda düzeylerden biri referans kategori olarak alınmakta ve faktörlerin yorumlanması buna göre yapılmaktadır. Prognostik faktörlerin düzeylerine göre oransal risklerinin belirlenmesi amacıyla her bir prognostik faktöre ait parametre tahmin değerinin üsteli alınır. TNM evresi IIA olan hastaların TNM evresi (0) olanlara göre yaklaşık 1,87 kat, TNM evresi IIB olan hastaların TNM evresi (0) olanlara göre yaklaşık 2 kat,



ŞEKİL 1: Dört sağkalım modeli için tanısal grafikleri.

Östrojen reseptör durumu negatif olanların Östrojen reseptör durumu pozitif olanlara göre 1.605 kat ve Aksiller Lenf Nodu Tutulumu Pozitif olanların Aksiller Lenf Nodu Tutulumu negatif olanlara göre 2 kat sağkalım süresi daha azdır.

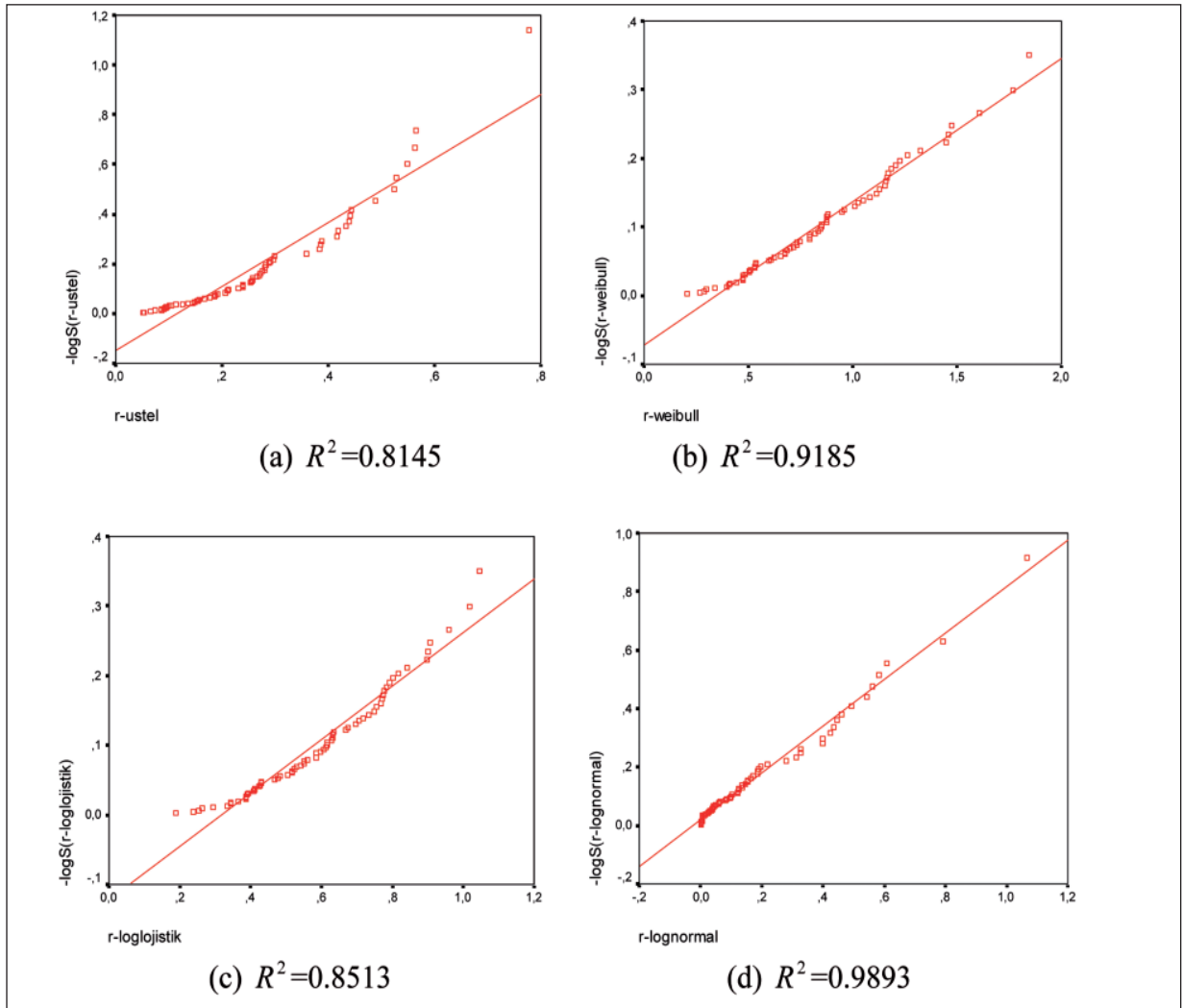
TARTIŞMA VE SONUÇ

Tıbbi çalışmalarda parametrik sağkalım modeli sık kullanılmaktadır. Kurulan parametrik sağkalım modelinin doğruluğu ancak doğru tanımlanan sağkalım süresinin dağılımı ile gerçekleşir. Doğru dağılımın belirlenmesinde test istatistiklerinin yanı sıra sağkalım süresine ait olasılık yoğunluk, hazard fonksiyonu ve sağkalım fonksiyonunu ve sağkalım fonksiyonunu doğrusallaştıran tanısal grafiklerinden de yararlanılabilir. Parametrik sağkalım modelinde sağkalım süresinin uygun dağılımının belirlenmesi ile ilgili birçok çalışma bulunmaktadır.

TABLO 4: Parametrik sağkalım modelleri için $-2\log L$ ve AIC değerleri.

Kriter	Üstel	Weibull	Log-lojistik	Log-normal
$-2\log L$	894,8	870,4	867	864,2
AIC	918,8	896,4	893	890,2

dır. Raman ve ark.nın 2012 yılının Şubat ayında yaptıkları çalışmada 522 meme kanserli hastanın sağkalım süresi üzerinde evre, düşük sayısı, çocuk sayısı, kemoterapi, radyoterapi, ailedeki kanser hikayesi ve menepoz durumu değişkenlerinin etkileri incelenmiş ve bu amaçla üç farklı dağılım (Weibull, Üstel ve Log-normal) uygulanmıştır. Uygun dağılım en küçük $-2\log L$ değerine bakılarak karar verilmiştir. İncelenen prognostik faktörlerden; evre ve kemoterapi değişkenlerin etkisinin



ŞEKİL 2: Dört sağkalım modeli için Cox-Snell artık grafikleri.

TABLO 5: Log-normal sağkalım modelinin sonuçları.

Bağımsız Değişkenler	β	σ_β	Z	P	Exp(β)	Exp(β) Alt sınır	Exp(β) Üst sınır
Yaş (Yıl)	-0.0029	0.0060	-0.4739	0.6360	0.9971	0.9854	1.0089
Aile Hikayesi (Meme kanseri)	-0.0377	0.2150	-0.1754	0.8610	0.9630	0.6319	1.4677
Aile Hikayesi (Diğer kanser)	0.3860	0.2937	1.3145	0.1890	1.4711	0.8272	2.6160
Menopoz Durumu	0.2267	0.1849	1.2259	0.2200	1.2545	0.8731	1.8024
Tümör Büyüklüğü	-0.0253	0.0077	-3.2726	0.001	0.9750	0.9604	0.9898
TNM Evresi IIA	0.6284	0.2140	2.9367	0.003	1.8746	1.2324	2.8515
TNM Evresi IIB	0.7648	0.2603	2.9381	0.003	2.1486	1.2900	3.5787
Histolojik Tip (Duktal)	-0.0047	0.1621	-0.0288	0.977	0.9953	0.7244	1.3675
Östrojen Reseptör Durumu (Negatif)	0.4731	0.1770	2.6737	0.0075	1.605	1.1345	2.2705
Progesteron Reseptör Durumu (Negatif)	0.2158	0.1732	1.2462	0.2130	1.2408	0.8837	1.7424
Tedavi TÜRÜ (MKC,MRM)	0.0358	0.1618	0.2210	0.8250	1.0364	0.7548	1.4232
Aksiller Lenf Nodu Tutulumu (Pozitif)	0.8022	0.1869	4.2922	0.000	2.2304	1.5463	3.2172

Z: Wald istatistiği

önemli olduğu saptanmıştır.¹⁸ 2010 yılında Hayat ve ark.nın yaptığı çalışmada ise, meme kanserine yakalanan hastaların sağkalım süresinin dağılımının belirlenmesinde sadece yaş prognostik faktörü ele alınarak beş sağkalım modeli (Log-normal, Log-lojistik, Weibull, Gamma ve Gompertz) hazard fonksiyon grafikleri ve AIC test istatistiği ile karşılaştırılmıştır. Bu çalışmada yaş prognostik faktörü önemsiz bulunmuş ancak sağkalım süresinin Gompertz dağıldığı hem hazard fonksiyon grafiği yardımıyla hem de AIC test istatistiğinin yararlanılarak saptanmıştır.¹⁹ Royston 2001 yılında 686 meme kanserli hastanın sağkalım süresine etki eden prognostik faktörler olarak yaş, menepoz durumu, tümör büyüklüğü, tümör evresi, lenf nodu tutulumu, östrojen reseptör durumu ve progesteron reseptör durumu ele alınmış ve 6 parametrik model (Weibull, Üstel ve Log-normal, Log-lojistik, Gompertz ve Gamma) den en küçük deviance değerine sahip olan Log-normal sağkalım model ile Cox modelin karşılaştırmasını yapmış ve Log-normal sağkalım modelin daha dayanıklı sonuçlar verdiğini belirtmiştir.²⁰ 2006 yılında ise Chapman ve ark. 415 meme kanserli hastaların sağkalım süresine yaş, ağırlık, tümör büyüklüğü, histolojik tip, östrojen reseptör durumu, progesteron reseptör durumu, tedavi türü, lenf nodu tutulumu prognostik faktörlerin etkilerinin saptanmasında Log-normal sağkalım modeli ile Cox modelini karşılaştırmış ve Log-normal sağkalım modelinin daha avantajlı olduğunu tespit etmiştir.^{21,22}

Grafiksel yöntemler görselliği sağlayarak birçok dağılım içerisinden uygun olabilecek dağılım sayısını azaltarak, seçim yapmayı kolaylaştırmaktadır. Sağkalım süresinin dağılımının belirlenmesinde sağkalım fonksiyonunu doğrusallaştıran tanısal ile ilgili Meme kanseri verileri üzerinde herhangi bir çalışma bulunmamaktadır. Bu nedenle bu çalışmanın önemi ortaya çıkmakla birlikte, bebek ve çocuk ölümlerinin, sağkalım fonksiyonunu doğrusallaştıran tanısal grafik yöntemi ile Weibull dağılımına sahip olduğu ve kolay bir biçimde saptandığını belirten sadece bir çalışma bulunmaktadır.²³ Ayrıca rüzgar güllerinin kuşların uçuşlarını

etkilemesi üzerine yapılan başka çalışmada Üstel, Weibull, Log-lojistik ve Log-normal sağkalım modelleri ele alınmış ve sağkalım süresi için en iyi model sağkalım fonksiyonlarını doğrusallaştıran tanısal grafikler ve test istatistikleri ile saptanmaya çalışılmıştır.²⁴ Bizim çalışmamızda olduğu gibi bu çalışmanın sonucunda sağkalım süresi için uygun dağılımın bulunmasında, sağkalım fonksiyonunu doğrusallaştıran tanısal grafikleri ile test istatistiklerinin sonuçlarının örtüştüğü saptanmıştır.

Çalışmamızda; Üstel, Weibull, Log-lojistik ve Log-normal sağkalım modelleri meme kanseri ve-rişi için karşılaştırıldı. Sağkalım süresine en uygun dağılımın belirlenmesi için görsel yöntem olarak sağkalım fonksiyonunu doğrusallaştıran tanısal grafiğinden yararlanıldı. Sağkalım süresine etki eden prognostik faktörleri kapsayan en uygun modelin belirlenmesinde uyum iyiliği grafiklerinden Cox-Snell artık grafiği ile test istatistiği yöntemi olarak (-2LogL) ve AIC kullanıldı.

Tüm modeller için elde edilen tanısal grafiklerinden en iyi doğrusal görünümü, en yüksek belirleme katsayısı, en küçük -2LogL ve AIC değeri ile en doğrusal Cox-Snell artık grafiğini veren modelin Log-normal sağkalım model olduğu saptanmıştır.

Meme kanseri tanısı konan hastaların sağkalım sürelerinin Log normal dağıldığı belirlendiğinden parametrik sağkalım modeli kullanılmıştır. Sağkalım sürelerine etki eden prognostik faktörler; yaş, aile hikayesi, menepoz durumu, tümör büyüklüğü, TNM evresi, histolojik tip, östrojen reseptör durumu, progesteron reseptör durumu, tedavi türü (radyoterapi ve hormon terapisi) ve aksiller lenf nodu tutulumu olarak belirlenmiştir. Bu prognostik faktörlerden Tümör büyüklüğü, TNM evresi IIA, TNM evresi IIB, Östrojen reseptör durumu negatif olanların ve Aksiller lenf nodu pozitif olanların %95 güvenle istatistiksel olarak önemli olduğu ve sağkalım süresine etki ettiği belirlenmiştir.

Parametrik sağkalım modellerinde en uygun modelin belirlenmesinde test istatistiğinin yanı sıra tanısal grafiklerinden de yararlanılması daha güvenilir sonuçlar sağlamaktadır.

KAYNAKLAR

1. Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53(282):457-81.
2. Feigl P, Zelen M. Estimation of exponential survival probabilities with concomitant information. *Biometrics* 1965;21(4):826-38.
3. Collet D. *Modelling Survival Data in Medical Research*. 2nd ed. Florida: CRC Press; 2003. p.1-408.
4. Block AD, Lemis LM. Parametric model discrimination for heavily censored survival data. *IEEE Transactions on Reliability* 2008;57(2): 248-59.
5. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 1972;34(2):187-220.
6. Kleinbaum DG, Klein M. Parametric survival models. In: Gail M, Krickeberg K, Samet J, Tsiatis A, Wang W, eds. *Survival Analysis-A Self-Learning Text*. 2nd ed. New York: Springer; 2005. p.260-312.
7. Grambsch PM, Therneau T. Proportional hazards test and diagnostics based on weighted residuals. *Biometrika* 1994;81(3):515-26.
8. Efron B. The efficiency of Cox's likelihood function for censored data. *J Am Stat Assoc* 1977;72(359):557-65.
9. Nardi A, Schemper M. Comparing Cox and parametric models in clinical studies. *Stat Med* 2003;22(23):3597-610.
10. Oakes D. The asymptotic information in censored survival data. *Biometrika* 1977;64(3): 441-8.
11. Ata N, Karasoy D, Sözer MT. [Parametric and nonparametric survival models for nonproportional hazards]. *İstatistikçiler Dergisi* 2008; 1(3):125-34.
12. Klein JP, Moeschberger ML. *Survival Analysis Techniques for Censored and Truncated Data*. 2nd ed. New York: Springer-Verlag; 1997. p.1-502.
13. Lee ET, Wang JW. Some well-known parametric survival distribution and their application. *Statistical Methods for Survival Data Analysis*. 3rd ed. Hoboken, New Jersey: Wiley & Sons, Inc; 2003. p.134-62.
14. London D. *Survival Models and Their Estimation*. 3rd ed. Winsted, Conn: ACTEX Publications; 1997. p.177-215.
15. Lawless JF. *Statistical Models and Methods for Lifetime Data*. 2nd ed. Hoboken: NJ Wiley;2003.p.1-630.
16. Cox DR, Oakes D, *Analysis of survival data*. London: Chapman & Hall; 1998. p. 13-24.
17. Lee ET, Wang JW, *Statistical Methods for Survival Data Analysis*. 3rd ed. Hoboken, New Jersey: John Wiley & Sons Inc; 2003. p.1-534.
18. Raman TT, Venkatesan P. Accelerated failure time frailty model in survival analysis. *International Journal of Science and Technology IJST* 2012;2(2):65-9.
19. Aktürk Hayat E, Suner A, Uyar B, Dursun Ö, Orman MN, Kitapçioğlu G. Comparison of five survival models: breast cancer registry data from Ege University Cancer Research Center. *Türkiye Klinikleri J Med Sci*, 2010;30(5):1665-74.
20. Royston P. The Lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors. *Statistica Neerlandica* 2001;55(1):89-104.
21. Chapman JA, Lickley HL, Trudeau ME, Hanna WM, Kahn HJ, Murray D, et al. Ascertaining prognosis for breast cancer in node-negative patients with innovative survival analysis. *Breast J* 2006;12(1):37-47.
22. Chapman JW, Meng D, Shepherd L, Parulekar W, Ingle JN, Muss HB, et al. Competing causes of death from a randomized trial of extended adjuvant endocrine therapy for breast cancer. *J Natl Cancer Inst* 2008;100(4): 252-60.
23. Luther NY. A simple linearization procedure for local fit of the Weibull survival distribution: a note on an article by Choe. *Asian Pac Cens Forum* 1982;9(2):5-8.
24. Bispo R, Bernardino J, Marques TA, Pestana D. Discrimination between parametric survival models for removal times of bird carcasses in scavenger removal trials at wind turbines sites. In: Lita da Silva J, Caeiro F, Natário I, Braumann CA, Esquível ML, João João Tiago Mexia JJ, eds. *Advances in Regression, Survival Analysis, Extreme Values, Markov Processes and Other Statistical Applications Studies in Theoretical and Applied Statistics*. 1st ed. Berlin Heidelberg: Springer Science & Business; 2013. p. 65-72.