

# Grup Ardışık Test Yöntemlerinin Sağkalım Analizinde Uygulanması ve Harcama Fonksiyonlarının Güç Analizi

## APPLICATION OF GROUP SEQUENTIAL TEST METHODS IN SURVIVAL ANALYSIS AND POWER ANALYSIS OF SPENDING FUNCTION

Dr. Yüksel TERZİ,<sup>a</sup> Münevvere SARI,<sup>a</sup> Ahmet SABRİ ÖĞÜTLÜ,<sup>a</sup> Dr. Mehmet Ali CENGİZ<sup>b</sup>

<sup>a</sup>İstatistik Bölümü, Afyonkarahisar Kocatepe Üniversitesi, Fen-Edebiyat Fakültesi, AFYONKARAHİSAR

<sup>b</sup>İstatistik Bölümü, Ondokuz Mayıs Üniversitesi, Fen-Edebiyat Fakültesi, SAMSUN

### Özet

**Amaç:** Grup ardışık test yöntemlerini sağkalım analizi çalışmalarında uygulamak.

**Gereç ve Yöntemler:** Çalışmada 224 akciğer kanseri tanısı alan iki grup (KHAK ve KHDAK) hasta prospektif olarak 1060 gün takip edildi. Akciğer kanseri teşhisi konan hastalar grup ardışık test yöntemleri ile adım adım incelendi ve her adımda iki kanser grubu hastaların sağkalım süreleri karşılaştırıldı. Ayrıca farklı gözlem sayıları ve farklı sağkalım oranları için 5 harcama fonksiyonunun güç analizlerinin karşılaştırmak için simülasyon çalışması yapıldı.

**Bulgular:** İki farklı akciğer kanseri olan hastalar her adımda (265 günde bir) sağkalım süreleri bakımından karşılaştırıldı. Grup ardışık test yöntemleri kullanılarak hangi adımda çalışmanın sonlandırılacağına karar verildi. Ayrıca güç analizi yapılarak grup ardışık test yöntemlerinde kullanılan harcama fonksiyonlarından hangisinin daha iyi olduğuna bakıldı. Grup ardışık test yöntemlerinde kullanılan harcama fonksiyonlarından en yüksek güç analizi sonucunu O'Brien&Fleming verdi.

**Sonuç:** Klinik çalışmalarda iki tedavi yöntemi/grubu karşılaştırılıyor, çalışmanın sonu beklenmeden hangi tedavi yönteminin/grubunun daha iyi sonuç verdiği grup ardışık test yöntemleri ile tespit edilebilir. Böylece belli bir aşamadan sonra hastalar diğer tedavi yöntemine/grubuna yönlendirilerek, hastaların sağkalım süreleri uzatılabilir.

**Anahtar Kelimeler:** Sağkalım analizi; güç

### Abstract

**Objective:** Application of group sequential test methods in survival analysis.

**Material and Methods:** In the study, two groups of 224 patients diagnosed as lung cancer (KHAK and KHDAK) have been followed prospectively for 1060 days. These patients have been examined using group sequential test methods step by step and two groups' survival times have been compared for each step. In addition considering varied observation numbers and survival rates a simulation study is done for comparing 5 spending functions' power analysis.

**Results:** Patients of two different types of lung cancer have been compared with respect to their survival times for each step (every 265 days). Determination of the step in which the study terminates has been fulfilled with group sequential test methods. In addition, spending functions used in these methods have been tested for superiority using power analysis. Of all the spending functions O'Brien Fleming produced the highest power analysis score.

**Conclusion:** In the case of two treatment methods/groups comparisons, without waiting for end of the study, which treatment methods/groups to yield better could be determined with group sequential test methods. Hence, patients' survival times could be extended directing patients to other treatment methods/groups after a particular step.

**Key Words:** Survival analysis; power

Türkiye Klinikleri J Med Sci 2007, 27:846-852

Geliş Tarihi/Received: 04.06.2007 Kabul Tarihi/Accepted: 23.08.2007

20-24 Mayıs 2007'de Antalya'da düzenlenen 5. İstatistik Kongresi'nde sözlü bildiri olarak sunulmuştur.

**Yazışma Adresi/Correspondence:** Dr. Yüksel TERZİ  
Afyonkarahisar Kocatepe Üniversitesi,  
Fen Edebiyat Fakültesi,  
İstatistik Bölümü, AFYONKARAHİSAR  
yukselet@aku.edu.tr

Copyright © 2007 by Türkiye Klinikleri

İstatistikte amaç örnek veri toplamak ve bu verilerden hareketle anakütle hakkında çıkarımlar yapmaktır. Bunun için sabit (belli) sayıda gözlemler toplanır. Bazen gerekli verinin tamamını toplamak zor veya zaman açısından uygun olmayabilir. Bu özellikle klinik denemelerde önemlidir. Eğer 2 tedavi ya da 2 ilaç türünü karşı-

laştırmak isteniyorsa, çalışma sonunda hangi tedavi yönteminin ya da hangi ilacın daha iyi sonuç verdiği karar verilebilir. Ancak daha iyi sonucu tespit etmek için çalışmanın sonu beklendiğinde, bazı hastalarda istenmeyen sonuçlar (ölüm gibi) görülebilir. Bunun için çalışma bitmeden daha erken bir aşamada hangi tedavi yönteminin ya da ilacın daha iyi olduğu tespit edilerek, hastalar o tedavi ya da ilaca yönlendirilebilir. Böylece daha sağlıklı bir sonuca ulaşılmış olur.

Ayrıca klinik denemeler çok masraflı olabileceği için erken sonlanmanın yani çalışmanın sonunu beklemeden belli bir aşamada karar vermenin maddi ve zaman yönünden faydaları da bulunmaktadır. İşte böyle durumlarda bireyleri sürekli izleme veya ardışık olarak takip etme ihtiyacı ortaya çıkmaktadır.

Sağkalım analizi çalışmaları, son yıllarda oldukça sık kullanılmaktadır. Bu çalışmalar daha çok ölüm ile sonuçlanan (kanser gibi) hastalıklar üzerinde uygulanmaktadır. İnsanlar üzerinde yapılan bu tür çalışmalarda gerçek anlamda sağlıklı veri elde etmek zor ve zaman almaktadır. Bu tür çalışmalarda veri elde etmenin zaman, maliyet, çalışma süresince olabilecek aksaklıklardan dolayı ardışık yöntemlerden yararlanılmıştır. Ardışık yöntemler (birimler ardışık olarak gelmesi) çalışmanın erken durdurulmasına imkan sağlamasından dolayı kullanışlıdır.

Ardışık yöntemlerde çalışma sonuçlarının sürekli değişmesinden dolayı, grup ardışık yöntemleri geliştirilmiştir. Grup ardışık test yöntemleri üzerinde ilk çalışmaları Armitage ve ark. yapmışlardır. Daha sonraları ise Pocock, O'Brien&Fleming bu çalışmaları geliştirmişlerdir.<sup>1,2</sup> Lan&DeMets ise grup ardışık testinin alfa maliyet fonksiyonu yaklaşımı ile ilgili çalışmalar yapmışlardır.<sup>3</sup> Rebussion ve ark. ise farklı grup tasarımları için örneklem genişliği ve güç analizi hesabı için metotlar geliştirmişlerdir.<sup>4</sup> Bu metotlarla iki sağkalım eğrisini karşılaştırmak için kullanılan grup ardışık düzenleri için örneklem genişliğini ve testin gücü hesaplanabilmektedir. Slud&Wei ve Lan&DeMets grup büyüklüklerinin farklı olduğu durumlarda nasıl bir yöntem izleneceği üzerine çalışmalar yapmışlardır.<sup>5</sup>

Bu çalışmada amaç, grup ardışık test yöntemlerini sağkalım analizinde gerçek verilerle uygulamasını yapmaktır. Ayrıca farklı gözlem sayıları ve değişik sağkalım oranları için 5 harcama fonksiyonunun güç analizlerini bulmaktır.

## Gereç ve Yöntemler

Grup ardışık test yöntemlerinin sağkalım analizinde uygulaması için prospektif olarak yapılan bir çalışmada, akciğer kanseri tanısı konan toplam 224 hasta araştırıldı. Hastaların takip süreleri gün olarak hesaplanmıştır. Bu hastaların %94'ü erkek, %6'sı kadından oluşmaktadır. Hastaların ortalama yaşı  $60.63 \pm 9.49$  olarak bulunmuştur. Hastaların en sık başvuru nedenleri; öksürük (%19), nefes darlığı (%16), göğüs ağrısı (%14), göğüs dışı ağrılar (%13), hemoptizi (%11), akciğer grafi anormallidir (%10). Çalışmada kullanılan veriler sağdan sansürlü veriler olup hastalar, küçük hücreli akciğer kanseri (KHAK) ve küçük hücre dışı akciğer kanseri (KHDAK) olmak üzere 2 gruptan oluşmaktadır.<sup>6</sup>

## Sağkalım Analizinde Kullanılan Parametrik Olmayan Testler

Sağkalım analizinde sansürlü verilerin yer aldığı 2 grubun karşılaştırılmasında çok sık kullanılan bir testtir. Logrank testinde farklı gruplardaki bireylerin hazard oranlarının tüm zaman boyunca aynı olduğu varsayılır. Yokluk hipotezi  $H_0: S_1(t) = S_2(t)$  yani 2 grubun sağkalım eğrileri (fonksiyonları) arasında fark yoktur biçiminde kurulur.

$$Z_k(t) = \sum_{A \leq T_i \leq T} W(T_i) r_{ik} \left( \frac{d_{ik}}{r_{ik}} - \frac{d_i}{r_i} \right), k = 1, 2, \dots, K - 1 \quad (1)$$

(1) formülündeki ölümlere farklı ağırlıklandırmalar ( $W_i$ ) verilebilir. Logrank testinde tüm ölümlere eşit ağırlık verilir.  $w_i = 1$  ise logrank testi,  $w_i = r_i$  olursa Gehan testi ve  $w_i = (r_i)^{1/2}$  alınırsa Tarone-Ware testi elde edilir.<sup>7</sup>

## Grup Ardışık Test Yöntemleri

Grup ardışık test yöntemi üzerine çalışmalar yapan Pocock ve O'Brien&Fleming grup sayılarını eşit olduğu durumlar için çalışma yapmışlardır.<sup>3</sup>

Ardışık metot çalışma boyunca verinin periyodik olarak analizidir. Bu metot bir deneyde sü-

rekli sınamayı içerir. Eğer 2 tedavi yöntemi karşılaştırılıyorsa, veriler her tedavi grubundan bir gözlem alınarak çiftler halinde analiz edilir. Her çiftin ardından bir test istatistiği hesaplanır ve bir sonlanma sınırıyla karşılaştırılır. Eğer test istatistiği sınır geçerse, bu durumda çalışmaya (denemeye) son verilir. Aksi takdirde deneme bir sonraki analize kadar devam eder. Bir denemeyi sonlandırma kararı sadece bir tedavinin sonucunun diğerinden anlamlı bir şekilde daha iyi veya daha kötü olmasına bağlıdır.

Klasik ardışık metotlar her veri çiftinden sonra, veri sürekli olarak değerlendirildiğinden dolayı klinik deneylerde çok kullanışlı değildir. Bunun yerine grup ardışık test yöntemleri tercih edilmektedir.

Grup ardışık test yöntemleri ile deneysel bir çalışmada 2 tedavi yöntemi karşılaştırılıyorsa, çalışmanın sonu beklenmeden hangi tedavi yönteminin daha iyi bir sonuç verdiği tespit edilebilir. Belli bir aşamadan sonra hastalar diğer tedavi yöntemine yönlendirilir. Böylece hastaların sağkalım süreleri uzatılmış olabilir.

### Pocock Grup Ardışık Testi

Pocock, her bir grupta aynı sayıda örneklemin olduğu durumlarda, 2 grubun karşılaştırılması üzerine çalışmalar yapmıştır. Pocock'a göre grup sayısının 5'ten büyük olması grup büyüklüklerini azaltacağından dolayı uygulamada fazla bir avantaj sağlamamaktadır.<sup>1</sup>

Maksimum K tane grup testin gücü  $(1-\beta)$  ve  $\alpha$  anlamlılık düzeyiyle beraber önceden belirlenir. Yeterli anlamlılık tespit edilse ya da edilmese de K. analizde denemeye son verilir. i. grupta birikmiş veri için aşağıdaki istatistik hesaplanır.

$$T_i = \frac{S_i}{\sqrt{i}}$$

Burada  $S_i$  bağımsız ve aynı standart normal dağılımlı rassal değişkenlerin toplamıdır. X rassal değişkeni A ve B gibi 2 deneme grubuna sahip olsun. Bu durumda  $S_i$  değişkeni aşağıdaki gibi bulunur.

$$X_j = \frac{1}{\sqrt{n_0}} \left[ \sum_{l=1}^{n_0} \frac{X_{Al} - X_{Bl}}{\sqrt{2\sigma}} \right],$$

$$S_i = \sum_{j=1}^i X_j$$

(3)

Her analiz sonlandırma sınırı olan Z ile karşılaştırılır.

$$P(|T_1| < Z, |T_2| < Z, \dots, |T_{k-1}| < Z, |T_k| \geq Z | H_0) = \alpha$$

Bir denemeyi sonlandırma ardışık bir süreç gösterir. Eğer  $|T_1| \geq Z$  ise ilk analizde çalışma sonlandırılır demektir.  $|T_1| < Z$ ,  $|T_2| \geq Z$  ise çalışma 2. analizde durdurulur.

### O'Brien&Fleming Grup Ardışık Testi

O'Brien&Fleming tarafından geliştirilen bu testte,  $H_0$  hipotezini reddetmek için gerekli olan nominal anlamlılık düzeyi çalışmanın devam ettiği süre içerisinde arttığı için, başlarda  $H_0$  hipotezini reddetmek zordur, ancak sonraki çözümlemelerde reddetmek daha kolaylaşmaktadır.<sup>2</sup>

Test düzeni ve uygulaması bakımından Pocock'un testi ile benzerdir, ancak sınır değerlerinde farklılık bulunmaktadır  $C_i = P(N, \alpha) \sqrt{N/i}$ ,  $i = 1, \dots, N$  olmak üzere testin işleyişi ve örneklem büyüklüğünün hesaplanması Pocock'un testi ile aynıdır.

### $\alpha$ - Harcama Fonksiyonları

Lan ve DeMets alfa harcama fonksiyonunu ortaya koydular.<sup>3</sup> Bu harcama fonksiyonuna göre;  $Z_1, Z_2, \dots, Z_k$  test istatistikleri dizisi için  $b_1, b_2, \dots, b_k$  sınır değerleri belirlenir. Bu sınır değerleri ardışık hipotez testlerinin kritik değerleridir. Çalışmanın her bir aşamasından sonra  $Z_k$  ve  $b_k$  değerleri kıyaslanır, eğer  $|Z_k| < b_k$  ise çalışma devam ettirilir,  $|Z_k| \geq b_k$  olduğu durumda ise ortalamaların eşit olduğu  $H_0$  hipotezi reddedilir ve çalışma durdurulur.

$\alpha$ - harcama fonksiyonunda yer alan zaman göstergesi "t" geçen zamanın maksimum deneme süre-

cine oranını veya toplanan örneklem oranını ifade eder ( $t = n/N$  veya  $t = d/D$ ). Zaman göstergesi  $t$  bir oranı ifade ettiğinden 0 ve 1 aralığında değerler alır.

Alfa harcama fonksiyonunun sahip olduğu karakteristikler  $\alpha(0) = 0$  ve  $\alpha(1) = \alpha$  ki bu karakteristik deneme tamamlandığında sabit bir  $\alpha$  seviyesini garanti etmektedir yani  $\Pr(|Z_1| \geq b_1 \text{ or } |Z_2| \geq b_2, \dots, \text{ or } |Z_k| \geq b_k) = \alpha(t)$ . Beş farklı harcama fonksiyonu aşağıdaki gibidir:

- i) O'Brien&Fleming :  $2 - 2\Phi(Z_{\alpha/2} / \sqrt{t})$
- ii) Pocock :  $\alpha \ln(1 + (e - 1)t)$
- iii) Alfa\*Time :  $\alpha t$
- iv) Alfa\*Time<sup>1.5</sup> :  $\alpha t^{1.5}$
- v) Alfa\*Time<sup>2</sup> :  $\alpha t^2$

### Bulgular

Grup ardışık test yöntemlerinin sağkalım analizinde uygulaması için OMÜ Tıp Fakültesi'ne 1 Ocak 2001 ile 31 Aralık 2002 tarihleri arasında tedavi için gelen 224 akciğer kanseri tanısı olan hasta prospektif olarak araştırıldı.<sup>8</sup> Çalışmada kullanılan veriler sağdan sansürlü verilerdir. Hastalar KHAK ve KHDAK olmak üzere 2 gruptan oluşmaktadır.<sup>6</sup>

52 tane KHAK ve 172 tanede KHDAK olan toplam 224 akciğer kanseri hastası, sağkalım süreleri bakımından hangisinin daha çok yaşam sürdürdüğünü belirlemek için ardışık olarak incelendi. Toplam izlenme süresi 1060 gün olup, 4 adımda her 265 gün için hastalar ardışık olarak ayrı ayrı analiz edildi. Burada amaç 1060 günü beklemeden her 265 günde bir, iki akciğer kanseri olan hastaları sağkalım süreleri bakımından karşılaştırmak ve daha erken sürede sağlıklı sonuç elde etmektir.

İki hasta grubuna ait özet istatistikler Tablo 1'de gösterilmiştir.

Tablo 1'e göre ortanca sağkalım sürelerine bakıldığında KHAK olan hastaların sağkalım süresi 255 gün iken, KHDAK olan hastaların ortanca sağkalım süresinin ise 216 gün olduğu görülmektedir. İki akciğer grubu sağkalım analizi ile test edildiğinde Logrank testine göre sağkalım

**Tablo 1.** İki hasta grubuna ait özet istatistikler.

	Sansürlü	Ölü	Toplam	Ortanca
KHAK	17	35	52	255
KHDAK	27	145	172	216
<b>Toplam</b>	<b>44</b>	<b>180</b>	<b>224</b>	

fonsiyonları arasında anlamlı fark olduğu görüldü ( $p < 0.05$ ).

Grup ardışık test yöntemleriyle 3 duruma göre çözüm yapılabilmektedir. Durum-I de her bir adımda sabit sayıda ölüm gerçekleşene kadar beklenir ve ondan sonra diğer aşamaya geçilir. Durum-II de testin gerçekleştiği zaman aralıkları her adımda eşit alınır. Durum-III'de ise rastgele belirlenen zaman aralıkları ile çözüme gidilir.<sup>8</sup>

Bu çalışmada durum 2'ye göre analiz yapılmıştır. Bu tür çalışmalarda adım sayısı en çok 5 alınmaktadır. Çalışmadaki toplam çalışma süresi 1060 gün olduğundan, uygulama için  $1060/4 = 265$  gün alındığında 4 adım verimiz için daha uygun bulundu. Her adımda (her 265. günde) muhakkak ölüm olayının gerçekleşmesi gerekmektedir. Her bir adımdaki ölüm oranları ( $t_i$ ) hesaplandı. Her 265 gün için analiz sonuçları Tablo 2'de gösterilmiştir.

Tablo 2'de iki hasta grubu 4 adıma göre her 265 günde bir ayrı ayrı olarak incelendi.  $N = 1$  durumunda ilk 265 günde toplam 122 hasta ölmüştür. I. adımdaki (265. gün) ölüm olasılığı  $t_i = 122/180 = 0.67$  bulunur. Benzer şekilde II. adımda (53. gün) 0.95, III. adımda (795. gün) 0.99 ve IV. adımda (1060. gün) 1.00 elde edildi. Grup ardışık test yöntemlerine göre çözüm yapıldığında  $\alpha = 0.05$  için 5 farklı harcama fonksiyonuna göre iki grup arasında fark olmadığı görüldü ( $p > 0.05$ ). Ancak  $\alpha = 0.10$

**Tablo 2.** Dört adıma göre analiz sonuçları.

N	$t_i$	Ölümler			Logrank
		KHAK	KHDAK	Toplam	
N= 1 265. gün	0.67	27	95	122	<b>p= 0.048</b>
N= 2 530. gün	0.95	34	137	171	p= 0.378
N= 3 795. gün	0.99	35	144	179	p= 0.980
N= 4 1060. gün	1.00	35	145	180	<b>p= 0.043</b>

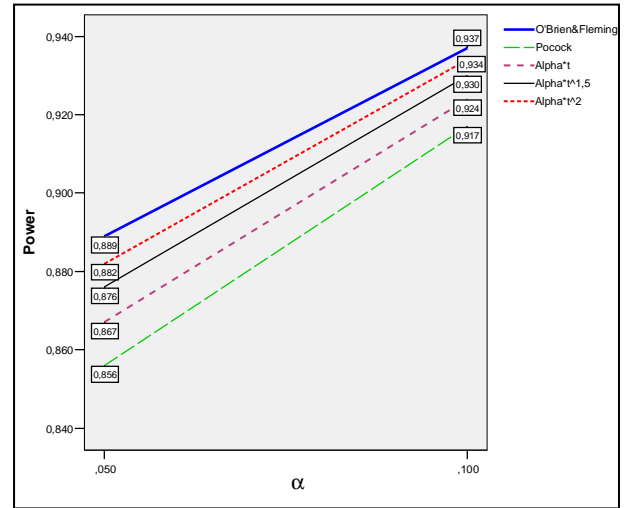
**Tablo 3.**  $\alpha=0.05$  ve  $\alpha=0.10$  için grup ardışık test sonuçları.

$\alpha$	Adım	Log-rank	O'Brien& Fleming		Pocock		$\alpha t$		$\alpha t^{1.5}$		$\alpha t^2$	
			Üst Sınır	Alt Sınır	Üst Sınır	Alt Sınır	Üst Sınır	Alt Sınır	Üst Sınır	Alt Sınır	Üst Sınır	Alt Sınır
0.05	1	1.977	2.502	-2.502	2.072	-2.072	2.126	-2.126	2.205	-2.205	2.283	-2.283
	2	0.883	2.062	-2.062	2.264	-2.264	2.197	-2.197	2.138	-2.138	2.106	-2.106
	3	1.655	2.096	-2.096	2.381	-2.381	2.293	-2.293	2.214	-2.214	2.164	-2.164
	4	2.022	2.123	-2.123	2.427	-2.427	2.333	-2.333	2.250	-2.250	2.196	-2.196
0.10	1	<b>1.977</b>	<b>1.929</b>	-1.929	<b>1.771</b>	-1.771	<b>1.832</b>	-1.832	<b>1.920</b>	-1.920	2.006	-2.006
	2	0.883	1.758	-1.758	1.946	-1.946	1.878	-1.878	1.820	-1.820	1.790	-1.790
	3	1.655	1.802	-1.802	2.059	-2.059	1.968	-1.968	1.889	-1.889	1.839	-1.839
	4	<b>2.022</b>	<b>1.831</b>	-1.831	2.104	-2.104	<b>2.008</b>	-2.008	<b>1.923</b>	-1.923	<b>1.869</b>	-1.869

alındığında 1. adımda (265. gün) Logrank testine göre Pocock,  $\alpha t$  ve  $\alpha t^{1.5}$  harcama fonksiyonlarına göre iki grup arasında fark olduğu ortaya çıkmıştır. 2. adımda (530. gün) ve 3. adımda (795. gün) ise iki grup arasında anlamlı fark bulunmadı. 4. adımda ise (1060. gün) yine Logrank testinde O'Brien & Fleming,  $\alpha t$ ,  $\alpha t^{1.5}$  ve  $\alpha t^2$  harcama fonksiyonlarına göre iki akciğer kanseri arasında bir fark olduğu görüldü. 1. ve 4. adıma göre KHAK olan hastaların KHDAK olan hastalara göre sağkalım sürelerinin daha yüksek olduğu söylenebilir

Grup ardışık testlerin analizi, güç analizleri ve simülasyon çalışmaları Ncss-Pass programında yapıldı. Simülasyon denemesi 1000 defa yapılmıştır.

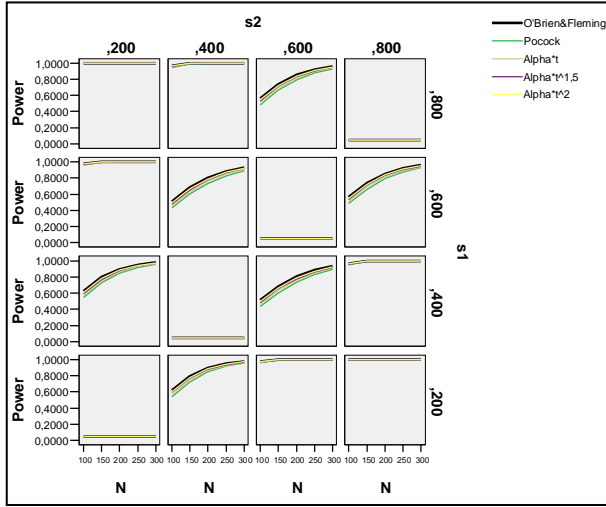
Tablo 3'te farklı N değerleri için 5 harcama fonksiyonunun güç analiz sonuçlarına bakıldı. Tablo 3'te  $\alpha=0.05$  ve  $\alpha=0.10$  için Logrank testi ile 5 harcama fonksiyonunun alt ve üst sınır değerleri verilmektedir.  $\alpha=0.05$  için grup ardışık test yöntemlerine göre iki grup (KHAK-KHDAK) arasında fark olmamaktadır. Ancak  $\alpha=0.10$  alındığında 1. ve 4. adımlarda iki grup arasında fark olduğu görüldü. 1. adımda O'Brien&Fleming, Pocock,  $\alpha t$  ve  $\alpha t^{1.5}$  fonksiyonlarının üst sınır değerlerinin Logrank test istatistik değerini aştığından, 1. adımda (265. günde) iki akciğer kanseri hastası arasında anlamlı fark olduğu görüldü. Yine 4. adımda O'Brien&Fleming,  $\alpha t$ ,  $\alpha t^{1.5}$  ve  $\alpha t^2$  fonksiyonlarının üst sınır değerlerinin Logrank test istatistik değerinden yüksek olmaktadır. Yani 4. adımda (1060. günde) iki akciğer kanseri hastası arasında anlamlı fark olduğu söylenebilir.

**Şekil 1.** N= 224 akciğer kanseri verisi için  $\alpha=0.05$  ve  $\alpha=0.10$ 'a göre 5 harcama fonksiyonunun güç analizi sonuçları.

Çalışmada kullanılan N= 224 akciğer kanseri verisi için  $\alpha=0.05$  ve  $\alpha=0.10$ 'a göre 5 harcama fonksiyonunun güç analizi sonuçları Şekil 1'de verildi.

Şekil 1'e göre 224 hasta ile yapılan çalışmanın istatistiksel olarak ne kadar güçlü olduğuna bakıldı.  $\alpha=0.05$  ve  $\alpha=0.10$  için 5 harcama fonksiyonuna göre güç analizleri karşılaştırıldı.  $\alpha=0.05$  için en yüksek güç değerini O'Brien & Fleming (0.889) ve  $\alpha t^2$  (0.882), en düşük güç değeri ise Pocock (0.856) verdi.  $\alpha=0.10$  için ise en yüksek güç değerini O'Brien&Fleming (0.937) ve  $\alpha t^2$  (0.934), en düşük güç değeri ise Pocock (0.917) ile elde edildi.

Gerçek verilerle elde edilen sonuçların güvenilir olup olmadığını araştırmak için simülasyon



**Şekil 2.** Farklı N ve değişik sağkalım oranları için 5 harcama fonksiyonunun simülasyon çalışması ile güç analizlerinin karşılaştırılması.

çalışması yapılarak, farklı gözlem sayıları (N= 100, 150, 200, 250 ve 300) ile farklı sağkalım oranlarına ( $S_1$  ve  $S_2$ : 0.2 0.4 0.6 0.8) göre 5 harcama fonksiyonunun güç analizleri sonucu Şekil 2’de verilmiştir.

Şekil 2’ye göre iki gruba ait sağkalım oranları eşit olduğunda, bir grupta düşük diğer grupta çok yüksek olduğunda ve bir grupta yüksek diğer grupta çok düşük ise 5 harcama fonksiyonu arasında fark olmadığı görüldü. Ancak iki gruptaki sağkalım oranları birbirinden çok farklı değilse ( $S_1= 0.2$  ile  $S_2= 0.4$  gibi) O’Brien&Fleming harcama fonksiyonu çalışmamızda olduğu gibi en yüksek güç değerini verdi.

### Tartışma

Sağkalım analizi ve grup ardışık test yöntemleri son yıllarda sağlık alanı ile ilgili çalışmalarda çok sık kullanılmaktadır. Bu iki yöntemin bir arada kullanılması sağlık alanında önemli sonuçlar elde edilmesine sebep olacaktır. Çünkü tıbbi çalışmalarda iki tedavi yöntemi-grubu-ilaç karşılaştırılıyorsa, çalışmanın sonunda ancak hangi tedavinin-grubun-ilacın daha iyi sonuç verdiği tespit edilebilir. Grup ardışık test yöntemleri kullanılarak çalışmanın sonu beklenmeden, hangi tedavi yönteminin-grubunun-ilacın daha iyi sonuç verdiği tespit edilebilir. Böylece çalışma süresince belli bir aşamada hastalar diğer tedavi yöntemine-grubuna-

ilaca yönlendirilerek, hastaların sağkalım süreleri uzatılabilir. Bu tür çalışmalar genelde insanlar üzerinde uygulandığından, elde edilen sonuçların ne kadar güçlü olduğu çok önemlidir.

Bu çalışmada sağdan sansürlü 224 akciğer tanısı konan iki grup hastaya (KHAK ve KHDAK) ait verilere grup ardışık test yöntemleri kullanıldı. Sağkalım analizi yöntemlerine göre iki grup hastanın sağkalım sürelerinin farklı olduğu bulundu ( $p < 0.05$ ). KHAK olan akciğer kanseri hastalarının KHDAK olanlara göre sağkalım sürelerinin daha yüksek olduğu görüldü. Grup ardışık test yöntemleri kullanıldığında ise hastalar 1060 günün sonunda değil de her 265 günde bir karşılaştırıldı. Her 265. gün bir adım kabul edilerek, grup ardışık test yöntemlerinde kullanılan 5 harcama fonksiyonuna göre her adım için analiz yapıldı. En yüksek güç değeri O’Brien&Fleming harcama fonksiyonu ile bulundu.

Gerçek verilerle bulunan bu sonucu genelleştirebilmek için farklı N değerleri, değişik sağkalım oranlarına göre simülasyon çalışması yapıldı ve 5 harcama fonksiyonunun güç analizi sonuçları karşılaştırıldı. Simülasyon çalışması sonucunda iki gruba ait sağkalım oranları eşit olduğunda, bir grupta düşük diğer gruptaki sağkalım oranı çok yüksek olduğunda ve bir grupta yüksek diğer grubun sağkalım oranı çok düşük ise 5 harcama fonksiyonu arasında fark olmadığı görüldü. Ancak iki gruptaki sağkalım oranları birbirinden yüksek oranlarda farklı değilse (0.2 ile 0.4 gibi) O’Brien&Fleming harcama fonksiyonu çalışmamızda olduğu gibi en yüksek güç değerini verdi.

Grup ardışık test yöntemleri sansürlü verilerin yer aldığı sağkalım analizi çalışmalarında kullanılarak, ölümlerle sonuçlanan hastalıklarda (kanser gibi) çalışma sonucunu beklemeden belli süre aralıklarında kanser türleri ya da tedavi grupları hakkında önemli bilgilere ulaşılabilir.

### KAYNAKLAR

1. Pocock SJ. Group Sequential Methods in the Design and Analysis of Clinical Trials. Biometrika 1977;64: 191-9.
2. O’Brien PC, Fleming TR. A Multiple Testing Procedure for Clinical Trials. Biometrics 1979;35:549-56.
3. Lan KKG, DeMets DL. Discrete Sequential Boundaries for Clinical Trials. Biometrika 1983;70:659-63.

4. Rebusion DM, Demets DL, Kim K, Lan KKG. Computations for Group Sequential Boundaries Using the Lan-DeMets Spending Function Method. *Controlled Clinical Trials* 2000; 21:190-207.
5. Slud EV, Wei LJ. Two-Sample Repeated Significance Tests Based on the Modified Wilcoxon Statistics. *JASA* 1982;77:855-62.
6. Terzi Y, Cengiz MA, Terzi E, Bek Y. Akciğer Kanseri Olan Hastalarda Prognostik Faktörlerin Sağkalıma Etkisi. *Mersin Ün. Tıp Fakültesi Dergisi* 2005;6: 344-9.
7. Klein JP, Moeschberger ML. *Hypothesis Testing. Survival Analysis-Techniques for Censored and Truncated Data.* 2<sup>nd</sup> ed. New York: Springer; 1997. p.194-6.
8. Demirhan YP, Bacanlı S. Group Sequential Test of Non-Parametric Statistics, for Survival Data. *Hacettepe Journal of Mathematics and Statistics* 2005;34: 67-74.