



Hastalıkların Yere Göre Kümelenmesinde Kullanılan Yöntemler

Erdem Karabulut*, Reha Alpar*, Enis Özayar**

*Hacettepe Üniversitesi Tıp Fakültesi, Biyoistatistik AD, Ankara

**Hacettepe Üniversitesi Tıp Fakültesi, Radyasyon Onkolojisi AD, Ankara

Hastalıklar belirli bir bölgede (yer kümesi) ya da belirli bir zaman aralığında/diliminde (zaman kümesi) yığılım gösterebilir. Hastalıkların yer, zaman ya da yer-zamanda kümelenip kümelenmediğini belirlemek için çok sayıda istatistiksel yöntem geliştirilmiştir. Bu çalışmada, hastalıkların yere göre kümelenip kümelenmediğini belirlemekte kullanılan Grimson, Besag-Newell, Kulldorff'un yersel scan, Lawson-Waller skor yöntemleri tanıtılmıştır. Yapılan uygulama ile yöntemlere ilişkin elde edilen bulgular karşılaştırmalı olarak incelenmiştir.

Uygulama verisi, Hacettepe Üniversitesi Onkoloji Hastanesi Radyasyon Onkolojisi Anabilim Dalında nazofarinks kanseri tanısı ile tedavi gören 248 hastadan elde edilmiştir.

Hastalığın yere göre genel kümelenmesi Grimson yöntemi ile önemsiz bulunmuştur. Yerel (lokal) kümeleme yöntemlerinin uygulanması sonucunda yöntemlere göre farklı bölgelerde olmakla birlikte, yerel hastalık kümelenmeleri önemli bulunmuştur. Kulldorff scan testi kullanıldığında seçilen farklı tarama penceresi büyüklükleri için farklı sonuçlar elde edilmiştir. Tarama penceresinin büyüklüğü arttıkça yöntem büyük kümelere duyarlı olmaktadır. Besag-Newell yönteminde k değerleri 2, 4, ve 6 için yerel kümelenme önemli bulunmazken 8, 10, 12, 14 ve 16 değerleri için yerel kümelenmeler önemli bulunmuştur.

Bir risk odağı çevresinde hastalık kümelenmesini incelemek için, verimizde önceden belirlenmiş bir odak olmamasına karşın, Ankara-Çankaya'nın koordinatları odak olarak varsayıp Lawson-Waller skor testi ile inceleme yapıldığında bu odak çevresinde kümelenme önemli bulunmuştur.

Anahtar Kelimeler: Hastalık kümelenmesi, Yere göre kümelenme, Yersel yapı, Nazofarinks kanseri.

Methods Used For Clustering of Disease Due To Geographical Distribution

Diseases may be aggregated in certain space (geographical clustering) or in certain time periods (time clustering). Several statistical methods have been developed to determine disease clustering in space, time, and space and time. In this study, Grimson, Besag-Newell, Kulldorff's spatial scan, Lawson-Waller score methods used in space clustering of disease were promoted. Findings that were obtained from the applications of some methods were comparatively examined.

Application data were obtained from 248 patients under treatment for nasopharynx cancer in Hacettepe University Oncology Hospital, Radiation Oncology Department between October 1993 and June 2002.

Global spatial disease clustering wasn't statistically significant according to Grimson's method. When local spatial clustering methods were applied, local disease clustering was found significant with all methods, in different areas. In Kulldorff's spatial scan test, different results were obtained by selecting different scanning window sizes. When scanning window size is increased, the method became more sensitive to large clusters. In Besag and Newell method, local clustering were not significant for k=2, 4, and 6, and local clustering were significant for k=8, 10, 12, 14 and 16.

Although there was no evidence of focus of risk in our data, in order to examine the disease clustering around a focus by Lawson and Waller score test, the coordinates of Çankaya-Ankara was set as the coordinates of a focus. Clustering was found significant according to Lawson and Waller score test around this focus.

Key Words: Disease clustering, Space clustering, Spatial pattern, Nasopharyngeal carcinoma.

Zaman zaman iç ve dış basında Çernobil nükleer kazasından sonra Karadeniz bölgesinde ya da çeşitli termik santraller etrafında kanser olgularının sayısında artış (kümelenme) olduğunu belirten yazılar çıkmaktadır. Bu yazılarda

belirtilen görüşlerin bir bölümü sezgisellik ve duygusallıkla kaleme alındığı için bilimsel bir nitelik taşımamakta, sadece yazarların kişisel görüşlerini yansıtmaktadır.

Hastalıkların bir bölgede bir risk odağı (termik santral, nükleer tesis, çöp yakma fırını vb.) çevresinde ya da belirli bir zaman aralığında kümelenip kümelenmediği test etmek amacıyla geliştirilmiş birçok istatistiksel yöntem vardır. Bu çerçevede *hastalık kümesi*; “yer ve zamanda birlikte gruplanan ve sağlık örgütüne bildirilmiş olan sağlık olaylarının, gerçek ya da gözle görülür şekilde sıra dışı yığılması (toplanması)” şeklinde tanımlanmıştır.¹

Hastalıkların (son zamanlarda özellikle kanserin) “yer, zaman ve yer-zamana göre kümelenip kümelenmediğini inceleme” amacına yönelik çalışmalar 1900’lü yılların başlarında başlamış ve ilk metodolojik yöntemler 1950’li yıllarda literatürde yer almıştır. Bu konuya olan ilgi 1960’lı yılların ortasında yoğunluk kazanmış olup bu ilgi günümüzde de devam etmektedir. Kümelenmenin olup olmadığına ilişkin yaygın kullanılan yöntemler arasında Knox, Moran I, Besag ve Newell, Cuzick ve Edwards, Kulldorff’un Scan yöntemleri sayılabilir.²

Bu tür çalışmalar yardımıyla hastalıkların küme(ler) oluşturduğu belirlendiğinde, kümelerdeki olguların ortak bir nedenle (etiyojijyle) ortaya çıktığı düşünülmektedir. Seçilen uygun bir kümeleme yöntemi ile hastalık kümesi/kümelerinin belirlenmesi hastalıkla ilgili ileride yapılacak çalışmalara ışık tutacak, hastalığın kümelendiği yer ve zaman aralıkları belirlenerek neden bu yer ve zaman aralığında olguların sayısının fazla olduğu ayrıntılı olarak incelenebilecektir. Küme araştırmaları sonucunda kümelerin ortaya çıktığı belirlendiğinde kıt olan kaynaklar devlet ve yerel sağlık birimlerince daha etkin bir şekilde kullanılabilir, kümeler yardımıyla hastalığın nedenlerine dayalı epidemiyolojik hipotezlerin kurulmasına yardımcı olunabilecek ve bu konuda ayrıntılı araştırmalar yapılabilecektir. Ayrıca, çok geniş bir bölgede ya da uzun zaman aralıklarında inceleme yapmak yerine sadece kümelenmenin önemli bulunduğu yer-zaman aralıklarında inceleme yapılabilecektir.

Kümelenme, olguların yanlış bildirim, tanımlanamayan demografik değişiklikler ya da olguların birden çok bildirim gibi nedenlerle yapay olarak da oluşabilmektedir. Bu durumda ilginin beklenilmeyen kümelenmelere yoğunlaşacağı açıktır. Genellikle kümelerin kolaylıkla açıklanabilmesi istenir.

Hastalıkların kümelenmesi ve yere göre veri analizi yöntemleri ile yapılan çalışmalar son yıllarda ciddi düzeyde artmıştır. Biyoistatistikçiler arasında da nükleer kuruluşların yakınında lösemi kümelerinin incelenmesi ve bunların haritalarının çıkartılarak atlaslarının oluşturulmasına olan ilginin gittikçe arttığı görülmektedir. Ancak, ülkemizde bu konuda yapılmış çalışmalar şu ana kadar çok sınırlı kalmıştır. Bunun en önemli nedenlerinden biri ülke, bölge ya da il bazında hastalıkların tam ve güvenilir şekilde kayıt edildiği sistemlerin olmamasıdır. Kurulmaya çalışılan kayıt sistemlerinde de kaydedilen olguların sayısı düşük kalmaktadır. Fidaner ve Ark. (2001), Sağlık Bakanlığı’nın ülke genelindeki kanser kayıtlarının tutulabilmesi için 1983 yılında bir ‘edilgen kanser kayıt sistemi’ kurduğunu, ancak burada toplam kanser olgularının yaklaşık dörtte birinin kaydedilebildiğini belirtmişlerdir.³

Hastalıkların nedenleri hakkında fazla bilgi olmadığında kümelenme için güçlü bir istatistik seçmek zordur. Örneğin, kümelenme bir olgudan diğerine hızlı bir yayılma nedeniyle mi oluşmakta, yoksa, bir olgu ortaya çıktığında enfeksiyonun odağı haline gelip ikincil olguların bir kümesini mi oluşturmaktadır? Bulaşma süresi nedir? Enfeksiyon dışında küme için, örneğin radyasyona maruz kalma gibi, bazı ortak nedenler olabilir mi? Bu nedenle, bu ve benzeri sorulara verilebilecek yanıtlara göre uygun yöntemin (istatistik) seçilmesine özen gösterilmelidir.⁴

Son zamanlarda diğer bölgelere göre hastalığın daha az görüldüğü (negatif kümelenme) bölgeleri inceleyen yöntemler de geliştirilmiştir. Bu bölgelerde de neden hastalığın daha az görüldüğü incelenerek hastalığa neden olabilecek değişkenler belirlenebilmektedir.

Küme Çalışmalarında Karşılaşılan Çeşitli Güçlükler

Belirlenen bir kümede, kümelenme nedenlerinin araştırılması sonucunda ortaya çıkacak yararın ne kadar olacağını belirlemek ve kümeyi neyin oluşturduğunu tanımını yapmak kolay olmamaktadır. Diğer taraftan, rutin olarak toplanan istatistiklerden ya da gözlemlerden küme belirlemek, olguların ayrıntılı incelenmesini gerektirmektedir. Ancak, bu incelemeyi yapmak; tanıların, zaman aralıklarının, bölgelerin ve düşünülmesi gereken demografik grupların çeşitliliği nedeniyle de oldukça zordur.

Kümelenmeyi değerlendirmek için geliştirilen yöntemlerdeki zorluklardan biri de uygun küme modelinin seçimi ve alternatif hipotezin

oluşturulmasıdır. Kümelenmenin test edilmesinde olağan soru, bir ya da daha fazla coğrafik bölgede hastalıkların gözlenen yapısının (pattern) sadece şans ile mi oluştuğudur.^{5,6}

Kümelenme, olguların yanlış bildirim, tanımlanamayan demografik değişiklikler ya da olguların birden çok bildirim gibi nedenlerle yapay olarak da oluşabilmektedir. Bu durumda, ilginin beklenilmeyen kümelenmelere yoğunlaşması ve de kümelerin kolaylıkla açıklanamaması sorunu ile karşılaşmaktadır.

Kümelerin Oluşturulmasına İlişkin Yaklaşımlar

Kümelerin oluşturulmasında değişik yaklaşımlar söz konusudur. Wartenberg ve Greenberg,⁷ küme tanımlamak için dört adım olduğunu belirtmişlerdir:

1. Araştırma için elde edilecek verileri tanımlamak.
2. Kümeleme tipini; yer, zaman ya da yer ve zamana göre belirlemek.
3. Kümelenmenin olmayacağı beklendiğinde olguların yapısını tanımlayan yokluk hipotezini belirlemek.
4. Kümeleme altında beklenen olguların yapısına karşılık gelen alternatif hipotezi belirlemektir.

Hastalıkları Yere Göre Kümeleme Yöntemleri

Yere göre (spatial) küme belirleme yöntemleri ikiye ayrılabilir. Birincisi, olası kümelenme yapısını belirlemek için tüm çalışma alanının incelenmesi, ikincisi ise riskin kaynağı olarak önceden belirlenen alanların (odak olarak adlandırılır) incelenmesidir. Besag ve Newell bu ayrımı yapabilmek için genel (general) ve odaklanmış (focused) terimlerini kullanmışlardır.^{8,9} Hastalıkların yere göre kümelenmesi için tüm çalışma alanının kullanıldığı testler de iki alt başlıkta incelenir. Bu kategoriler genel (general) ve yerel (local) testler olarak adlandırılır. Genel testler, çalışma bölgesinde sadece olguların kümelenip kümelenmediği ile ilgilenir. Yerel testler ise yine tüm çalışma alanında olguların kümelenip kümelenmediği hakkında bilgi verirken kümelenme varsa bunun yerinin neresi olduğunu da belirtir. Bu yaklaşımlara ilişkin testler *genel*, *yerel* ve *odaklanmış* alt başlıklarında incelenmektedir. Çalışmamızda, hastalıkların yere göre genel kümelenmesini incelemek için Grimson ve Besag-Newell yöntemleri, yerel kümelenmeleri incelemek için Besag-Newell ve Kuldorff'un scan yöntemleri ve bir odak çevresindeki kümelenmeyi incelemek için de Lawson-Waller skor yöntemleri kullanılmıştır.

Hipotezlerin Oluşturulması

Hastalıkların yere göre kümelenmesinde, belirli yerleşimlerde ya da alanlarda hastalığın daha yaygın olup olmadığı sorusu ile ilgilenilir.

Homojen populasyonlarda, kümelemeye yönelik olarak yapılacak testlerde yokluk hipotezi aşağıdaki gibi tanımlanır;

H_0 : Her birey diğer bireylerden bağımsız olarak, eşit olasılıkla hastalığa yakalanır.

Büyük homojen populasyonlarda ve nadir görülen hastalıklarda ise yokluk hipotezi şöyle tanımlanır;

H_0 : C_i kümesi $E(C_i) = n_i \lambda$ ($1 \leq i \leq I$) ile bağımsız Poisson rasgele değişkenleridir.

Burada λ ; her bir birey için ortak baseline insidans oranını gösterir. Genel testlerde, H_0 doğru değildir alternatif hipotezi kullanılırken, odaklanmış testlerde daha spesifik alternatif hipotez kullanılır. Bu hipotez: önceden tanımlanmış odağa yakın hücrelerdeki ortalama hızlar yüksektir şeklindedir.

Grimson Yöntemi

Grimson yöntemi çok yönlü bir yöntemdir. Hastalık hızları ve olgu-sayıları verilerinin her ikisini kullanarak yere göre kümeleri belirlemek için kullanılabilir. Hızlarla kullanıldığında, ortak sınırı olan alanlar ardışık olarak tanımlanır. Yüksek-riskli alanlar kümelenmediğinde test istatistiği büyük olur. Hem olguların hem de kontrollerin yerleşimleri bilindiğinde yöntem olgu-kontrol verileri ile de kullanılabilir. Fazla olgunun olduğu alanlar "yüksek-riskli" olarak etiketlenir ve bir diğerinden belirli uzaklıkta olanlar ardışık olarak düşünülür.

Test istatistiği A , etiketlenmiş alanlar arasında gözlenen ardışıklığın/komşuluğun sayısıdır. Etiketlenmiş hücrelerin sayısı (n), toplam hücre sayısı x 'e göre çok küçükse Poisson dağılımı A 'nın yokluk hipotezi altındaki dağılımı için mükemmel bir yaklaşım sağlar. Grimson ve arkadaşları yaptıkları bir simülasyon çalışması ile etiketlenmiş nesnelerin ardışıklıklarının sayısının Poisson dağılımına uyduğunu göstermiştir. n/x büyükse yokluk hipotezi altında A 'nın normal dağılıma yaklaştığı varsayılır ve z skoru aşağıdaki gibi hesaplanır. Genel olarak elde edilen z değerinin önemliliği standart normal dağılım yardımıyla test edilir.¹⁰

Grimson ve Rose 1991 yılında yaptıkları çalışmada, geliştirdikleri yöntemi DRG'da (dorsal root sensory ganglia) nöronlara uygulamışlardır. Çalışmadaki genel

amaç; gangliondaki hücre gövdelerinin yerleşimlerini, hücrelerin ayrı ayrı çevresel uzanımlarını taşıdığı özel çevresel sınırlara göre belirlemek iken, özel amaç; hücre gövdelerinin, ganglionlarda onların projeksiyonlarını içeren sınırlara göre nasıl organize olduklarını belirlemektir. Çalışmada etiketlenmiş hücre gövdelerinin kümelendiği bulunmuştur.¹⁰

Besag ve Newell'in Yöntemi

Bir olgunun oluştuğu bölgeye A_0 etiketi atansın. Diğer tüm bölgelere de A_0 'ın merkezinden kendi merkezlerine olan uzaklıklarına göre artan bir şekilde sıralanan A_i $i=1,2,\dots,n$ etiketleri verilsin. i . bölgedeki olguların sayısı y_i ve nüfusu t_i $i=1,2,\dots,n$ ile gösterilsin. Merkez koordinatlarının kesikli olması nedeniyle pratikte oluşabilecek eşit değerler (ties) rasgelelik ya da bölge birleştirilmesiyle ortadan kaldırılabilir. Yöntem için gerekli olan olguların yığılımlı toplamı (D_i) ve risk altındaki nüfusun yığılımlı toplamı (U_i) aşağıdaki gibi tanımlanır:

$$D_i = \left(\sum_{j=0}^i y_j \right) - 1, \quad U_i = \left(\sum_{j=0}^i t_j \right) - 1$$

Böylece; $D_0 \leq D_1 \leq \dots$ olguların A_0, A_1, \dots bölgelerinde toplanmış sayısı ve $U_0 \leq U_1 \leq \dots$ bu bölgelere karşılık gelen risk altındaki nüfusun toplanmış sayısıdır. $M = \min\{i : D_i \geq k\}$ olsun. Diğer bir deyişle, A_0, \dots, A_m bölgelerinde en az k olgu varken A_0, \dots, A_{m-1} bölgelerinde k 'dan daha az olgu vardır. M 'nin gözlenen küçük değeri A_0 etrafında bir kümelenmeyi gösterir. Her bir potansiyel küme için poisson dağılımı yardımıyla anlamlılık düzeyi hesaplanır. Besag ve Newell yönteminde öncelikle bireysel kümelerin belirlenmesiyle ilgilenilmesine rağmen, çalışma bölgesi bir bütün olarak ele alındığında belirlenen kümelerin sayısının beklenenden daha büyük olup olmadığı araştırılabilir. Besag ve Newell, geliştirdikleri küme belirleme yönteminin bir tarama aygıtı olarak tasarlandığını, epidemiyolojik olarak önemli kümeleri dışarıda bırakabileceğinden çoklu karşılaştırma için düzeltme kullanmadıklarını belirtmiştir. Bu yöntemle daha ayrıntılı incelemelerin yapılmasının istendiği bölgelerin sayısı belirlenebilmekle birlikte Besag ve Newell, düzenlenen araştırmanın temel amacı genel olarak bir kümelenmeyi test etmek olduğunda başka bir testin seçilmesini önermiştir.⁸

E.G. Hill ve arkadaşları (2000), Santa Clara (California) Şehrindeki 7 sayım bölgesine su sağlayan iki kuyunun organik solvent ile kirlenmesi sonucu riske maruz kalan bölgede 1 Ocak 1981 ile 31

Ağustos 1982 tarihleri arasında canlı doğumlardaki önemli kalp anomalilerini incelemiştir. Santa Clara Şehrinde 1981 yılında düzeyinde şiddetli kalp kusurlu doğumların genel kümelenmesi Besag-Newell yöntemi ile önemli bulunmamıştır ($P>0,05$).¹¹

L.A. Waller ve arkadaşları (1995), 1980-1990 yılları arasında İsveç'deki 2594 bölgede 0-15 yaş grubundaki çocuklarda görülen akut lymphocytic leukaemia'lı (ALL) 656 olguyu incelemiştir. Besag-Newell'in yönteminde kullanılan farklılık değerleri için hastalık olgularının kümelenmesi önemli bulunmamıştır ($P>0,05$).¹²

L.A. Waller ve B.W. Turnbull (1993), Upstate New York'ta aktif olmayan ve trichloroethylene içeren tehlikeli atık bölgesini şüphelenilen kaynak olarak düşünerek 1978-1982 yılları arasında bölgedeki lösemi insidansının bu kaynak etrafında kümelenmesini üç yöntemle (Besag ve Newell, Stone, Waller ve ark.) incelemiştir. Araştırmacılar, çalışma bölgesinin bölünmesini iki farklı şekilde yapmıştır: blok grup bölünmesi ve sayım bölgesi bölünmesi. Besag ve Newell yöntemi için k değeri 6, 8, 10 alındığında blok grup ve sayım bölgesi bölünmesinde farklı sonuçların elde edildiği belirtilmiştir.¹³

Kulldorff'un Yersel Scan Testi

Gerek Kulldorff'un yersel scan testinde gerekse diğer scan testlerinde sıklıkla kullanılacak bir tanım *pencere* tanımıdır. Pencere, basitçe; incelenecek coğrafik bölgede büyüklüğü önceden belirlenen ya da çalışma sürecinde değişikliğe uğrayabilen (genişleyebilen ve daralabilen) küçük geometrik alanlardır (daire, kare vb. gibi).

Kulldorff'un yersel scan testi, önceden belirlenmiş bir büyüklüğe kadar tarama penceresinin değişmesine olanak verir. Küme tarafından kaplanan alanın büyüklüğü hakkında önsel bilgi olmadığında değişken pencere büyüklüğünü kullanmak daha yararlıdır. Bu yöntemde, çapının büyüklüğü sürekli değişen çember çalışma bölgesi içinde hareket ettirilir. Her bir çemberin içindeki ve dışındaki hastalık olgularının gözlenen sayısı ve risk altındaki nüfusu yansıtan beklenen sayısı belirlenir. Bu sayılar temelinde her bir çember için olasılık hesaplanır. Olasılığın en büyük olan çember, en çok olası küme olarak gösterilir. Önemlilik Monte Carlo benzetimiyle değerlendirilir.¹⁴

15

M. Kulldorff (1999), New Mexico'da 1973-1991 yılları arasındaki kanser kayıtlarını kullanarak, 9254 akciğer

kanseri olgusunun kümelenmesini yersel scan istatistiği ile incelemiştir. En çok olası küme, New Mexico'nun güneybatısındaki Chaues, Eddy ve Lea'da istatistiksel olarak önemli bulunmuştur ($p=0,0001$).¹⁶

Lawson ve Waller Skor Testi

Lawson, Waller ve arkadaşları, verilen bir konum çevresinde hastalık kümelenmesinin olup olmadığını test etmek için bir yöntem geliştirmişlerdir. i . hücredeki olgular (c_i) için riske maruz kalmanın ölçüsü g_i olsun. Maruz kalmanın kesin kestirimleri olmadığında, g_i , hücre i 'nin merkezinden odağa olan uzaklığın reciprocal'i ($1/\text{uzaklık}$) olarak alınabilir. Test istatistiği U , her bir hücredeki gözlenen hastalık insidansının H_0 altında beklenenden sapmasının, maruz kalmanın ölçüsü (g_i) ile ağırlıklandırılmış toplamıdır. Test istatistiği U ;

$$U = \sum_{i=1}^I g_i (c_i - E_i) \quad (2.69)$$

ile verilir. Burada, yokluk hipotezi altında hücre i 'deki olguların beklenen sayısı $E_i = n_i c_+ / n_+$ 'dir. Yokluk hipotezi altında test istatistiği U 'nun beklenen değeri

sıfır ve yaklaşık varyansı $\sum_{i=1}^I g_i^2 E_i$ 'ye eşittir.

Standartlaştırılmış istatistik U^* :

$$U^* = \frac{U}{\sqrt{\text{var}(U)}} \quad (2.70)$$

ile verilir. Çok nadir görülen hastalıklar hariç U^* çoğu durumda asimtotik olarak standart normal dağılıma uyar.¹³

L.A. Waller ve arkadaşları (1995), 1980-1990 yılları arasında İsveç'te 0-15 yaş grubundaki çocuklarda görülen akut lymphocytic leukaemia (ALL) olgularının dört nükleer tesis etrafındaki kümelenmesini Lawson-Waller skor testi ile incelemiş ve nükleer tesisler etrafında kümelenmeyi önemsiz bulmuştur ($P>0,05$).¹²

L.A. Waller ve B.W. Turnbull (1993), Upstate New York'ta tehlikeli atık bölgesini şüphelenilen odak olarak düşünerek 1978-1982 yılları arasında bölgedeki lösemi insidansının bu kaynak etrafında kümelenmesini Lawson-Waller yöntemiyle incelemiştir. Lawson-Waller yönteminde, blok grup ve sayım bölgeleri kullanıldığında birbirine yakın p değerleri elde edildiği belirtilmiştir. Araştırmacılar, çalışma bölgesinin bölünmesiyle ilgili seçimin yersel

yapıyı inceleyerek elde edilen sonuçlar üzerine önemli etkisi olduğunu belirtmiştir.¹³

GEREÇ – YÖNTEM

Uygulamada, Hacettepe Üniversitesi Onkoloji Hastanesinde nazofarinks kanseri nedeniyle Radyasyon Onkolojisi Anabilim Dalı'nda tedavi görmüş olan hastalardan elde edilen veriler kullanılmıştır. Onkoloji hastanesinin açıldığı 1993 yılından 2002 Haziran ayına kadar toplam 270 nazofarinks kanserli hasta tedavi görmüştür. Bu hastalardan hem yerleşim yeri hem de tanı konulma tarihi bilinen 248 hasta çalışmaya alınmıştır.

Bu hastaların yerleşim yerleri arasındaki kuş uçuşu mesafeleri hesaplamak için gerekli olan ilçeler bazında yer koordinatları Maden Tetkik ve Arama Genel Müdürlüğü Uzaktan Algılama Merkezi'nden alınmıştır. Uygulamamızda kullanılan bilgisayar yazılımlarının bazılarında, yerleşim yerleri arasındaki mesafeyi hesaplayabilmek için coğrafi koordinatları (enlem ve boylam) kullanılırken bazılarında UTM (Universal Transverse Mercator) koordinatları kullanılmaktadır. UTM koordinatları kullanıldığında iki nokta arasındaki uzaklık doğrudan Öklit uzaklık formülü ile hesaplanabilmektedir. Bu nedenle, Besag-Newell ve Lawson-Waller skor testleri için Visual Fox Pro kullanılarak geliştirdiğimiz yazılımda, UTM koordinatları kullanılmıştır. Hastalık olgularının görüldüğü bölgelere ilişkin nüfuslar 1997 Nüfus Tespiti ve 2000 Nüfus Sayımı sonuçlarından alınmıştır. Hastalıkların yere göre kümelenmesini incelemek için SatScan V2.1.3, Cluster V3.1 ve tarafımızdan geliştirilen yazılım kullanılarak veriler incelenmiştir.

BULGULAR

Grimson yöntemiyle yere göre kümelenme incelenirken, bölgelerin etiketlenmesi için nazofarinks görülme sıklığı olan yüzbinde 1,34 kullanılmış ve hastalık görülme hızı bu değer üzerinde olan bölgeler etiketlenmiştir. Cluster V3.1 paket programı ile elde edilen sonuçlar aşağıda Tablo 1'de verilmiştir. Tablo 1'den yere göre kümelenmenin önemli olmadığı görülmektedir ($P=0,084$).

Kulldorff'un scan istatistiği kullanılarak hastalıkların yere göre kümelenmesi incelendiğinde, tarama penceresinin büyüklüğünün alabileceği en büyük değere bağlı olarak analiz sonuçları değişmektedir. Tarama penceresinin alabileceği en büyük değer büyük seçildiğinde yöntem büyük kümelere, küçük

Karabulut ve ark

seçildiğinde küçük kümelerle duyarlı olmaktadır. SatScan 2.1.3 paket programı ile tarama penceresinin büyüklüğü, çalışma bölgesinin %5'inden %50'sine kadar genişletilerek incelemeler yapılmıştır. Tarama penceresinin büyüklüğü çalışma bölgesinin %5'i seçildiğinde en çok olası kümenin merkezi Ankara-Çankaya (P=0,002) ve ikincil kümenin merkezi Konya-Akşehir'de (P=0,034) kümelenme önemli bulunmuştur. Pencere büyüklüğü %25 ve %50 seçildiğinde ise sırasıyla sadece en çok olası küme olarak belirlenen Karabük ve Ankara-Polatlı'da kümelenmeler önemli bulunmuştur (P=0,001).

Tablo 1. Grimson yöntemi sonuçları

Etiketleme için kullanılan hız (Yüzbinde)	1,34
Gözlenen Ardışıklık Sayısı	80
Beklenen Ardışıklık Sayısı	72,87
Yüksek Risk Ardışıklıklarının St. Sapması	5,17
Z değeri	1,379
P	0,084

Yerel kümelenmeyi incelemek için kullanılan Besag ve Newell yönteminde k değeri 2, 4, 6, 8, 10, 12, 14 ve 16 seçilerek test yapılmıştır. Test işlemleri geliştirilen yazılım kullanılarak yapılmış ve sonuçlar Tablo 2'de verilmiştir. k=8, 10, 12, 14 ve 16 seçildiğinde bazı bölgelerde yerel kümelenmeler, k=12, 14 ve 16 seçildiğinde ise genel yere göre kümelenme önemli bulunmuştur.

Bir odak çevresindeki kümelenmeyi incelemek için Lawson ve Waller'in skor testi kullanılmıştır. Bu testi kullanabilmek için önceden elimizde herhangi bir bilgi olmamasına karşın Ankara-Çankaya'nın hastalık kümelenmesi için bir odak olduğunu varsayalım. Ankara-Çankaya odak olarak alındığında geliştirilen yazılım ile elde edilen sonuçlar aşağıda Tablo 3'te verilmiştir. Tablodan da görülebileceği gibi, odak olarak seçilen Ankara-Çankaya çevresinde hastalık olgularının kümelenmesi önemli bulunmuştur (P<0,001).

TARTIŞMA

Nazofarinks kanserinin genel yere göre kümelenmesi Grimson yönteminde önemli bulunmazken, Besag-Newell'in yönteminde k değeri 12, 14 ve 16 için genel yere göre kümelenme önemli bulunmuştur. Ancak, Besag ve Newell genel olarak yere göre kümelenme olup olmadığı incelenmek istendiğinde başka yöntemlerin kullanılmasını önermişlerdir. Bu nedenle, veride genel olarak yere göre bir kümelenmenin olmadığına karar verilmiştir.

Tablo2. Besag - Newell yöntemi sonuçları (önemli bulunan kümeler)

K	Bölge Adı	Yerel		Genel	
		M	P	R	P
K=8	Kastamonu	6	0,042	2	0,145
	Konya-Selçuklu	0	0,044		
K=10	Konya-Selçuklu	0	0,006	1	0,264
K=12	Çankırı	5	0,036	5	0,010
	Konya-Cihanbeyli	3	0,028		
	Konya-İlgin	3	0,031		
	Konya-Karatay	1	0,027		
K=14	Konya-Selçuklu	1	0,027		
	Konya-Cihanbeyli	3	0,005	4	0,052
	Konya-İlgin	3	0,006		
	Konya-Selçuklu	3	0,042		
K=16	Konya-Tuzlukçu	4	0,023		
	Karaman	4	0,024	7	0,024
	Karaman-Ermenek	4	0,046		
	Kastamonu-Azdavay	11	0,046		
	Kırıkkale	3	0,030		
	Konya-Cihanbeyli	5	0,041		
Konya-İlgin	4	0,005			
	Konya-Tuzlukçu	4	0,005		

Tablo 3. Lawson ve Waller'in skor testi sonuçları

Odak Noktasının X Koordinatı	728
Odak Noktasının Y Koordinatı	4747
Hastalık Görülme Sıklığı	0,0000134
U	17,143
Var(U)	9,718
Test İstatistiği U*	5,499
P	<0,001

Yerel (lokal) yere göre kümeleme yöntemlerinde hem kümelenmenin olup olmadığı hem de kümelenmenin yeri belirlenebilmektedir. Ayrıca, genel yere göre testlerde belirlenemeyen yerel kümelenmeler belirlenebilmektedir. Besag ve Newell yöntemi uygulandığında k değeri 8, 10, 12, 14 ve 16 için çeşitli bölgelerde (özellikle Konya'nın ilçelerinde) yerel yere göre kümelenmeler önemli bulunmuştur.

Kulldorff yersel scan yönteminde seçilen tarama penceresi büyüklüğüne göre test sonuçları değişmektedir. Bu nedenle, hastalık kümelerinin büyüklüğü hakkında bir öngörümüz varsa pencere büyüklüğü bu öngörüye göre belirlenebilir. Kulldorff'un yersel scan istatistiği uygulanırken tarama penceresinin büyüklüğü, çalışma bölgesinin büyüklüğünün %5'inden %50'sine kadar olacak şekilde çeşitli değerler seçilmiştir. Tarama penceresinin büyüklüğü küçük seçildiğinde test küçük kümelenmelere duyarlı olmaktadır. Tarama penceresinin büyüklüğü %5 seçildiğinde olası küme olarak belirlenen kümelerin içerisinde, hem en çok olası küme olarak belirlenen küme hem de ikincil kümeler olarak belirlenen kümelerden biri önemli bulunmuştur. İkincil kümelenmenin önemli

Hastalıkların Yere Göre Kümelenmesinde Kullanılan Yöntemler

bulduğu kümeler ile en çok olası küme üst üste çıkışabilir. Bu durumda, ikincil kümelerin pek fazla açıklayıcılığı olmayacağı için bu kümeler gözardı edilebilir. Seçilen diğer tarama penceresi büyüklüklerinde sadece en çok olası küme olarak belirlenen kümeler önemli bulunmuştur.

Elimizde risk kaynağı olarak belirlenmiş bir odak olmamasına karşın uygulama yapmak amacıyla Ankara-Çankaya odak olarak seçilip bu odak çevresinde olguların kümelenip kümelenmediği Lawson-Waller skor testi ile incelendi. Test sonucunda odak olarak seçilen Ankara-Çankaya çevresinde kümelenme önemli bulunmuştur. Ankara'ya uzak yerlerden Hacettepe Hastanesine gelen hastaların sayısının Ankara'ya yakın yerlerden gelen hastaların sayısından daha az olacağı göz önünde bulundurulduğunda böyle bir sonucun çıkması doğaldır.

Literatürde, kullanılan yöntemlerin güçlerinin karşılaştırıldığı çeşitli çalışmalar da bulunmaktadır. Ancak, yöntemlerin çeşitliliği ve farklı parametreleri kullanmaları nedeniyle hangi durumda hangi yöntemin en iyisi olduğunu belirlemek zordur. Bu yöntemlerin gücü, verinin yer-zamanda dağılımına göre ya da kullanılan yöntem için seçilen parametrelere göre farklılık göstermektedir. Bu nedenle, seçilen kümeleme tipine göre hastalıkların kümelenip kümelenmediği incelenirken birkaç farklı yöntem kullanıp değerlendirme yapılabilir.

Kümeleme çalışmalarından doğru sonuç elde edebilmek için “nedenleri tam olarak bilinmeyen ya da seyrek görülen hastalıkların” kayıtlarının merkezi olarak tutulacağı kayıt sistemleri oluşturulmalı ve bu kayıt sistemleri ortaya çıkan hastalık olgularının hepsini kapsayacak şekilde düzenlenmelidir. Bu

kayıtlarda, olguların görüldüğü bölgeler ile ilgili yer bilgilerinin de bulunması çok önemlidir. Olguların yer koordinatlarının sadece il-ilçe bazında değil, olguların adreslerine göre daha ayrıntılı belirlenmesi incelemelerde daha fazla bilgi sağlayabilecektir. Bu nedenle, sağlık olaylarıyla ilgili bir Coğrafi Bilgi Sistemi'nin oluşturulması yararlı olabilir.

KAYNAKLAR

1. Metz L.M. ve McGuinness S. Responding to Reported Clusters of Common Disease: The Case of Multiple Sclerosis. *Canadian Journal of Public Health* 1997; 88(4): 277-79.
2. Özdamar Kazım. Hastalıkların Yer, Zaman ve Yer-Zaman'da Kümelenmelerinin Belirlenmesinde Yararlanan Yöntemler. 5. Ulusal Biyoistatistik Kongresi Bildiri Kitabı, Eskişehir 2000; 234-61.
3. Fidaner C., Eser S.Y. ve Parkin D.M. Incidence in Izmir in 1993-1994: First Results From Izmir Cancer Registry. *European Journal of Cancer* 2001; 37: 83-92.
4. Ederer F., Myers M.H. ve Mantel N. A Statistical Problem in Space and Time: Do Leukemia Cases Come in Clusters. *Biometrics* 1964; 20: 626-38.
5. Rothman K.J. A Sobering Start for The Cluster Busters' Conference. *American Journal of Epidemiology* 1990; 132: S6-13.
6. Wakefield J. ve Quinn M. Editorial: Disease Clusters and Ecological Studies. *Journal of Royal Statistics Society A2001*; 164: 1-2.
7. Jacquez G.M. ve Ark. The Analysis of Disease Clusters, Part I. *Infection Control and Hospital Epidemiology* 1996; 17(5): 319-27.
8. Besag J. ve Newell J. The Detection of Clusters in Rare Disease. *Journal of Royal Statistics Society. A* (1991); 154:143-55.
9. Marshall R.J. A Review of Methods for the Statistical Analysis of Spatial Patterns of Disease. *Journal of Royal Statistics Society. A* (1991); 154:421-41.
10. Grimson R.C. ve Rose R.D. A Versatile Test for Clustering and A Proximity Analysis of Neurons. *Methods of Information in Medicine* 1991; 30: 299-303.
11. Hill E.G. ve Ark. A Comparison of Three Tests to Detect General Clustering of Rare Disease in Santa Clara County, California. *Statistics in Medicine* 2000; 19: 1363-378.
12. Waller L.A. ve Ark. Detection and Assessment of Clusters of Disease: An Application to Nuclear Power Plant Facilities and Childhood Leukaemia in Sweden. *Statistics in Medicine* 1995;14:3-16.
13. Waller L.A. ve Turnbull B.W. The Effects of Scale on Tests for Disease Clustering. *Statistics in Medicine* 1993; 12: 1869-884.
14. Kulldorff Martin. A Spatial Scan Statistic. *Communication in Statistics* 1997; 26(6): 1481-496.
15. Kulldorff M. ve Ark. Evaluating Cluster Alarms: A Space-Time Scan Statistic and Brain Cancer in Los Alamos, New Mexico. *American Journal of Public Health* 1998; 88: 1377-380.
16. Kulldorff Martin. Geographic Information System (GIS) and Community Health: Some Statistical Issues. *Journal of Public Health Management Practice* 1999; 5(2): 100-06.

Yazışma Adresi:

Dr. Erdem Karabulut
Hacettepe Üniversitesi, Tıp Fakültesi
Biyoistatistik Anabilim Dalı, Sıhhiye / Ankara
Tel : 312 305 14 67
E-posta : ekarabul@hacettepe.edu.tr