

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM ANABİLİM DALI

**YAKALA TEKRAR YAKALA YÖNTEMİYLE PREVALANS
VE DİAGNOSTİK DOĞRULUK TAHMİNİ**

Havva Didem OVLA

YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. Bahar TAŞDELEN

MERSİN-2012

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM ANABİLİM DALI

**YAKALA TEKRAR YAKALA YÖNTEMİYLE PREVALANS
VE DİAGNOSTİK DOĞRULUK TAHMİNİ**

Havva Didem OVLA

YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. Bahar TAŞDELEN

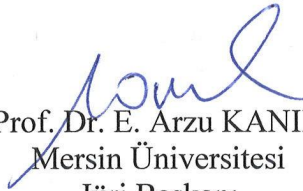
Tez No: 230

MERSİN-2012

Mersin Üniversitesi Sağlık Bilimleri Enstitüsü

Biyoistatistik Ve Tıbbi Bilişim Anabilim Dalı Yüksek Lisans programı çerçevesinde yürütülmüş olan “Yakala Tekrar Yakala Yöntemiyle Prevalans ve Diagnostik Doğruluk Tahmini” başlıklı çalışma, jürimiz tarafından yüksek lisans tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 19/12/2012


Prof. Dr. E. Arzu KANIK
Mersin Üniversitesi
Jüri Başkanı


Doç. Dr. Bahar TAŞDELEN
Mersin Üniversitesi
Jüri Üyesi


Yrd. Doç. Dr. İlter HELVACI
Mersin Üniversitesi
Jüri Üyesi

Bu tez, Enstitü Yönetim Kurulu'nun 25.12.2012 tarih ve 2012/393 sayılı kararı ile kabul edilmiştir.



TEŞEKKÜR

Yüksek lisans eğitimime başladığım andan itibaren her konuda yardımlarını ve desteğini gördüğüm, başta danışman hocam Sayın Doç. Dr. Bahar TAŞDELEN'e ve Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı Başkanı Sayın Prof. Dr. Emine Arzu KANIK'a sonsuz teşekkürlerimi sunarım.

Tezimin hazırlanması sırasında bana anlayış gösteren bölüm arkadaşlarıma teşekkürlerimi sunarım.

Hayatımın her aşamasında olduğu gibi bu tez çalışmam sırasında da benden manevi desteğini esirgemeyen aileme ve bana güvenen dostlarım Ceren, Didem, Emre, Gizem ve Mine'ye teşekkür etmeyi bir borç bilirim.

Saygılarımla

Havva Didem OVLA

Mersin, 2012

İÇİNDEKİLER

KABUL VE ONAY	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
SİMGELER VE KISALTMALAR DİZİNİ	x
ÖZET	xi
ABSTRACT	xii
1. GİRİŞ	1
2. GENEL BİLGİLER	3
2.1. Örneklem Nedir?	3
2.1.1. Olasılıklı Örneklem Yöntemleri	3
2.1.1.1. Basit Rasgele Örneklem	3
2.1.1.2. Sistematiik Örneklem	4
2.1.1.3. Tabakalı Rastgele Örneklem	4
2.1.1.4. Küme Örneklem	5
2.1.2. Olasılıksız Örneklem Yöntemleri	5
2.1.2.1. Gelişigüzel Örneklem	6
2.1.2.2. Amaçlı Örneklem	6
2.1.2.3. Kota Örneklemesi	7
2.1.2.4. Kartopu Örneklemesi	7
2.2. Sağlık Kayıtları Sırasında Karşılaşılan Sorunlar	7
2.3. Sağlık İstatistikleri İçin Veri Toplama Yöntemleri	10
2.3.1. Aktif Yöntem	10
2.3.2. Pasif Yöntem	10
2.3.3. Aktarım	10
3. MATERYAL VE YÖNTEM	12
3.1. Materyal	12

3.2. Yöntem	15
3.2.1. Yakala Tekrar Yakala (Capture Recapture) Yöntemi	15
3.2.1.1. Yakala Tekrar Yakala Yönteminin Varsayımları	16
3.2.1.2. Açık Popülasyon	18
3.2.1.3. Kapalı Popülasyon	18
3.2.1.4. Yakala Tekrar Yakala Yönteminin Avantajları ve Dezavantajları	19
3.2.2. Yakala Tekrar Yakala Yönteminin Uygulanışı	19
3.2.2.1. Lincoln-Petersen Tahmin Yöntemi	20
3.2.2.2. Jolly-Seber Tahmin Yöntemi	23
3.2.2.3. Schnabel Tahmin Yöntemi	26
3.2.2.4. Schumacher-Eschmeyer Tahmin Yöntemi	27
3.2.3. Yakala Tekrar Yakala Tahmininde Yanlılık Tipleri	29
3.2.4. Tarama Testleri İçin Yakala Tekrar Yakala Yaklaşımı	31
3.2.4.1. Walter'ın Kısmi Gizli Sınıf Modeli	33
3.2.4.2. Lincoln-Petersen Tahmini	34
3.2.4.3. Bağımlılık Parametrelerine Dayalı Tahmin	34
3.2.4.4. Odds Oranına Dayalı Tahmin	35
3.2.5. Logaritmik Doğrusal (Log-linear) Model	36
3.2.5.1. Model Seçim Kriterleri	39
3.2.5.2. Logaritmik Doğrusal Model Kısıtlılıkları	42
4. BULGULAR	44
4.1. Tahmin Yöntemlerinin Uygulanışı	44
4.1.1. Lincoln-Petersen Tahmin Yöntemine Ait Uygulama	44
4.1.2. Jolly-Seber Tahmin Yöntemine Ait Uygulama	45
4.1.3. Schnabel Tahmin Yöntemine Ait Uygulama	47
4.1.4. Schumacher-Eschmeyer Tahmin Yöntemine Ait Uygulama	48
4.2. Yakala Tekrar Yakala Yaklaşımıyla Tarama Testlerinin Doğruluk Tahmini	49
4.3. Logaritmik Doğrusal Modeller Kullanılarak Prevalans Tahmini	51
4.4. Literatür Taraması Bulguları	53
5. TARTIŞMA	65
6. SONUÇ VE ÖNERİLER	68
7. KAYNAKLAR	70

EK.....	80
ÖZGEÇMİŞ	81

ŞEKİLLER DİZİNİ

Şekil 3.1. Yakala tekrar yakala uygulama yöntemleri.....	20
Şekil 4.1. Literatür taraması için akış şeması.....	54

ÇİZELGELER DİZİNİ

Çizelge 3.1. Lincoln-Petersen tahmini için simülasyon ile elde edilen veri seti (1: işaretlenen, 0: işaretlenmeyen)	12
Çizelge 3.2. Jolly-Seber tahmini için kullanılan hipotetik veri seti	12
Çizelge 3.3. Schnabel tahmini için kullanılan hipotetik veri seti	13
Çizelge 3.4. Schumacher-Eschmeyer tahmini için kullanılan hipotetik veri seti	13
Çizelge 3.5. Meme kanserinden şüphelenilen 20211 hastaya ait fizik muayene ve mamografi sonuçları	13
Çizelge 3.6. Meme kanserinden şüphelenilen ve biyopsi yapılan hastaların, fizik muayene ve mamografi sonuçları	14
Çizelge 3.7. Simülasyon ile elde edilen veri seti (1: işaretlenen, 0: işaretlenmeyen)	14
Çizelge 3.8. Lincoln-Petersen tahmin yönteminde seçim	21
Çizelge 3.9. Jolly-Seber tahmin yönteminde seçim	24
Çizelge 3.10. Schnabel ve Schumacher-Eschmeyer tahmin yönteminde seçim	26
Çizelge 3.11. Korelasyon yanlılığı gösterimi	30
Çizelge 3.12. Tarama testleri için yakala tekrar yakala yaklaşımı	33
Çizelge 3.13. İki yönlü tablo için gözlenen frekanslar	37
Çizelge 3.14. $k=3$ için yakala tekrar yakala dizaynı	38
Çizelge 3.15. Üç yönlü tablo için gözlenen frekanslar	38
Çizelge 3.16. Üç boyutlu çapraz tablo için bağımsız logaritmik doğrusal modeller ve serbestlik dereceleri	41
Çizelge 4.1: Lincoln-Petersen için yakalama sonuçları	44
Çizelge 4.2. Jolly-Seber için yakalama sonuçları	46
Çizelge 4.3. Schnabel için yakalama sonuçları	47
Çizelge 4.4. Schumacher-Eschmeyer için yakalama sonuçları	48

Çizelge 4.5. Meme kanserinden şüphelenilen 20211 hastaya ait fizik muayene ve mamografi sonuçları	49
Çizelge 4.6. Meme kanserinden şüphelenilen ve biyopsi yapılan hastaların fizik muayene ve mamografi sonuçları.	50
Çizelge 4.7. Logaritmik doğrusal model için simülasyon sonucu elde edilen veriler.....	51
Çizelge 4.8. Tüm yakalamalar için logaritmik doğrusal model sonuçları.....	52
Çizelge 4.9. $\hat{y} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AC} + \lambda_{jkc}^{BC}$ modeli kullanılarak elde edilen tahmini frekanslar.	53
Çizelge 4.10. Çalışmalarda incelenen hastalık veya tıbbi durumların dağılımı.	55
Çizelge 4.11. Çalışmalarda kullanılan modellerin dağılımı	58
Çizelge 4.12. Çalışmalarda incelenen popülasyon tiplerinin dağılımı	58
Çizelge 4.13. Çalışmalarda tahmin edilen istatistiklerin dağılımı.....	58
Çizelge 4.14. Çalışmalarda incelenen örnek genişliklerinin dağılımı.	59
Çizelge 4.15. Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp araştırmalarının, çalışılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.	60

SİMGELER VE KISALTMALAR DİZİNİ

α	: Alfa
θ	: Teta
LP	: Lincoln-Petersen
MLE	: Maximum Likelihood Estimation
AIC	: Akeike Information Criteria
BIC	: Bayesian Information Criteria
CRC	: Capture Recapture
YTY	: Yakala Tekrar Yakala

ÖZET

Yakala Tekrar Yakala Yöntemiyle Prevalans ve Diagnostik Doğruluk Tahmini

Hastalık değerlendirmelerinde, özellikle insidans ve prevalans çalışmalarında kullanılan en yaygın yol liste kayıtlarını kullanmaktır. Ancak genellikle kullanılacak tam bir liste yoktur. Birden çok listenin birbiriyle çakışan kayıtlara ve eksik bilgilere sahip olması sebebiyle de yakala tekrar yakala (YTY) yöntemi kullanılmaktadır. Uzun yıllar ekolojistler tarafından yaban hayatındaki hayvanların sayımında kullanılan yakala tekrar yakala yöntemi, son zamanlarda epidemiyologlar tarafından hastalık prevalans ve insidans belirleme çalışmalarında ya da evsizler gibi özel popülasyonların belirlenmesinde kullanılmaya başlanmıştır. Değişen koşullara göre kullanılacak yakala tekrar yakala yöntemi de değişiklik göstermektedir. Bunlar Lincoln-Petersen, Jolly-Seber, Schanabel ve Schumacher-Eschmeyer yöntemleridir. Ayrıca kaynak bağımlılığını incelemek için logaritmik doğrusal modellere de değinilmiştir.

Sensitivite ve spesifite değerleri düşük, referans tanı testleri sonrasında altın standart ile değerlendirilmesine karar verilen hasta sayısında yapılabilecek hataları önlemek adına yakala tekrar yakala yönteminin kullanımı oldukça pratiktir.

Bu çalışmada tüm yöntemlerin nasıl uygulanacağı hipotetik ve benzetim tekniğiyle üretilen veriler üstünde gösterilmiştir. Yakala tekrar yakala yöntemiyle ilgili günümüze kadar yapılan yayınlar taranmış, sağlıkla ilgili olanlar çalışılan hastalık/tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama veya liste sayısı ve örnek genişlikleri bakımından incelenmiştir. Tarama sonucu elde edilen sağlıkla ilgili 255 çalışmanın yaklaşık %38'inde insidans, %28'inde prevalans ve %27'sinde mortalite hesabı yapıldığı görülmüştür.

Anahtar sözcükler: Yakala tekrar yakala, prevalans, diagnostik doğruluk, Lincoln-Petersen, logaritmik doğrusal modeller.

ABSTRACT

Estimated Prevalence and Diagnostic Accuracy through Capture Recapture Method

In the disease assessments, especially the incidence and prevalence studies, the most common way is to use the list records. However, usually there is not a complete list to use. Since multiple lists have overlapping and missing information, capture recapture method (CRC) is used. CRC method that has been used for census of wildlife animals by ecologist for many years, has recently also been employed by epidemiologists for determining the prevalence and incidence of the disease or a specific population, such as homeless. CRC method to be used for changing conditions also varies. These could be listed as Lincoln-Petersen, Jolly-Seber, Schanabel and Schumacher-Eschmeyer methods. Additionally, log-linear models can also be applied for evaluating the resource dependency.

CRC method is practical to apply in order to prevent in the patient number that is assessed via gold standard following the diagnostic/reference tests with low sensitivity and specificity values.

In this study, above mentioned methods are explained via hypothetical and simulation data. Accordingly all the up to date studies regarding CRC method have been examined and health related studies are analyzed with respect to medical conditions, population type, estimation method used, number of catches or lists, and sample size. As a result of scanning, it is observed that 38% of 255 health related studies are examined for incidence, 28% for prevalence and 27% for mortality calculation.

Key words: Capture recapture, prevalence, diagnostic accuracy, Lincoln-Petersen, log-linear models.

1. GİRİŞ

İlgilenilen bir popülasyondaki birey sayısının belirlenmesinde tek tek saymak oldukça zaman alıcı, yorucu ve pahalı bir yoldur. Ayrıca örnekleme seçilen bireylerin sabit tutulamadığı durumlar söz konusu ise yapılan tahminlerin ne kadar güvenilir olduğu tartışılmalıdır.

Örneğin herhangi bir bölgede bulunan belirli bir kuş türü popülasyonunun, herhangi bir denizde bulunan balık popülasyonunun, herhangi bir bölgede belirli bir hastalığı taşıyanların sayısı belirlenmek istendiğinde sayılan her olguyu tekrar sayma ihtimalini ortadan kaldırmak için belli bir yerde tutmak gerekir. Bunun gibi olguları sabit tutamadığımız durumlar için geliştirilmiş popülasyon genişliği tahmin yöntemi Yakala Tekrar Yakala (Capture Recapture, mark-recapture, tag-recapture veya multiple record system) yöntemidir.

Yakala tekrar yakala yönteminin tarihsel gelişimine göz atacak olursak; ilk olarak John Graunt 1660'larda İngiltere'nin popülasyonunu tahmin etmek için ve benzer şekilde Pierre Simon Laplace 1786 yılında hali hazırdaki metotları takip ederek Fransa popülasyonunu tahmin etmek için değinmişlerdir. Fakat tam anlamıyla 1896 yılında balık popülasyonu çalışmasında Johannes Petersen tarafından uygulanmıştır (1). Teorik bilgilerde ki bazı eksiklikler ise 1930 yılında Frederick Lincoln'un bu yöntemi ele almasıyla tamamlanmıştır.

Yakala tekrar yakala yöntemi geleneksel olarak biyometride özellikle de hayvan popülasyonu kestiriminde kullanılmaktadır. Son zamanlarda bu teknik sağlık çalışmalarında, popülasyondaki hastalık ve sakatlık oranlarının daha doğru belirlenmesinde artarak kullanılmaktadır (2). Verimsiz, pahalı ve neredeyse imkansız olan evrensel sayım ve geleneksel epidemiyolojik sürveyans sistemlerine alternatif olarak yakala tekrar yakala yöntemi etkili ve uygun maliyetlidir. Yakala tekrar yakala metotlarının diyabet, çeşitli kanserler, HIV, felç, bağırsak hastalıkları ve tüberküloz gibi tıbbi durumların prevalanslarının tahmininde, ayrıca uyuşturucu madde kullananlar, hayat kadınları, teröristler ve evsizler gibi gizli popülasyon tahmininde kullanımı giderek popüler hale gelmektedir.

Uygulamada genellikle hastane listeleri, tedavi merkezi kayıtları veya eczane kayıtları gibi tamamlanmamış iki ya da üç kaynak (liste) birleştirilir. Listeler birleştirilirken hastanın isim-soyisim baş harfleri ya da doğum tarihi gibi bilgiler kullanılarak çakışmalar belirlenebilir. Önemli olan en mükemmel birleştirmeyi yapmak için kullanılacak en doğru bilgi kriterini belirlemektir. Aksi halde elde edilen tahminlerin geçerliliği etkilenebilir (3).

2. GENEL BİLGİLER

2.1. Örneklem Nedir?

Herhangi bir olayı incelemek için popülasyonu temsil edebilecek birimlerin belli kurallar çerçevesinde popülasyondan seçilme işlemine örneklem denir. Popülasyonun tamamıyla çalışmak çoğu zaman imkansız, zor veya pahalıdır. Eğer popülasyonu temsil edebilecek yeterli örneklem seçilirse, elde edilen sonuçlar genellenebilir. Seçim işlemini yaparken olasılıklı ve olasılıklı olmayan örneklem yöntemlerinden faydalanılır.

2.1.1. Olasılıklı Örneklem Yöntemleri

Popülasyondaki birimlerin eşit olasılıklarla ve rastgele bir şekilde örneklem seçildikleri yöntemlerdir. Bunlar basit rastgele örneklem yöntemi, sistematik örneklem yöntemi, tabakalı rastgele örneklem yöntemi ve küme örneklem yöntemidir.

2.1.1.1. Basit Rastgele Örneklem

Seçilecek örneklemdeki rastgeleliği sağlamak için rastgele üretilmiş sayıların kullanıldığı olasılıklı bir örneklem yöntemidir. Bunun için 1927 yılında istatistikçi Tippett tarafından geliştirilen “Rastgele Sayılar Tablosu” yaygın olarak kullanılmakla birlikte, istatistik yazılımlarındaki gelişmelere bağlı olarak rastgele sayı üreten programlar da sıklıkla kullanılmaya başlanmıştır. Basit rastgele örneklemede, popülasyon çok büyük değilse seçim işlemi kolaydır. Ancak bu yöntemin kullanılabilmesi için popülasyondaki birimlerin listelenmesi gerekmektedir (4).

Popülasyon çok büyük ise listelemek ve seçim işlemi zorlaşacağından basit rastgele yerine sistematik örneklemenin kullanımı daha uygundur (5). Bununla birlikte, popülasyon dağınık bir şekilde bulunuyorsa birimlere ulaşmak, araştırmanın uygulama

aşaması ve sonuçların değerlendirilmesi zorlaşacağından basit rastgele yerine küme örnekleme kullanımı daha uygundur (5). Eğer popülasyondaki birimler farklı özelliklerine göre sınıflanmışlarsa sonuçlar tüm özellikleri yansıtmayabilir. Bu durumda ise popülasyon farklılık gösteren özellik bakımından sınıflara ayrılır ve basit rastgele örnekleme yerine tabakalı rastgele örnekleme yöntemi kullanılır (5).

2.1.1.2. Sistemik Örnekleme

Popülasyonun çok büyük olduğu ve basit rastgele örnekleme yöntemiyle seçim yapmanın çok zor olduğu durumlarda sistemik örnekleme yöntemi kullanılır. Bu yöntem daha çok, işçi kayıtları, hasta ve hastalık kayıtları gibi birim sayısı çok fazla olan popülasyonlarda tercih edilir (5). N birimlik popülasyondan n birimlik örneklem seçimi yapılacağına başlangıç sayısı olarak rastgele sayılar tablosundan, 1 ile N/n arasında bir sayı seçilir ve her N/n kişide bir kişi örnekleme alınır.

Başlangıç sayısı bazı çalışmalar için oldukça önemlidir. Özellikle örnekleme seçilecek kayıtlar küçükten büyüğe veya eskiden yeniye şeklinde sıralanmışlarsa her bir başlangıç sayısında farklı sonuçlar elde edilebilir. Üzerinde çalışılacak kayıtlar haftalık veriler gibi periyodik bir döngü içerisindeyse de her hafta aynı günün seçilmesiyle sonuçlar yanlı olabilir (5).

2.1.1.3. Tabakalı Rastgele Örnekleme

Popülasyondaki birimler bazı özelliklere (cinsiyet, sosyo-ekonomik düzey vb.) göre farklılık gösteriyorsa seçilecek örneklem bu farklılığı yansıtabilecek şekilde seçilmelidir. Bu durumda tabakalı rastgele örnekleme yöntemi kullanılır (4). İlgilenilen özelliğe ait tabakalar belirlendikten sonra n birimlik örneklem her tabakadaki birim sayısı ile orantılı olacak şekilde rastgele seçilir. Bu yöntemde oluşturulacak tabakaların kendi içlerinde homojen, aralarında heterojen olması istenmektedir. N_i , i. tabakadaki birim sayısı olmak üzere, i. tabakanın ağırlığı N_i/N olarak hesaplanmaktadır.

2.1.1.4. Küme Örneklemesi

İlgilenilen popülasyon çok dağınık bir bölgede yerleşmiş ise yapılacak araştırmanın sonuçları pahalı ve zor koşullar altında elde edilebilir. Bu sebepten istenilen sonuca daha az sayıda çalışanla, daha az zamanda ve daha ucuz yollarla ulaşabilmek için küme örneklemesi yöntemi kullanılır (6). Her ne kadar ekonomi ve zaman yönünden avantaj sağlasa da elde edilen sonuçlardaki hata payı diğer örneklemesi yöntemlerine göre daha büyüktür. Oluşturulan kümeler her bölge birimini temsil edebilecek nitelikte olmalıdır. Kümeler küçük ve mümkün olduğunca aynı büyüklükte olmalıdır. Bu örneklemesi yönteminde ise kümeler arası homojenlik, küme içinde heterojenlik olması istenmektedir. Hangi kümenin örneklemesi alınacağına rastgele sayılar tablosu kullanılarak basit rastgele örneklemesi ile karar verilebilir. Kümeler belirli bir özelliğe göre tabakalandırılarak tabakalı örneklemesi yöntemi kullanılarak da seçilebilir (5).

2.1.2. Olasılıksız Örneklemesi Yöntemleri

Araştırmada zaman, masraf, örneklem çerçevesinin yetersizliği, ilgilenilen özelliğin popülasyonda az görülmesi ya da anket çalışmalarında cevapsızlık oranının çok fazla olması gibi sebeplerden dolayı olasılıklı örneklemesi yöntemleri kullanılamayabilir (7). Çalışılan bazı konularda gönüllü katılımcıların dışında kimse çalışmaya alınamayacağından istenirse de olasılıklı örneklemesi yöntemleriyle çalışılamayabilir.

Hasta memnuniyet anketlerinin çoğu kez rastgele yapıldığı savunulsa da, tam rastgele seçim söz konusu olmadığından ve hastaneye gelen tüm hastaların seçilme şansı eşit olmadığından yapılan örneklemesi çoğunlukla olasılıksız örneklemesidir.

Herhangi bir olasılık kuralına dayanmayan, deneklerin çalışmayı yapan kişi tarafından seçildiği örneklemesi yöntemlerine olasılıksız örneklemesi denir. Örneklemeye giren deneklerin seçilme olasılıkları belli olmadığından varyans hesabı yapılamaz, bu da örneklemesi hatasını etkiler. Seçim süreci önyargılı olabileceğinden, tahminler ancak

subjektif olarak deęerlendirilebilirler. Bu sebeple bilimsel arařtırmalarda bu yöntemler kolay olmasına raęmen tercih edilmezler.

Olasılıksız örnekleme yöntemlerinin,

- Arařtırmacı sınırlı bütçe, zaman ve işgücüne sahipse,
- Popülasyon belli bir özellięe sahip olduęu zaman,
- Arařtırmacı pilot çalışma ya da keşif çalışması yapmayı planladığı zaman
- Popülasyon kısıtlı ve randomizasyonun mümkün olmadığı zaman,
- Arařtırmanın amacı elde edilen sonuçları tüm popülasyona genellemek olmadığında,

kullanılması önerilebilir (7).

2.1.2.1. Gelişigüzel Örnekleme

Herhangi bir konuda popülasyondaki birimlerden en çabuk rastlanılanların seçilmesiyle yapılan örnekleme yöntemidir. Ulaşılabilen her birey örnekleme alınır. İnternet ve telefonla yapılan anketler, gönüllü katılımcıların girdiğı çalışmalar bu örnekleme tipine örnek verilebilir (5).

2.1.2.2. Amaçlı Örnekleme

Popülasyon geniş bir coğrafyaya yayılmış ve ulaşılması para ve zaman yönünden sıkıntılı ise tercih edilen bir örnekleme yöntemidir. Bu örnekleme yönteminde örnekleme seçilecek birimler çalışmayı yapan kişi tarafından belirlenir. Arařtırmacı kişisel görüşü çerçevesinde N birimlik popülasyondan yeterli gördüğü n tane seçim yaparak popülasyon parametresini tahmin etmeyi amaçlar.

2.1.2.3. Kota Örneklemesi

Popülasyon çok büyük ve imkanlar sınırlı olduğunda kullanılan bir örnekleme yöntemidir. Popülasyon incelenecek nitelik açısından homojen gruplara ayrılır. Bu gruplardan popülasyon içindeki oranlarına göre birimler seçilir. Diğer olasılıksız örnekleme yöntemlerine göre popülasyonu temsil etme gücü daha yüksek bir örneklem oluşturulabilir (4).

2.1.2.4. Kartopu Örneklemesi

Genelde gizli tutulan ve nadir görülen özelliklerin araştırıldığı çalışmalarda kişisel bağlantılar ve sosyal ilişkilerden faydalanılarak yapılan örnekleme kartopu örnekleme denir. Örnekleme alınan birey, ilgilenilen konudaki başka bir birey tanıyorsa o da örnekleme alınır. Aynı işlem sıradaki için ve n birimlik keyfi örneklem oluşuncaya kadar tekrarlanır (5). Sağlık çalışmalarında bu örnekleme yöntemi daha çok kayıtların düzgün tutulmadığı veya halkın bir sağlık kuruluşuna gitmeyi reddettiği durumlarda araştırmacıların son çare olarak tercih ettikleri bir yöntemdir.

2.2. Sağlık Kayıtları Sırasında Karşılaşılan Sorunlar

Ülkelerin gelişmişlik düzeyinin önemli bir göstergesi olan doğum ve ölüm hızlarının tahmini çoğu kez büyük bir sorundur

Ülkemizde mortalite ile ilgili çalışmalarda Türkiye İstatistik Kurumu'ndan (TÜİK) elde edilen bilgiler kullanılmaktadır. Çünkü ölümlerde defin kâğıdı doldurulurken ölüm istatistik formu da doldurmak zorunludur. Ancak bazı durumlarda eksik doldurulduğu için gerçek ölüm sebeplerine ulaşmak zor olabilmektedir. Bazılarında ise hastanenin kayıtlarında farklı bir ölüm sebebi yazarken defin kâğıtlarında farklı bir sebep yazmaktadır. Bu ise sonuçların doğruluğunu ve güvenilirliğini etkilemektedir. Diğer bir önemli istatistik ise doğum kayıtlarıdır. Bölgenin gelişmişlik düzeyine göre hastanede kayıtlı olan doğumlarla nüfus

müdürlüklerindeki doğum kayıtlarının birbirini tutmadığı durumlar söz konusu olabilmektedir (8).

Tekrarlanan ve eksik veri durumunda veri setinin ve sonuçların kalitesi düşeceğinden, bu sorunu ortadan kaldırmak için yakala tekrar yakala yöntemi kullanılmaktadır.

Hasta sayıları ve hastalık tanıları gibi veriler, Sağlık Bakanlığı'na bağlı aile sağlığı merkezleri, ana çocuk sağlığı aile planlaması merkezleri, verem savaş dispanserleri gibi birinci basamak sağlık kuruluşlarından, ayrıca özel muayene ve hastanelerden toplanabilir. Ancak bilgileri aktaracak birimler sadece kendilerine gelen kişileri kayıt altına aldığından ve aynı kişi birden fazla birime gitmiş olabileceğinden tekrarlanan durumlar söz konusu olmaktadır (8).

Çalışmalarına Şubat 2011'de başlayan ve AB Komisyonu tarafından desteklenen, Avrupa Hastalarına Yönelik Akıllı Hizmetler (Smart Open Services for European Patients) portalına Türkiye'den giriş yapan doktorlar Avrupa ülkesindeki bir hastanın kimlik numarasını kullanarak sorgulamasını yapabilmekte ve Türkiye ile Avrupa arasında hasta veri kayıtlarını paylaşabilmektedir (9). Bu sorgulamada, Avrupa'dan gelen hastaya ait temel sağlık verilerine, hastanın da onayıyla Türkiye'deki bir hekim tarafından erişilebilmektedir.

Belirli bir toplumun kayıt merkezlerinden elde edilen hastalık istatistikleri, diğer toplumlarla veya farklı zamanlarla karşılaştırılabilir olmalıdır. Karşılaştırılabilirliğin temelinde ise kayıt işleminin evrensel standartlar ve kurallar çerçevesinde yapılması vardır. Kayıt altına alınacak vakalara ait değişkenlerin tanımları kayıt işlemi başlamadan önce yapılmalıdır. Değişkenlere ait birimler tüm kayıtlarda aynı olmalıdır. Kayıt listesinin doğruluğu için veriler tam (completeness) ve geçerli (validity) olmalıdır. Hedef nüfustaki bireylerin hepsi liste kayıtlarında bulunuyor ise veri tamdır ve bunu kontrol etmek için kullanılan yöntemlerden biri yakala tekrar yakala yöntemidir. Birden çok kaynak olduğu durumda kullanımı oldukça kolaydır (10).

Hastalıklara ilişkin bilgi toplamak ve hastalığa sahip kişilerde izlem kolaylığı için sağlık kuruluşlarındaki hasta vakalar ve ölüm vakaları kayıt altına alınır. Kayıt işlemi hastaneler, klinikler, laboratuvarlar, toplum sağlığı merkezleri gibi sağlık birimlerinde ve ölüm kayıt formları, adli tıp kurumları gibi birimlerde tutulmaktadır. Bu sayede hastalığa sahip ve hastalıktan ölmüş vakaların kayıtları birleştirilebilir ve

karşılaştırılabilir. Çalışma bir hastane ya da klinik merkezli olabilir. Herhangi bir il, ilçe ya da ülke merkez alınarak da bilgiler toplanabilir. İlgilenilen hastalığın toplum üzerindeki etkisini, ortaya çıkış süresini, şeklini ve sonuçlarını bu kayıtlar sayesinde elde edebiliriz. Özellikle aynı hastalık üzerinde uygulanan değişik tedavilerin etkinliğinin karşılaştırmasında oldukça yararlıdır.

Hastalığın toplumdaki durumunu ortaya koymaya çalışan analitik ve tanımlayıcı araştırmalar için seçilecek örneklemin kayıtlardan elde edilmesi hatayı en aza indirmeye yardımcı olur.

Tüm merkezlerin yanı sıra kanser vakalarıyla ilgili bilgileri kanser kayıt merkezleri kaydetmektedir. Ölüm kayıt formlarının dahil edildiği çalışmalarda ilgilenilen hastalığa ait insidans hızı gerçeğinden daha düşük olacaktır. Çünkü hastalığa ne zaman yakalandığıyla ilgili bilgiye bu formlarda yer verilememektedir (10).

Belirli bir hastalık için büyük popülasyonların taranması, hastalık surveyansı ve koruyucu hekimlik için oldukça önemlidir. Meme kanseri, prostat kanseri, rahim kanseri, kardiyovasküler hastalıklar ve diyabet gibi pek çok hastalık toplum sağlığı açısından önemlidir ve gelişmiş ülkelerde popülasyonların bu hastalıklar bakımından taranması gerekmektedir. hastalık taraması sırasında öne çıkan en önemli sorun bazı hastalıklar için altın standartın olmayışı olarak bilinse de son yıllarda özellikle biyoistatistikçiler tarafından üzerinde önemle durulan konu diagnostik değerlendirme yanlılığıdır.

Tarama testleri kullanılırken karşılaşılan pek çok yanlılıktan sadece birisi olan doğrulama yanlılığı (verification bias) seçilmiş denek kullanmaktan kaynaklanır. İki tanı testinden birinin pozitif sonuç verdiği bir hastaya ileri bir tarama testi uygulanırken ikisinin de negatif sonuç verdiği bir hastaya ileri test uygulanmaz. Bu durumda iki testin de negatif sonuç verdiği hastaların durumu doğrulanmadığından tanı testlerinin performansı (doğruluk, sensitivite, spesifite) değerlendirilirken yanlı sonuçlar elde edilir. Yakala tekrar yakala tahmin yöntemi kullanılarak sonucu doğrulanamayan bireylerin ne kadarının hasta olduğu tahminlenebilir.

2.3. Sağlık İstatistikleri İçin Veri Toplama Yöntemleri

2.3.1. Aktif Yöntem

Kayıt merkezinin personelinin (kanser kayıt elemanlarının) sağlık kurumlarıyla doğrudan ilişki kurarak ilgilenilen olguları saptaması ve bu olgulara ait bilgileri standartlarına uygun olarak derlemesidir. Burada verinin doğrudan kaynaktan elde edilmesi söz konusudur. Bu yöntemdeki en önemli birey ilgilenilen konuyla ilgili eğitim gören kayıt elemanlarıdır. Çünkü verinin doğruluğu ve güvenilirliği kayıt elemanına bağlıdır (10).

2.3.2. Pasif Yöntem

İlgilenilen olgulara ait verilerin hastane, klinik, laboratuvar gibi ilgili kurumda çalışan personel tarafından veri toplama merkezlerine kendiliğinden gönderilmesi yöntemidir. Verinin tam ve geçerli olması açısından sıkıntılı bir yöntemdir.

2.3.3. Aktarım

Toplanan ve elektronik ortama girilen verilerin bilgisayar ağı kullanılarak kayıt merkezinin verileri arasına dahil edildiği veri toplama yöntemine aktarım (linkage) denir. Mersin Üniversitesi Araştırma ve Uygulama Hastanesi'nde kullanılan NUCLEUS bilgi sistemi buna bir örnektir. Hastaların genetik, radyolojik, histolojik, patolojik, fizik muayene gibi değerlendirme sonuçları bilgisayar ağı kullanılarak NUCLEUS sistemine kaydedilmektedir. Çoğu gelişmiş ülkelerde de bu yöntem kullanılmaktadır. Ancak gelişmekte olan ülkelerin çoğunda nüfusun tamamına ait güvenilir bilgi yoktur. Ayrıca cinsiyete, yaşa, etnik gruplara göre tabakalandırma sağlıklı olmayabilir. Bunların yanı sıra gelişmekte olan ülkelerde çok çeşitli nedenlerle nüfus hareketleri olmakta, sürekli göçler yaşanmaktadır. Çoğu zaman bu hareketler kaydedilememektedir. Tekrar eden

kayıtları saptayabilmek için kişileri ayırt edebilmek gereklidir. Nüfusun hareketliliğinden ayrı olarak kişileri birbirinden ayırt etmek de çok zordur. Gelişmiş ülkelerde her kişinin bir kimlik numarası vardır. Oysa az gelişmiş ülkelerin çoğunda böyle bir sistem olmadığından kişileri ayırt etmede isim, soyisim, baba ismi gibi kriterler kullanılır. Bu da tam bir ayırt etme sağlamayabilir (10).

3. MATERİYAL VE YÖNTEM

3.1. Materyal

Materyal 1: Yakala tekrar yakala yöntemi kullanılarak eksik veri ve popülasyon genişliğinin tahmin aşamalarını Lincoln-Petersen, Jolly Seber, Schnabel ve Schumacher yöntemlerinin her biri için ayrı ayrı göstermek amacıyla veri setleri üretilmiştir.

Lincoln-Petersen tahmin yönteminin kullanılmasını göstermek amacıyla, $p=0.50$, $n=100$ ve yakalama ya da liste sayısı $k=2$ olacak şekilde Ek'deki makro ile veri simülasyonu yapılmıştır. Elde edilen veri seti Çizelge 3.1'de verilmiştir.

Çizelge 3.1. Lincoln-Petersen tahmini için simülasyon ile elde edilen veri seti (1: işaretlenen, 0: işaretlenmeyen)

1. Liste	2. Liste	Sayı
1	1	28
1	0	38
0	1	29
0	0	?

Jolly-Seber tahmin yönteminin kullanılmasını göstermek amacıyla, Çizelge 3.2'deki hipotetik veri seti kullanılmıştır.

Çizelge 3.2. Jolly-Seber tahmini için kullanılan hipotetik veri seti

t	Toplam	İşaretli Yakalanan	Yeni İşaretlenen
1. Yakalama	50	-	50
2. Yakalama	55	14	41
3. Yakalama	38	14	24
4. Yakalama	45	25	20
5. Yakalama	39	25	14

Schnabel tahmin yönteminin kullanılmasını göstermek amacıyla, Çizelge 3.3'deki hipotetik veri seti kullanılmıştır.

Çizelge 3.3. Schnabel tahmini için kullanılan hipotetik veri seti

t	Toplam	İşaretli Yakalanan	Yeni İşaretlenen
1. Yakalama	39	-	39
2. Yakalama	43	8	35
3. Yakalama	44	15	29
4. Yakalama	44	19	25
5. Yakalama	41	24	17
6. Yakalama	45	26	19

Çizelge 3.4. Schumacher-Eschmeyer tahmini için kullanılan hipotetik veri seti

t	Toplam	İşaretli Yakalanan	Yeni İşaretlenen
1. Yakalama	39	-	39
2. Yakalama	43	10	33
3. Yakalama	44	8	36
4. Yakalama	44	11	33
5. Yakalama	41	9	32
6. Yakalama	45	5	40

Tarama testlerinin doğruluk tahmini için yöntemin kullanılmasını göstermek amacıyla, Strax, Venet, Shapiro ve Gross'un (1967) çalışmasına ait veri seti örnek olarak seçilmiştir (11). Bu çalışmada meme kanseri şüphesiyle hastaneye başvuran 20211 kadın test 1 (Fiziksel muayene) ve test 2 (mamografi) ile taranarak en az bir testten pozitif sonuç alanlara biyopsi yapılmıştır. Çalışma sonuçları Çizelge 3.5 ve Çizelge 3.6'da verilmiştir.

Çizelge 3.5. Meme kanserinden şüphelenilen 20211 hastaya ait fizik muayene ve mamografi sonuçları

		Mamografi	
		+	-
Fiziksel Muayene	+	23	168
	-	116	19904

Çizelge 3.6. Meme kanserinden şüphelenilen ve biyopsi yapılan hastaların, fizik muayene ve mamografi sonuçları

Biyopsi ile sağlıklı bulunanlar

Biyopsi ile hasta bulunanlar

		Mamografi	
		+	-
Fiziksel Muayene	+	13	144
	-	95	$n_{00}^1=?$

		Mamografi	
		+	-
Fiziksel Muayene	+	10	24
	-	21	$n_{00}^2=?$

Materyal 2: Prevalans tahmininin logaritmik doğrusal modeller kullanılarak analizini ve sonuçların yorumlanışını göstermek amacıyla, yakalama oranı $p=0,30$, $n=50$ ve yakalama sayısı $k=3$ olacak şekilde Ek'deki makro ile veri simülasyonu yapılmıştır. Elde edilen veri seti Çizelge 3.7'de verilmiştir.

Çizelge 3.7. Simülasyon ile elde edilen veri seti (1: işaretlenen, 0: işaretlenmeyen)

1. Yakalama	2. Yakalama	3. Yakalama	Sayı
1	1	1	11
1	1	0	1
1	0	1	3
1	0	0	3
0	1	1	12
0	1	0	26
0	0	1	6
0	0	0	?

Materyal 3: Çalışma kapsamına dahil edilmek üzere Mersin Üniversitesi internet erişim ağı üzerinden “capture recapture” anahtar tümcüsüyle Pubmed veri tabanında sistematik olarak geriye dönük bir tarama yapılmıştır.

Dahil edilme kriterleri;

1. Çalışmanın tam metin olması,
 2. Materyali insan olan uygulamalı tıbbi bir çalışma olması,
- olarak belirlenmiştir.

Dışlanma kriterleri;

1. Çalışmanın tam metin olmaması,
2. Çalışmanın uygulamalı ve tıbbi olmaması (teorik çalışma, biyolojik çalışma, hayvan çalışması, derleme, gözden geçirme, editöre mektup)
3. Uygulamalı, tıbbi fakat gen çalışması olması,

olarak belirlenmiştir.

3.2. Yöntem

3.2.1. Yakala Tekrar Yakala (Capture-Recapture) Yöntemi

Veri kalitesi, en uygun örneklemin oluşturulmasıyla doğru orantılıdır. Örneklem oluşturmak için popülasyonun doğru belirlenmesi ve ilgili tüm birimleri içermesi gerekmektedir. Bazı durumlarda popülasyonun içerdiği bireyleri saptamada zorluklar yaşanabilir. Bunlardan biri popülasyondaki birim sayısının tam olarak bilinmemesidir. Bu sorunu gidermede yakala tekrar yakala (capture-recapture) tahmin yönteminden faydalanılmaktadır. Bazı bilimsel araştırmalarda örnekleme yöntemi olarak belirtilen yöntem, tekrar tekrar örneklenecek sayması güç olan bir popülasyonun asıl sayısının tahmin edildiği bir yöntemdir. Esas kullanım alanı vahşi hayvan popülasyonlarının tahmin edilmesi olsa da, nüfus sayımında, doğum ve ölüm hızlarının hesaplanması gibi epidemiyolojik araştırmalarda da kullanılan bir örnekleme yöntemidir. Takip edilecek popülasyona ait iki ve ikiden çok izlem sistemi veya rastgele seçilen yakalamaların çakışmasından elde edilen sayı ile toplam sayı tahmin edilir. İkili yakalama (örneklem) söz konusu olduğunda, ilgilenilen popülasyondan bir örnek grubu yakalanır, sayılır ve işaretlendikten sonra popülasyona tekrar bırakılarak diğer birimlerle yeterince karışması sağlanır. Daha sonra aynı işlem ikinci örnek grubu için işaretli birimler ve ilk defa yakalanan birimler olmak üzere sayılarak elde edilen örneklemden popülasyon genişliği tahminlenebilir (12).

Hastalık değerlendirmelerinde, özellikle insidans ve prevalans çalışmalarında kullanılan en yaygın yol liste kayıtlarını kullanmaktır. Ancak genellikle kullanılacak

tam bir liste yoktur. Birden çok listenin birbiriyle çakışan kayıtlara ve eksik bilgilere sahip olması sebebiyle de yakala tekrar yakala yöntemi kullanılmaktadır.

Bu yöntem, hayvan popülasyonunu belirleme ihtiyacıyla ortaya çıkmıştır. Daha sonra demografik bir sayım olan Londra nüfusunu belirlemek için 1662 yılında Graunt tarafından kullanılmıştır. Benzer şekilde Laplace bundan 124 yıl sonra yöntemin teorik kısmındaki matematiksel ifadeleri düzenlemiş ve Fransa'nın nüfusunu belirlemek için kullanmıştır (1786). Fakat tam anlamıyla 1896 yılında balık popülasyonu çalışmasında Petersen tarafından uygulanmıştır (1). Teorik bilgilerdeki bazı eksiklikler ise 1930 yılında Lincoln'un bu yöntemi ele almasıyla tamamlanmıştır. Daha sonra Chandrasekar ve Deming tarafından 1949 yılında, Hindistan nüfusundaki doğum ve ölüm oranlarını belirlemek için kullanılmıştır (13).

Liste kaynaklarının kullanılması yöntemi 1968 yılında Wittes tarafından çalışılmıştır. Cormack (1968, 1972, 1981, 1989), El-Khorazaty ve arkadaşları (1977), Burnham ve Overton (1979), Chao (1986), Darroch ve arkadaşları (1993) popülasyon genişliği belirleme tahmini üzerine çalışmışlardır (14).

Uzun yıllar ekolojistler tarafından yaban hayatındaki hayvanların sayımında kullanılan yakala tekrar yakala yöntemi, son zamanlarda epidemiyologlar tarafından hastalık prevelans ve insidans belirleme çalışmalarında ya da evsizler gibi özel popülasyonların belirlenmesinde kullanılmaya başlanmıştır (15).

Toplumda bireyler isim, soy isim, cinsiyet, doğum tarihi ve sosyal sağlık numarası gibi özellikleriyle işaretlenmektedir. Birden çok yerden bilgi sağlanan bir çalışmadaki listelerde çakışan birimler belirlenebilirken, hiçbir kaynaktan yer almayan birim belirlenmemektedir. Ayrıca cevapsızlık oranının yüksek olduğu anket çalışmalarında eksik veriyi tamamlamak için de yakala tekrar yakala yönteminden faydalanılmaktadır (15).

3.2.1.1. Yakala Tekrar Yakala Yönteminin Varsayımları

İşaretleme: İşaretlemler kaybolmayacak, kalıcı şekilde yapılmalıdır. (balık çalışmasında suda çözünen boya kullanılmaması, hasta protokol numarasının elektronik ortama da kaydedilmesi gibi.) (16)

Kaynak (Seçim) Sayısı: İncelenecek durum veya vaka için birden fazla kaynak (liste) ya da tek bir topluluktan birden fazla rastgele örneklem (yakalama) seçimi olmalı, (bir hasta birden fazla hastanede birden fazla ilde görülmüş olabilir.)

Rastgele örnekleme: Yakalama işleminde birimler eşit şansa sahip olmalı, yani sağlıklı olan birimin seçilme şansı diğerlerinden yüksek olmamalıdır. Kaynaklar, toplumdaki vakaları rastgele yakalamış olmalıdır. Yakalanması zor ya da imkansız olan gizli/kısıtlı popülasyon içermemelidir. Gizli popülasyon rastlanılan topluluklar:

- Enjekte uyuşturucu, afyon, kokain kullananlar, (sero-yaygınlık)
- Cezaevindeki mahkumlar,
- Hayat kadınları (yabancı uyruklular),
- Aids gibi sıra dışı hastalıklara sahip olanlar,
- Eşcinseller,
- Perinatal dönem bebekleri (17).

Gizli popülasyonla yapılacak çalışmalarda daha çok anket gibi kesitsel çalışma tipi planlanır. Yapılan çalışma sonucunu daha geniş bir popülasyona genellemek, gizli popülasyon durumunda neredeyse imkansızdır.

Tanımlayıcı değişken: Vakalar tek tanımlayıcıyla eşleştirilebilir olmalı (örneğin tanımlayıcı olarak *hasta ismi* belirlenmiş ise aynı isme sahip birden çok birey olabilir ve ayırt etmekte zorluk çıkabilir, oysa tanımlayıcı olarak hasta protokol numarası seçilmiş olursa bireylerde karışıklık oluşmaz.)

Nüfus Hareketliliği: Eğer toplumda nüfus hareketliliği fazla ise, en az dört bağımsız yakalama işlemi gerçekleştirilmelidir.

Bağımsızlık: Birimin ilk örnekleme yakalanma şansı ile bir sonrakilerde yakalanma şansı eşit ve bağımsız olmalıdır. Bir bilginin bir kayıt formunda bulunma olasılığı diğer grupta bulunma olasılığını etkilememelidir. Bazı durumlarda her birimin yakalanma şansının eşit olmadığı bazı araştırmacılar tarafından belirtilmiştir. Mesela hem ölüm kayıt verileri hem de hastaneden taburcu olma verilerinin kullanıldığı kanser hasta sayısını hesaplamak amacıyla yapılan bir çalışmada, ağır kanser vakalarının daha çok hastaneye başvurabileceği ve daha fazla ölüm görülebileceği varsayımından yola çıkarak, bu kişilerin yakalanma olasılıklarının diğer kanser hastalarına göre daha fazla olabileceği (overestimate) bildirilmiştir.

Örnekleme çerçevesi belirleme yanlılığından (sampling frame bias) kaynaklanan bu hata, sonuçlara pozitif karıştırıcı (confounder) etkisinde bulunarak yüksek tahmine (overestimation) sebep olur. Benzer şekilde negatif karıştırıcı etkisindeki tahminler ise düşük (underestimation) sonuçlar verir (18).

3.2.1.2. Açık Popülasyon

Doğal ve daha sık rastlanılan popülasyonlardır. Bu popülasyon tipinde doğum, ölüm, iç ve dış göç oluşabilmekte, coğrafi sınırlar değişebilmektedir (19). Açık popülasyon genelde uzun zaman alacak çalışmalar için tercih edilmektedir. Örneğin diyabet teşhisi almış hastalar üzerinde yapılan bir çalışmaya hala tedavi altında olan hastalar da dahil edilirse popülasyon tipi açık olmaktadır.

Açık popülasyonların kullanıldığı en popüler model Jolly Seber (Jolly 1965; Seber 1965, 1982; Pollock et al. 1990; Schwarz and Seber 1999) modelidir. Açık popülasyon yakala tekrar yakala modellerinin ilk temellerini ise Jackson 1939, 1940, 1944, 1948; Fisher ve Ford 1947; Leslie ve Chitty 1951; Leslie ve ark. 1953 yıllarında yaptıkları çalışmalarla atmışlardır.

3.2.1.3. Kapalı Popülasyon

Çalışılan zaman periyodu içinde iç ve dış göçün göz ardı edilebilecek kadar küçük olduğu ve coğrafi sınırların değişmediği popülasyonlardır.

Kapalı yakala tekrar yakala modelinde doğum, ölüm, göç gibi zamanla popülasyon yoğunluğunu değiştiren durumların olmadığı varsayılır. Akraba evlilikleri üstüne yapılan çalışmalar, neredeyse hiç göç almayan ülkelerde (Finlandiya, Yeni Zelanda, Avustralya gibi) yapılan çalışmalar kapalı popülasyon tipine örnek verilebilir.

Demografik kapalılık varsayımı genelde, üreme olmayan bir sezon esnasında kısa bir sürede verinin toplanmasıdır. Seber (1982, 1986, 1992) ve Schwarz (1999) belirli bir hayvan topluluğunun sayısını tahmin etmek için kullanmış ve geliştirmişlerdir.

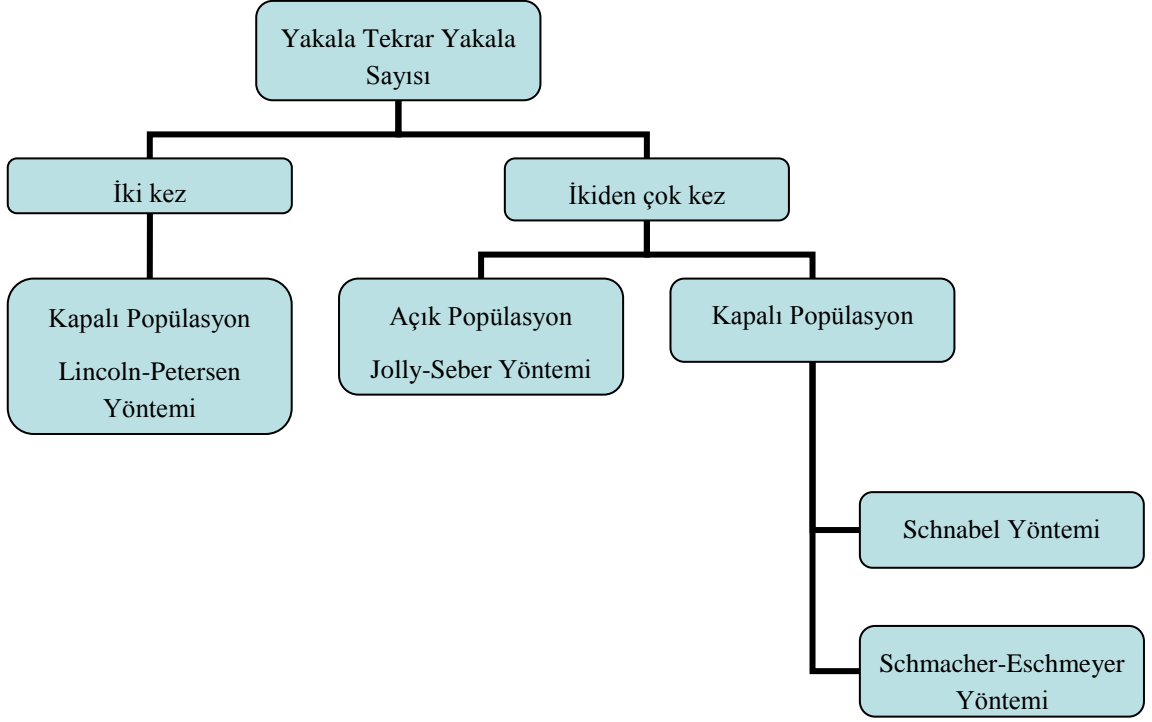
3.2.1.4. Yakala Tekrar Yakala Yönteminin Avantaj ve Dezavantajları

Yakala tekrar yakala yönteminin en önemli avantajları arasında; geçici veya kalıcı işaretler kullanılabilmesi, popülasyon genişliğindeki değişimin izlenebilmesi, matematiksel hesaplama kolaylığı, hasta kabul ve taburcu kayıt, aile hekimi listeleri gibi bilgisayarlaştırılmış listelerde popülasyon tespit yüksekliği sayılabilir.

Yakala tekrar yakala yönteminin dezavantajları olarak ise; analiz yapılırken ekstra dikkat gerektirmesi, zaman alıcı olması, özel ekipman ve eğitim gerektirmesi, denek davranışlarını etkileyerek yanlılığa sebep olması, gizlilik ilkesinin kolay ihlal edilebilmesi sayılabilir (20, 21).

3.2.2. Yakala Tekrar Yakala Yönteminin Uygulanışı

Birden çok kaynak verisi birleştirilerek ya da bir kaynaktan birden çok rastgele örneklem seçimiyle popülasyon tahmininde bulunan yakala tekrar yakala yönteminde popülasyon tipinin açık ya da kapalı olmasına göre uygulanacak model değişiklik göstermektedir. Sadece 2 kez yakalama yapılarak popülasyon tahmin edilmeye çalışılacaksa açık popülasyon bu duruma uygun değildir (22).



Şekil 3.1. Yakala tekrar yakala uygulama yöntemleri

Çok kaynaklı çalışmalar için matematikçiler tarafından önerilen liste sayısı bazı literatürlerde 3'den fazla olması yönünde iken, son zamanlarda yapılan çalışmalarda optimal sayının 3 olması önerilmektedir (23).

3.2.2.1. Lincoln-Petersen Tahmin Yöntemi

Kapalı popülasyon söz konusu olduğunda ve 2 kaynak liste birleştirileceğinde ya da bir topluluktan iki defa yakalama yapıma durumunda Lincoln-Petersen popülasyon tahmin yöntemi kullanılır.

Klasik yakala tekrar yakala tahmininde ilk olarak bir popülasyondan M sayıda denek seçilir, kayda alınır ve işaretlenerek geri iade edilir. İkinci adımda C sayıda yeni bir seçim yapılır ve bunlar birinci yakalamadan işaretli olanlar (R) ve olmayanlar (x,y) olarak kayda alınır. Elimizde olan M , C ve R ile N toplam popülasyon sayısını tahmin etmek için Lincoln-Petersen tahmin yöntemi kullanılır (23, 24).

Bunu bir çapraz tablo ile gösterecek olursak;

Çizelge 3.8. Lincoln-Petersen tahmin yönteminde seçim

		2. Seçim		Toplam
		+	-	
1. Seçim	+	R	x	M
	-	y	a	
Toplam		C		N

Lincoln-Petersen tahmin yönteminde yakalamaların birbirinden bağımsız olduğu varsayımı altında Odds oranı 1'e eşit kabul edilir. Odds oranının hesaplanmasındaki amaç farklılık tespitinden çok ilişki incelemektir. [Olasılık / 1-Olasılık] hesabıyla elde edilen odds oranında olasılık 0-1 aralığında değer alır. İlgilenilen olayın gerçekleşme olasılığı ile gerçekleşmeme olasılıklarının eşit yani birbirinden bağımsız olduğu durumda ise odds oranı 1 olarak hesaplanır. Yakalama işlemlerinde her deneğin seçilme şansının eşit olması ve herhangi bir deneğin seçilme şansının seçilmeme şansına eşit olması ile bağımsızlık varsayımı yerine getirilir. Denklem 3.1'de odds oranının 1 olarak kabul edilmesi iki yakalamanın birbiriyle ilişkili olmadığını yani bağımsızlığı gösterir.

Denklem 3.1. Yakalamaların bağımsızlığı

$$Odds = 1 = \frac{R.a}{xy}$$

Buradan her iki yakalamada da kayda alamadığımız için bilinmeyen "a" yı çekeriz:

Denklem 3.2. Bilinmeyen göze tahmini

$$a = \frac{xy}{R}$$

Tablodaki değerleri a ile toplayarak N 'yi elde ederiz.

Denklem 3.3. Lincoln-Petersen yöntemiyle popülasyon tahmini

$$N = \frac{\overbrace{xy}^a}{R} + x + y + R \rightarrow N = \frac{xy + xR + yR + R^2}{R} \rightarrow N = \frac{\overbrace{(x+R)}^M \overbrace{(y+R)}^C}{R}$$

$$N_{LP} = \frac{MC}{R}$$

Tahminin Sekar ve Deming (1949) tarafından önerilen varyansı aşağıdaki gibidir.

Denklem 3.4. Lincoln-Petersen yöntemiyle popülasyona ait varyans tahmini

$$Var(N_{LP}) = \frac{MCxy}{R^3}$$

Burada Petersen'in (1896) dikkate almadığı bir nokta, İkinci seçimde daha önce hiç işaretlenmemiş birimlerin gelmesi ve bu durumda n_{12} 'nin sıfır değerini almasıdır. Bu durumda eşitlik tanımsız hale geleceğinden, Chapman (1951) tarafından eşitlikte düzeltme yapılmıştır.

Denklem 3.5. Chapman yöntemiyle popülasyon tahmini

$$N_{Chap} = \frac{(M+1)(C+1)}{R+1} - 1$$

Seber (1970) ve Wittes (1972) tarafından da Chapman tahmin edicisine bir varyans önerilmiştir.

Denklem 3.6. Chapman yöntemiyle popülasyona ait varyans tahmini

$$Var(N_{Chap}) = \frac{(M+1)(C+1)xy}{(R+1)^2(R+2)}$$

Chapman tahmincisinin iki seçimdeki ortak birey sayısı 7'den az olduğunda kullanılması da önerilmektedir.

3.2.2.2. Jolly - Seber Tahmin Yöntemi

Jolly Seber tahmin yöntemi 1965 yılında Jolly ve Seber'in büyük toplulukları tahmin etme ilgileriyle ortaya çıkmıştır (Jolly 1965; Seber 1965, 1982, 1986, 1992; Pollock ve ark. 1990, Schwarz and Arnason 1996). Daha sonra Cormack-Jolly-Seber modelleri kullanılarak yaşam oranlarının tahmini için yakala tekrar yakala ana fikri önerülmüştür (25).

İlgilenilen popülasyonun kapalılığının sağlanmadığı yani doğum, ölüm, göç gibi faktörlerin sabitlenemediği durumda popülasyon genişliklerinin tahmininde Jolly Seber tahmin yöntemi kullanılır (20). Daha gerçekçi bir durum olan açık popülasyonda tahmin yapan bu yöntem kullanılırken yakala tekrar yakala durumu en az 3 kez tekrarlanır. Ancak bu defa her yakalamadaki bireylere farklı işaretlemeler yapılmaktadır. İşaretlenen bireyler diğer yöntemlerden farklı olarak çalışmanın sonuna kadar takip edilir. Bu metodun kullanılabilmesi için bazı varsayımların yerine gelmesi gerekmektedir:

- Her bireyin her yakalanmada seçilme şansı eşit olmalıdır.
- İşaretli her bireyin t örneklemesinden $t+1$ örneklemesine kadar yaşama şansı eşittir.
- Bireyler birbirlerini çalışmadan çıkaramazlar.
- Sayımlar sırasında işaretlemeler kaybolmaz.
- Örnekleme seçimleri arasındaki zaman göz ardı edilebilir (19).

t . örneklem sırasında çekilen n_t adet bireyden m_t tanesi daha önce yakalanmış ve işaretlenmiş, u_t tanesi de ilk defa t . örnekleme yakalanmıştır.

k : çalışmadaki örneklem sayısı ($t=1, \dots, k$)

m_t : t örnekleminde yakalanan işaretli birey sayısı

u_t : t örnekleminde yakalanan işaretli olmayan birey sayısı

n_t : t örnekleminde yakalanan toplam birey sayısı (m_t+u_t)

s_t : t örnekleminde sonra geri bırakılan birey sayısı ($n_t -$ yaralanma, ölüm ya da işaretlenme silinmesi)

r_t : t örnekleminde yakalanıp s_t olarak geri bırakılan bireylerden sonraki örneklemlerde tekrar yakalananların sayısı ($t=1, \dots, (k-1)$)

z_t : t örneklemeden önce işaretlenmiş, t örneklemesinde yakalanmamış ancak sonraki örneklemlerde yakalanmış bireylerin sayısını gösterir.

Çizelge 3.9. Jolly-Seber tahmin yönteminde seçim

t	m_t	u_t	n_t	s_t	r_t	z_t
1	0	u ₁	n ₁	s ₁	r ₁	0
2	m ₂	u ₂	n ₂	s ₂	r ₂	z ₂
.
.
k	m _k	u _k	n _k	s _k	0	0

(+1 düşük örnek genişliği durumu için düzeltme)

Denklem 3.7. Popülasyon oran tahmini

$$\hat{\alpha}_t = \frac{m_t + 1}{n_t + 1}$$

t zamanında işaretlenmiş popülasyon oranının tahmini,

Denklem 3.8. İşaretli popülasyon tahmini

$$\hat{M}_t = \frac{(s_t + 1)z_t}{r_t + 1} + m_t$$

t. örnekleme yapılmadan önceki işaretli popülasyon tahmini ve t zamanındaki popülasyon genişliğinin tahminidir. Bu tahmin ilk ve son örneklemler haricindeki tüm örneklemlerde hesaplanır.

Denklem 3.9. Jolly-Seber yöntemiyle popülasyon tahmini

$$\hat{N}_t = \frac{\hat{M}_t}{\hat{\alpha}_t}$$

Bu tahminin varyansı ise denklem 3.10 ile gösterilir (26).

Denklem 3.10. Jolly-Seber yöntemiyle popülasyona ait varyans tahmini

$$Var(\hat{N}_t) = \hat{N}_t(\hat{N}_t - n_t) \left\{ \frac{\hat{M}_t - m_t + s_t + 1}{\hat{M}_t + 1} \left(\frac{1}{r_t + 1} - \frac{1}{s_t + 1} \right) + \left(\frac{1}{m_t + 1} - \frac{1}{n_t + 1} \right) \right\}$$

$\hat{\phi}_t$; t ve t+1 zamanları arasında popülasyonun yaşama olasılığı denklem 3.11 ile hesaplanır.

Denklem 3.11. t ve t+1 zamanları arasında popülasyonun yaşama olasılığı

$$\hat{\phi}_t = \frac{\hat{M}_{t+1}}{\hat{M}_t - m_t + s_t}$$

Bu olasılığın varyansı ise denklem 3.12 ile gösterilmektedir (27, 28).

Denklem 3.12. Popülasyonun yaşam olasılığına ait varyans

$$Var(\hat{\phi}_t) = \hat{\phi}_t^2 \left\{ \frac{\hat{M}_{t+1} - m_{t+1} + 1(\hat{M}_{t+1} - m_{t+1} + s_{t+1} + 1)}{(\hat{M}_{t+1} + 1)^2} \left(\frac{1}{r_{t+1} + 1} - \frac{1}{s_{t+1} + 1} \right) + \frac{\hat{M}_t - m_t + 1}{\hat{M}_t - m_t + s_t + 1} \left(\frac{1}{r_t + 1} - \frac{1}{s_t + 1} \right) + \frac{1 - \hat{\phi}_t}{\hat{M}_{t+1} + 1} \right\}$$

\hat{B}_t ; t periyodundan t+1 periyoduna kadarki sürede yeni doğan sayısı denklem 3.13 ile gösterilir (29).

Denklem 3.13. t ve t+1 zamanları arasında yeni doğan sayısı

$$\hat{B}_t = \hat{N}_{t+1} - \hat{\phi}_t(\hat{N}_t - n_t + S_t)$$

3.2.2.3. Schnabel Tahmin Yöntemi

Kapalı popülasyonlarda Lincoln-Petersen tahmin yönteminin 2'den çok defa tekrarlanması olarak bilinen bu yöntem 1938 yılında Schnabel tarafından geliştirilmiştir. Petersen tahminini ağırlıklandırarak geliştirilen bu yöntemin kullanımında farklı yakalamalarda aynı tür işaretleme yapılabilir. Sonuç değişken olarak işaretli ve işaretli olmayan birey sayısı dikkate alınır. Ayrıca bu yöntemin kullanımı için $y=R/C$ ve $x=M$ olmak üzere x ve y arasında pozitif doğrusal korelasyon olması koşulu aranır.

Çizelge 3.10. Schnabel ve Schumacher-Eschmeyer tahmin yönteminde seçim

	Toplam	İşaretli Yakalanan	Yeni İşaretlenen	M
1. Yakalama	C_1	R_1	U_1	M_1
2. Yakalama	C_2	R_2	U_2	M_2
...				
...				
(t-1). Yakalama	C_{t-1}	$R_{(t-1)}$	$U_{(t-1)}$	$M_{(t-1)}$
t. Yakalama	C_t	R_t	U_t	M_t

s: serideki toplam örnek sayısı olmak üzere,

Her bir t örneği için;

C_t : t örneğinde yakalanmış bireylerin toplam sayısı,

R_t : t örneğinde yakalanmış işaretli bireylerin sayısı,

U_t : t örneğinde ilk kez işaretlenmiş ve bırakılmış bireylerin sayısı,

M_t : t örneği alınmadan önce popülasyondaki işaretli birey sayısını göstermektedir.

Birey kaybı varsa U_t değerinden çıkarılmalıdır. $C_t=R_t+U_t$ eşitliği sağlanır.

Denklem 3.14. t örneği alınmadan önce popülasyondaki işaretli birey sayısı

$$M_t = \sum_{i=1}^{t-1} U_i$$

Denklem 3.14 için $t=6$ durumunda $M_6=U_1+U_2+U_3+U_4+U_5$ ile elde edilmektedir.

Böylece Schnabel'in (1938) Petersen tahmininden yola çıkarak önerdiği N'nin ve varyansın tahmini denklem 3.15 ve denklem 3.16'daki gibidir.

Denklem 3.15. Schnabel yöntemiyle popülasyon tahmini

$$\hat{N} = \frac{\sum_{t=1}^s C_t M_t}{\sum_{t=1}^s R_t}$$

Denklem 3.16. Schnabel yöntemiyle popülasyona ait varyans tahmini

$$\text{Var}\left(\frac{1}{\hat{N}}\right) = \frac{\sum_{t=1}^s R_t}{\left(\sum_{t=1}^s C_t M_t\right)^2}$$

Eğer yakalanan her bir örnekteki işaretli popülasyon sayısı %10'dan az ise popülasyon genişliği tahmini denklem 3.17 ile elde edilir.

Denklem 3.17. Çakışma azlığı durumunda düzeltilmeli popülasyon tahmini

$$\hat{N} = \frac{\sum_{t=1}^s C_t M_t}{\left(\sum_{t=1}^s R_t\right) + 1}$$

3.2.2.4. Schumacher - Eschmeyer Yöntemi

Schumacher ve Eschmeyer'in (1943) öne sürdüğü N popülasyon tahmini ise denklem 3.18 ile hesaplanmıştır (32). Notasyonlar Schanabel yönteminde olduğu gibidir. Ancak bu defa $y=R/C$ ve $x=M$ olmak üzere x ve y arasında pozitif doğrusal korelasyon olma koşulu aranmaz.

Denklem 3.18. Schumacher–Eschmeyer yöntemiyle popülasyon tahmini

$$\hat{N} = \frac{\sum_{t=1}^s C_t M_t^2}{\sum_{t=1}^s R_t M_t}$$

Örneklemlere ait varyans ise denklem 3.19 ile hesaplanır. (Zar, 1996).

Denklem 3.19. Schumacher–Eschmeyer yöntemiyle popülasyona ait varyans tahmini

$$Var(\hat{N}) = N^2 \left[\frac{NS^2}{\sum_{t=1}^s R_t M_t} \right]$$

Burada S^2 denklem 3.20 ile bulunmaktadır.

Denklem 3.20.

$$S^2 = \frac{\sum_{t=1}^s \left(\frac{R_t^2}{C_t} \right) - \frac{\sum_{t=1}^s (R_t M_t)^2}{\sum_{t=1}^s (C_t M_t^2)}}{s - 1}$$

$$StandartHata(\hat{N}) = \sqrt{Var(\hat{N})}$$

olduğu bilinir ve tahminlenen N 'nin güven aralığı;

$$N \pm t \cdot \sqrt{Var(\hat{N})}$$

şeklinde elde edilir.

Yapılan tahminin iyi bir tahmin olması için her bir örnekteki R/C ve M/C oranlarının %10'dan az olması beklenmektedir (30,31, 32, 33).

3.2.3. Yakala Tekrar Yakala Tahmininde Yanlılık Tipleri

Yakala tekrar yakala popülasyon tahmin yöntemlerinin en önemli ortak varsayımı bireylerin yakalanma şanslarının eşit olmasıdır. Bunun sağlanabilmesi için popülasyonla ilgili bazı ön bilgilere ihtiyaç duyulmaktadır.

Tuzaktan Çekinme Eğilimi (Trap Shy): Ortamda yakalama amacıyla bir çeşit tuzak kullanılıyor ve bu tuzak ilgilenilen popülasyonun davranışlarında değişikliğe sebep oluyorsa sonuçlar homojen olmayabilir. Tuzak çekingenliğinde çakışan denek sayısı düşük olacağından (under-matching) popülasyon olduğundan daha fazla tahmin edilir (over-estimate) (34). Enjeksiyon korkusu olan bir hastanın tedavi olmak için hastaneye gidememesi tuzak çekingenliğine bir örnektir. Toplum sağlığı merkezi ve özel hastane kayıtlarının birleştirilmesi sonucunda da çakışan denek sayısı düşük olmaktadır.

Tuzağa Tekrar Girme Eğilimi: (Trap Happy): Tuzağa yakalanmaya alışan deneklerin sebep olduğu bir durumdur. Gruplar arasında pozitif bir bağımlılık varsa yani bir grupta bulunmak diğer grupta bulunma olasılığını arttırıyorsa (Toplum sağlığı merkezinden hastaneye sevk gibi) çakışan denek sayısında artış (over-matching) söz konusudur. Bu da tahmin edilmek istenen popülasyonun olduğundan daha az tahmin edilmesine (under-estimate) neden olmaktadır.

Tuzak Yerleşimi: Tuzağın yerleşim yeri örnekleme girecek deneklerin aynı olmasına sebep olmayacak şekilde belirlenmelidir. Muayeneye gidecek bir hastanın özel bir rahatsızlığı varsa polikliniğin girişi, bekleme salonu ve hastaların çağırılma şekli, hastanın oraya yeniden gitmesi ya da gitmemesi üstüne etkilidir (19).

İlk Kaynak-Çakışma İlişkisi: İki ya da daha fazla kaynak kullanarak prevalans veya insidans çalışması yapan bir araştırmacının kaynaklardaki tüm mümkün çakışmaların sayısını elde edebilmesi için en uygun yöntem yakala tekrar yakala yöntemidir. İlk kaynaktaki veriler, sıradaki tüm kaynakları etkilemekte iken, çakışmanın azlığı yakala tekrar yakala sonuçlarını etkilemektedir (35).

Hayvanlarla ilgili çalışmalarda her bir yakalamadaki birimler aynı şekilde işaretlenir, farklı yakalamalar farklı işaret veya renklerle belirlenir. Aksi halde sonuçlarda yanlılığa sebep olur. Hayvanlar dışında bir popülasyonla çalışıldığında

yakalama yapmak yerine kayıtlar ve veri tabanlarındaki ilgili popülasyonların çakışmasıyla ilgilenilmektedir.

Prevalans ve insidans tahmini için iki ve daha çok kaynaktan veri kullanan ya da toplayan araştırmacı sadece kaynak çakışmalarını takip etmekle kalmayıp aynı zamanda bunları bir rapor şeklinde yayınlamalıdır. Böylece bulgulardaki tüm farklı vakalar bildirilerek elde edilen toplam tahminin inandırıcılığı artar (35).

İlk defa yakalanma şansıyla, daha sonraki yakalanma şanslarının birbirlerini etkilemediği varsayılır. Ancak hastalıklarla ilgili durumlarda vakalar tamamen bağımsız değildir. Kronikleşmiş veya çok ciddi hastalıklarda kliniğe bir kez gelen bir hastaya konulan tanı hastanın tekrar tekrar gelme olasılığını etkilemektedir. Hem yüksek mortaliteye sahip hem de daha uzman merkezlere sevk eğilimli hastalıklarda hastane uzmanlarının tuttuğu raporlar ve ölüm kayıtları birbirinden çok da bağımsız olmayacaktır. Bu tür problemlerle başa çıkan istatistiksel yöntemler vardır. Bunlardan biri de logaritmik doğrusal modellerdir.

Liste seçimi: Çalışmanın yapılacağı listelerin seçimi önemlidir, yanlış seçim korelasyon yanlılığına (correlation bias) sebep olur.

X ve Y listelerinin örnek genişliği sırasıyla n_1 ve n_2 olsun ($n_2 > n_1$).

Eğer bir A bireyi hem $X(n_1)$ listesinde olup hem de $Y(n_2)$ listesinde yer alıyorsa pozitif korelasyon söz konusudur ve beraberinde aşırı uyum (overmatching) problemi oluşur. Bu ise sonucun eksik tahminlenmesine (underestimation bias) sebep olur.

Çizelge 3.11. Korelasyon yanlılığı gösterimi

		X	
		+	-
Y	+	n_{12}	x
	-	y	a

İki kaynakta ortak birey çok fazla ise $\rightarrow a = \frac{xy}{n_{12}} \rightarrow n_{12}$ çok büyük olduğu için a

oldukça küçük çıkar (36).

Tam tersi olarak A bireyi X listesinde yer almadığı halde Y listesinde varsa, negatif korelasyon söz konusudur ve beraberinde eksik uyum (undermatching) problemi oluşur. Bu durumda sonuçlar olduğundan aşırı tahminlenir (overestimation bias). Bu durumda odss oranı 1'den büyük çıkar. Çünkü N büyüdükçe x de büyür.

İki kaynaktan ortak birey sayısı çok az ise $\rightarrow a = \frac{xy}{n_{12}} \rightarrow n_{12}$ çok küçük olduğu için a oldukça büyük çıkar (36).

Homojenlik: Karşılaştırma yapılacak listelerde istenilen ortak özellikler arttıkça çalışmanın homojenliği artar. Bu sayede listelerin birleştirilmesindeki doğruluk artar (3). Tanımlayıcı bilgilerin doğruluğu ve verinin diagnostik kalitesinin karşılaştırılmasını bu yöntem sağlar. Düşük seçicilik, yüksek duyarlılık durumunda liste kayıtlarına bildirim artar (37).

3.2.4. Tarama Testleri İçin Yakala Tekrar Yakala Yaklaşımı

Hastalığın erken teşhisindeki tıbbi tarama prosedürlerinin klinik etkinliğini değerlendirmek için sık sık geniş popülasyonlar üstüne uygulanması gerekir. Hastalıklı birini gözden kaçırabilme (yanlış negatif) olasılığı ya da sağlıklı birine yanlış tanı koyma (yanlış pozitif) olasılığını tahmin etmek amacıyla çoklu bir tarama programından elde edilen veriye Yakala tekrar yakala metodu uygulanabilir.

Bu yanlış sınıflama olasılıkları, taramaların etkili bir şekilde çalıştırılıp çalıştırılmadığını belirleme ihtiyacı doğurur. Gereksiz tedavi ve geç kalınmalar önlenmeye çalışılır. Çünkü hastalıklı birine sağlıklı diyerek tedavisine başlamamak ciddi sorunlar yaratırken, sağlıklı birine hastalıklı diyerek gereksiz yere riskli ve ileri tetkikler uygulanır ve tedavi verilebilir.

İki veya daha fazla tanı testi kullanılarak değerlendirilen hastaların gerçek durumunun değerlendirilmesi tüm hastalar için mümkün olmayabilir. Çoğu kez Altın Standart uygulaması zor veya pahalı bir yöntem veya cerrahi bir işlem olabilir. Bu durumda bazı hastalıklar için hastalar öncelikle en az iki teste tabi tutulur, bu testlerden biri bile pozitif sonuç vermiş ise ileri tanı yöntemi uygulanır. Ancak testler negatif sonuç vermiş ise uygulanmaz. Bu durumda test sonuçları negatif olan kişinin gerçek

durumu hiçbir zaman bilinmez. Bu durumda hastalığın gerçek prevalansını tahmin etmek için kullanılan farklı yöntemler mevcuttur. Bunlar;

1. Walter'ın kısmi gizli sınıf modeli
2. Lincoln-Petersen tahmini
3. Bağımlılık parametrelerine dayalı tahmin
4. Odds oranına dayalı tahmin

İçinde belirli bir D hastalığına sahip kişilerin bulunduğu “n” sayıda birey içeren bir popülasyon T_1 ve T_2 adında iki diagnostik test kullanılarak taranır. Bireyler “d” farklı kategoriye ayrılır. Eğer testlerden biri pozitif sonuçlanırsa, o zaman genellikle doğru hastalık statusünü değerlendirmek için diğer tıbbi tanı testleri uygulanır. Her iki testin de negatif olarak belirlediklerine başka bir tıbbi tanı testi uygulanmaz. Oysa iki testin de sağlıklı olarak belirledikleri içinde yanlışlıklar olmuş ve tedaviye başlanması gereken hastalar karışmış olabilir (38). Uygulanacak diğer tıbbi tanı testi (altın standart) zaman ve maddi sebeplerden dolayı herkese yapılamayacağından sadece pozitif test sonucu alanlar üstünde devam eder. Bu ise çalışmada doğrulama yanlılığına (verification bias) sebep olur.

Test 1 için $j=0,1$ ve Test 2 için $k=0,1$ sonuçlarını ve $i=1 \dots d$ hastalık durumlarını göstermek üzere taranan bireylerin frekansları $x_{jk}^{(i)}$ ile gösterilir. Tüm hastalık durumlarında $x_{00}^{(i)}$ bilinmez, $p_{jk}^{(i)}$ koşullu olasılıkları gösterir. Eğer tüm $x_{jk}^{(i)}$ frekansları gözlenebiliyorsa $n_i = x_{++}^{(i)}$ olmak üzere, tüm $p_{jk}^{(i)}$ koşullu olasılıkları $\hat{p}_{jk}^{(i)} = \frac{x_{jk}^{(i)}}{n_i}$ ve q_i

olasılıkları $\hat{q}_i = \frac{n_i}{n}$ ile tahmin edilebilir (38).

Eğer 2 hastalık durumu söz konusuysa ($d=2$, sağlıklı:1, hasta:2) $p_{1+}^{(2)}$ test 1'in sensitivitesi, $p_{+1}^{(2)}$ test 2'nin sensitivitesi, $p_{0+}^{(1)}$ test 1'in spesifitesini, $p_{+0}^{(1)}$ test 2'nin spesifitesini ve ($n=n_1+n_2$) olmak üzere $q_2=n_2/n$ popülasyondaki hastalık prevalansını göstermektedir (37).

$$p_{1+}^{(2)} = \frac{x_{1+}^{(2)}}{n_2} \quad p_{+1}^{(2)} = \frac{x_{+1}^{(2)}}{n_2}$$

Çizelge 3.12. Tarama testleri için yakala tekrar yakala yaklaşımı

Altın Standartın Pozitifleri

Altın Standartın Negatifleri

		Test 2		
		+	-	
Test 1	+	x_{11}^1	x_{10}^1	x_{1+}^1
	-	x_{01}^1	$x_{00}^1=?$	x_{0+}^1
		x_{+1}^1	x_{+0}^1	n_1

		Test 2		
		+	-	
Test 1	+	x_{11}^2	x_{10}^2	x_{1+}^2
	-	x_{01}^2	$x_{00}^2=?$	x_{0+}^2
		x_{+1}^2	x_{+0}^2	n_2

3.2.4.1. Walter'ın Kısmi Gizli Sınıf Modeli

Walter (1999) test sonucu negatif olanların tahmini için gizli sınıf yaklaşımlı bir kısmi gizli sınıf (latent class) modeli geliştirmiştir (38). Tarama testlerinin hastalık üstündeki bağımsızlık koşulunu kullanan Walter modeli, $p_{11}^{(i)} = p_{1+}^{(i)} p_{+1}^{(i)}$ olduğunu varsayar. 2x2 bir tabloda bu eşitlik $p_{jk}^{(i)} = p_{j+}^{(i)} p_{+k}^{(i)}$ iki ayrı marjinal olasılığın birleşimi halini alır. Toplamda 3d-1 kadar parametre elde edilir, bunun d-1'i oran, 2d'si ise koşullu olasılıklardır. Örneğin hastalığın iki alt kategorisi varsa (sağlıklı, grade1, grade2), iki sensitivite, iki spesifite ve bir prevalans parametresi elde edilir. Olabilirlik denklem 3.21 ile sağlanır.

Denklem 3.21. Kısmi gizli sınıf modelinde olabilirlik hesabı

$$L = \prod_{i=1}^d (p_{1+}^{(i)} p_{+1}^{(i)} q_i)^{x_{11}^{(i)}} \prod_{i=1}^d (p_{0+}^{(i)} p_{+1}^{(i)} q_i)^{x_{01}^{(i)}} \prod_{i=1}^d (p_{1+}^{(i)} p_{+0}^{(i)} q_i)^{x_{10}^{(i)}} \left(\sum_{i=1}^d p_{0+}^{(i)} p_{+0}^{(i)} q_i \right)^{x_{00}^{(+)}}$$

Burada her bir $x_{00}^{(i)}$ frekansları bilinmese de $x_{00}^{(+)} = x_{00}^{(1)} + x_{00}^{(2)} + \dots + x_{00}^{(d)}$ eşit olduğu bilinmektedir. Olabilirlik fonksiyonunu maksimum yapan $p_{1+}^{(i)}$, $p_{+1}^{(i)}$ ve q_i tahminleri elde edilmeye çalışılır. Olabilirliği maksimize etmek için Newton-Raphson

prosedürü kullanılır, ancak alternatif olarak EM algoritması da kullanılır. Böylece $\hat{n}_i^* = n \times \hat{q}_i$ ($i=1, \dots, d$) eşitliği popülasyon genişliğinin tahminini verir (38).

3.2.4.2. Lincoln-Petersen Tahmini

Bu tahmin Lincoln ve Petersen tarafından farklı veri setlerinin bağımsızlığından yola çıkılarak elde edilmiştir (38). Chapman tarafından önerilen hesaplama kullanılır.

Denklem 3.22. Chapman yöntemiyle i . hastalık durumu için popülasyon tahmini

$$\hat{n}_i = \frac{(x_{1+}^{(i)} + 1)(x_{+1}^{(i)} + 1)}{x_{11}^{(i)} + 1} - 1$$

Tüm hastalık durumları için bu eşitlik hesaplanır ve toplanarak popülasyon genişliği elde edilir.

Denklem 3.23. Toplam popülasyon tahmini

$$n = \sum_{i=1}^d \hat{n}_i$$

3.2.4.3. Bağımlılık Parametrelerine Dayalı Tahmin

Bu model koşullu bağımlılık modelinden daha esnek bir modeldir. Ancak bağımlılıklara olan izin sınırlıdır (37). Modeldeki θ 'ya ilişkin tahmin yapılarak, θn_i elde edilir.

Taranan popülasyon genişliği denklem 3.24 ile hesaplanır ve her bir hastalık durumu için tahmin edilen genişlikler ortalamasına eşitlenir.

Denklem 3.24. Popülasyon genişliği hesabı

$$\sum_{i=1}^d E(\theta \hat{n}_i) = \theta \sum_{i=1}^d E(\hat{n}_i) = \sum_{i=1}^d n_i = n$$

Teorik olarak beklenen deęer $E(\hat{n}_i)$ yerine gözlenen deęer \hat{n}_i yazılarak

$$\hat{\theta} \sum_{i=1}^d \hat{n}_i = n \text{ ya da } \hat{\theta} = \frac{n}{\sum_{i=1}^d \hat{n}_i} \text{ olarak elde edilir.}$$

Denklem 3.26'dan yola çıkarak tüm $(i=1, \dots, d)$ hastalık durumlarına göre genişlik tahmini yapılandırılabilir ve toplanarak popülasyon genişliği denklem 3.27'deki gibi elde edilir (38).

Denklem 3.25. i. hastalık durumu için popülasyon tahmini

$$\hat{v}_i = \hat{n}_i \hat{\theta} = \frac{\hat{n}_i}{\sum_{i=1}^d \hat{n}_i} n$$

Denklem 3.26. Toplam popülasyon tahmini

$$\sum_{i=1}^d \hat{v}_i = n$$

3.2.4.4. Odds Oranına Dayalı Tahmin

Hastalık durumuna baęlı iki testin ilişkisi odds oranı kullanılarak ölçülebilir. odds oranı parametresi α olmak üzere, denklem 3.28 ile hesaplanır.

Denklem 3.27. Odds oranı parametresi

$$\hat{\alpha} = \frac{x_{00}^+}{\sum_{i=1}^d x_{10}^{(i)} x_{01}^{(i)} / x_{11}^{(i)}}$$

Tarama sonucu negatif olanların toplam sayısı $x_{00}^{(+)}$ bilinir ve n_i tahmini, denklem 3.29 ile hesaplanır (38).

Denklem 3.28. Tahmini popülasyon hesabı

$$\hat{v}_i = x_{11}^{(i)} + x_{10}^{(i)} + x_{01}^{(i)} + \hat{\alpha} \frac{(x_{10}^{(i)} + 1)(x_{01}^{(i)} + 1)}{+1} - 1$$

Literatürde bu dört yöntem simülasyon çalışmasıyla karşılaştırılmıştır (38). Karşılaştırma için en çok olabilirlik tahminleri (MLEs) Walter modeline dahil edilmiştir. Walter ve yakala tekrar yakala modellerindeki yanlılığın birbirine yakın olduğu ve örneklem sayısı arttıkça daha da yakınlaştığı görülmüştür (38). Örneğin örnek genişliği 5000'den fazla olduğunda ortaya çıkan yanlılık farklılığı göz ardı edilebilir.

En çok olabilirlik tahmin yöntemleri yerine yakala tekrar yakala tahmin yöntemlerinin kullanılması verimlilikte büyük bir kayıp oluşturmamaktadır. Bu ise elde edilmesi zor olan en çok olabilirlik tahmin yöntemlerinin kullanılmak zorunda olmadığını gösterir.

3.2.5. Logaritmik Doğrusal (Log-Linear) Model

İki yönlü çapraz tabloların bağımsızlık hipotezi testi 1960'ların sonlarına kadar Pearson ki-kare analiziyle yapıyordu.

Daha sonra tablolardaki değişkenler ikiden fazla kategori içermeye başladığı zaman değişkenler arasında interaksiyon ya da ilişki olup olmadığını belirleyebilmek için iki yönlü ki-kare analiz hesaplamalarını tekrarlayarak çok yönlü ki-kare analizi yapma ihtiyacı doğdu. Ancak L.A. Goodman (1970)'in logaritmik doğrusal (log-linear) modeller ile ilgili yayınladığı makaleler serisiyle kategorik veri analizinde oldukça değişim gerçekleşmiştir. Logaritmik doğrusal model aynı zamanda, yakala tekrar yakala verisi gibi çok kaynak içeren, tamamlanmamış bağımlı listelerden popülasyon tahmini yapan güçlü bir analiz metodudur (39).

Logaritmik doğrusal modellemenin temel stratejisi, kategorik değişkenlerin sınıflanmasında gözlenen frekansların modellenmesini içerir, gözlenen frekanslarla beklenen frekansların birbirine ne kadar benzeyip benzememesiyle ifade edilebilir (39).

Logaritmik doğrusal modelin basit teorisinde f_{ij} iki yönlü bir tablodaki gözlenen değerleri ve F_{ij} , beklenen değerleri göstermek üzere, satır ve sütundaki değişkenlerin ilişkisi Çizelge 3.13'deki gibi gösterilir (40).

Çizelge 3.13. İki yönlü tablo için gözlenen frekanslar

		B	
		1	0
A	1	f_{11}	f_{10}
	0	f_{01}	f_{00}

Beklenen değerler arasındaki ilişki Yule (1900)'un çapraz çarpım oranı (odds) önerisine göre denklem 3.20'deki gibi yazılabilir.

Denklem 3.29. Yule'un çapraz çarpım oranı

$$\alpha = \frac{f_{11}f_{00}}{f_{10}f_{01}}$$

A ve B kaynaklarının birbirinden bağımsız olması durumunda $\alpha=1$ 'dir. 2 x 2 çapraz tablosundaki beklenen değerlere ilişkin model logaritmik doğrusal formda şu şekilde ifade edilebilir.

Denklem 3.30. İki kaynaklı bağımsız logaritmik doğrusal model

$$Ln(F_{ij.}) = \mu + \lambda_i^A + \lambda_j^B$$

Burada, μ beklenen frekansların doğal logaritmasının genel ortalamasını λ_i^A ve λ_j^B ise A ve B kaynaklarının ana etkilerini gösterir.

İki kaynağın bağımsız olmadığı durumda ise modele λ_{ij}^{AB} etkileşimi eklenir.

Denklem 3.31. İki kaynaklı bağımlı logaritmik doğrusal model

$$Ln(F_{ij.}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

Kaynak (liste) sayısı k olmak üzere her bir kaynağın kullanıldığı 2^k boyutunda bir çapraz tablo oluşturulur. Bu tablodaki bir göze, kullanılan kaynaklarda yer almayan birimleri temsil ederken, diğerleri nüfusu kesinleşmiş birim sayılarını gösterir. İki kaynak söz konusu olduğunda kullandığımız basit matematiksel denklemlerle popülasyon tahmininde bulunmak imkansızdır. Çünkü kaynaklar arası bağımlılık ve heterojenliği kontrol etmek ancak logaritmik doğrusal modeller kullanılarak mümkün hale gelir (41).

İkiden çok kaynak durumunda harflendirme ve tablo gösterimi Çizelge 3.14'deki gibi gösterilmektedir (3).

Çizelge 3.14. $k=3$ için yakala tekrar yakala dizaynı

Kaynak1	Kaynak2	Kaynak3	Frekans
1	1	1	f_{111}
1	1	0	f_{110}
1	0	1	f_{101}
1	0	0	f_{100}
0	1	1	f_{011}
0	1	0	f_{010}
0	0	1	f_{001}
0	0	0	$f_{000}=?$

Birim sayısı bilinmeyen (f_{000}) boş gözenin tahmini için logaritmik doğrusal model kullanılır.

Çizelge 3.15. Üç yönlü tablo için gözlenen frekanslar

	C(1)		C(0)	
	B(1)	B(0)	B(1)	B(0)
A(1)	f_{111}	f_{101}	f_{110}	f_{100}
A(0)	f_{011}	f_{001}	f_{010}	f_{000}

Kaynakların birbirinden bağımsız olduğu varsayımıyla tablolar birbirine eşitlenebilir.

Denklem 3.32. Bağımsızlık durumunda çapraz çarpım gösterimi

$$\alpha = \frac{f_{111}f_{001}}{f_{101}f_{011}} = \frac{f_{110}f_{000}}{f_{100}f_{010}}$$

Beklenen değerlere ilişkin 3 kaynaklı etkileşim modelinde en genel model olarak, beklenen frekansların doğal logaritmasının genel ortalaması, her bir değişkenin ana etkileri, değişkenlerin ikişer interaksyonlu etkileri ve üç değişkenin etkileşimi yer alır.

Denklem 3.33. Üç kaynaklı bağımlı logaritmik doğrusal model

$$\ln(F_{ijk}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

Tüm interaksyonları içeren bu modele doymuş (saturated) model denir. Herhangi bir değişkenin etkisini izole ederek elde ettiğimiz modele ise “Bağımsız Model” denilmektedir (42, 43).

3.2.5.1. Model Seçim Kriterleri

Kullanılacak modelin belirlenmesinde dikkat edilmesi gereken iki nokta vardır. İlki seçilen modelin gereksiz parametreleri içermemesi, ikincisi ise modele alınması gereken parametrenin modelde yer almaması durumudur.

İlk olarak kullanılacak modelde tüm kaynaklar arasında ilişki varmış gibi doymuş model ele alınıp en yüksek dereceli etkileşimden başlanarak ilişkili olmayan terimler sırasıyla çıkarılabilir. Buna geriye dönük (backward) değişken eleme yöntemi denir. Ya da en bağımsız (doymamış) modelden başlanarak ilişkili terimlere ait etkileşimler sırasıyla modele eklenebilir. Buna da ileriye dönük (forward) değişken seçim yöntemi denir.

Kaynaklar arasında bir ilişki olup olmadığı (etkileşim) gözlenen frekanslara ait uygunluk (fitting) modelleriyle belirlenir. Uygunluk modeli testi için her bir gözlenen

frekansa ait beklenen frekans bulunur. Pearson ki-kare ve olabilirlik oran istatistiği (Likelihood ratio) (L^2) ile bu frekanslar karşılaştırılır. Likelihood ratio (L^2), çok yönlü tabloların bağımsızlık testinde daha güçlü bir test olduğundan daha yaygın bir kullanımı vardır.

L^2 istatistiği $(I-1)(J-1)$ serbestlik dereceli ki kare dağılımı gösterir ve iki kaynak durumunda formülü;

Denklem 3.34. İki kaynak durumunda L^2 hesabı

$$L^2 = 2 \sum f_{ij} \ln \left(\frac{f_{ij}}{F_{ij}} \right)$$

şeklinindedir. Üç kaynak olması durumunda ise denklem 3.36 ile hesaplanır.

Denklem 3.35. Üç kaynak durumunda L^2 hesabı

$$L^2 = 2 \sum f_{ijk} \ln \left(\frac{f_{ijk}}{F_{ijk}} \right)$$

Elde edilen L^2 değerlerinden daha büyük olan, modelin ilgili veriler için uygun olmadığı ve bu yüzden reddedilmesi gerektiğini gösterir. Karşılaştırma yapılacak iki model iç içe (nested) model olmalı ve daha kompleks olan modelden daha basit olan model çıkarılmalıdır. Model karşılaştırmalarında;

Denklem 3.36. Likelihood ratio ile model karşılaştırması

$$L_k^2 = L_{\text{model1}}^2 - L_{\text{model2}}^2$$

eşitliği kullanılır ve serbestlik derecesi, 1 numaralı modelden 2 numaralı modelin serbestlik derecesinin çıkarılmasıyla bulunur.

Kaynak 1 (A) için kategori sayısı I, Kaynak 2 (B) için J ve Kaynak 3 (C) için K olmak üzere üç boyutlu çapraz tablo için bağımsız logaritmik doğrusal modeller ve serbestlik dereceleri Çizelge 3.16'da verilmiştir.

Çizelge 3.16. Üç boyutlu çapraz tablo için bağımsız logaritmik doğrusal modeller ve serbestlik dereceleri

Model	Logaritmik doğrusal model ($\ln F_{ijk}$)	Serbestlik derecesi
$M^{(0)}$, [A] [B] [C]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C$	IJK-I-J-K+2
$M^{(1)}$, [A] [BC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC}$	IJK-I-JK+1
$M^{(2)}$, [B] [AC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC}$	IJK-IK-J+1
$M^{(3)}$, [C] [AB]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$	IJK-IJ-K+1
$M^{(4)}$, [AC] [BC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	K(I-1)(J-1)
$M^{(5)}$, [AB] [BC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC}$	J(I-1)(K-1)
$M^{(6)}$, [AB] [AC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$	I(J-1)(K-1)
$M^{(7)}$, [AB] [AC] [BC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$	(I-1)(J-1)(K-1)
$M^{(8)}$, [ABC]	$\mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$	0

Modelin uygunluğunu test etmek için kullanılan istatistiklerden diğerleri de Akeike ve Bayes bilgi kriteridir.

Akeike Bilgi Kriteri (Akeike İyformasyon Criteria): Akeike tarafından (1973) geliştirilen bu bilgi kriteri, modele eklenen değişkenlerin yarattığı farklılığın tespitinde kullanılan kriterlerden birisidir. Denklem 3.38 ile elde edilir (44,45). Her bir model için ayrı ayrı ve bağımsız olarak hesaplanır. Karşılaştırma için iç içe (nested) model zorunluluğu yoktur.

Denklem 3.37. Akeike bilgi kriteri değeri

$$AIC = -2[\text{Log}(L) - k]$$

Burada k, sabit terim sayısını gösterir. Tüm modeller için AIC hesaplanır ve en küçük AIC'ye sahip model ilgili veri için en uygun model olarak belirlenir.

Bayes Bilgi Kriteri (Bayesian İyformasyon Criteria): Akeike bilgi kriterine bayes perspektifinden bakan Schwarz (1978) tarafından geliştirilen bu kriter de tüm

modeller için hesaplanır ve en küçük BIC'ye sahip model ilgili veri için en uygun model olarak belirlenir (44,45).

Denklem 3.38. Bayes bilgi kriteri değeri

$$BIC = -2\text{Log}(L) + k\text{Log}(n)$$

Yine en küçük BIC hesaplanan model en uygundur sonucuna varılır.

3.2.5.2. Logaritmik Doğrusal Model Kısıtlılıkları

Değişken Sayısı: Logaritmik doğrusal modele çok fazla değişkenin dahil edilmesi yorumlamayı zorlaştırmaktadır (41).

Bağımsızlık: Bir hücredeki sıklık değerlerinden bağımsızdır.

Yeterli örnek genişliği: Logaritmik doğrusal modellerde en azından verideki hücre sayısının 5 katı kadar vakaya ihtiyaç duyulmaktadır. Örneğin 2 x 2 x 3 boyutlu bir çapraz tablo için en az 60 vakaya ihtiyaç vardır. Yeterli vaka yoksa arttırılması ya da değişkenlerden bir veya daha fazlasının elenmesi önerilir. Ancak gerekli değişkenlerin elenmesi bilgi kaybına sebep olacağından en etkili yol veri sayısının arttırılmasıdır.

Beklenen sıklık büyüklüğü: Tüm iki yönlü interaksyonlar için beklenen hücre sıklıkları 1'den büyük olmalıdır ve en fazla %20'si 5'den az olabilir. Bu gereksinim karşılanamazsa tip 1 hata oranı genelde yükselmez fakat gücün analiz sonuçlarını etkileyecek kadar düşebileceği bilinir (42).

Eğer düşük beklenen sıklıklar gözlemlenirse aşağıdaki adımlar izlenmelidir.

- Düşük beklenen değerlere ilişkin etkilerin test edilmesi için azalan güç kabul edilebilir.
- Göze birleştirme yapılabilir ancak bu durumda interaksyonların test edilmesine ilişkin güçte ciddi bir düşüş yaşanır ve sonuçta herhangi bir bulgu elde edilemez.
- Gözelerin sayısını azaltmak için değişkenler silinebilir ancak diğer değişkenlerle ilişkisi olan bir değişkeni silmemek gereklidir.

- Her bir hücreye sabit bir deęer eklenebilir fakat güç düşer ve tip 1 hata oranı ancak minimal düzeyde yükselir (39, 42).

4. BULGULAR

4.1. Tahmin Yöntemlerinin Uygulanışı

Yakala tekrar yakala yöntemi kullanılarak eksik veri ve popülasyon genişliğinin tahmin aşamalarını Lincoln-Petersen, Jolly-Seber, Schnabel, Schumacher-Eschmeyer yöntemleri ve logaritmik doğrusal modelle göstermek amacıyla her biri için ayrı ayrı veri setleri üretilmiştir.

4.1.1. Lincoln-Petersen Tahmin Yöntemine Ait Uygulama

Lincoln-Petersen tahmin yönteminin kullanılmasını göstermek amacıyla, $p=0,50$, $n=100$ ve yakalama sayısı $k=2$ olacak şekilde Ek'deki makro ile veri simülasyonu yapılmıştır. Elde edilen veri seti Çizelge 4.1'de verilmiştir.

Çizelge 4.1: Lincoln-Petersen için yakalama sonuçları

		Liste 2		Toplam
		+	-	
Liste 1	+	28	38	66
	-	29	a	$29+a$
Toplam		57	$38+a$	N

Burada Lincoln-Petersen tahmini ile bilinmeyen göze,

$$a = \frac{xy}{R} = \frac{38 \times 29}{28} = 40$$

olarak hesaplanmıştır. Toplam popülasyon genişliği ve varyans hesabı ise,

$$N_{LP} = \frac{MC}{R} = \frac{66 \times 57}{28} \cong 135$$

$$Var(N_{LP}) = \frac{MCxy}{R^3} = \frac{66 \times 57 \times 38 \times 29}{28^3} = 188,8 \quad \text{olarak bulunur.}$$

Chapman tahminiyle elde edilen popülasyon genişliği ve buna ait varyans ise,

$$N_{Chap} = \frac{(M+1)(C+1)}{R+1} - 1 = \frac{(66+1)(57+1)}{28+1} - 1 = 133$$

$$Var(N_{Chap}) = \frac{(M+1)(C+1)xy}{(R+1)^2(R+2)} = \frac{(66+1)(57+1)38 \times 29}{(28+1)^2(28+2)} = 169,7$$

olarak bulunur. Yani Lincoln-Petersen yöntemiyle 135 olarak tahmin edilen popülasyon genişliği, Chapman düzeltmesiyle 133 olarak bulunmuştur.

4.1.2. Jolly-Seber Tahmin Yöntemine Ait Uygulama

Jolly-Seber tahmin yönteminin kullanılmasını göstermek amacıyla, Çizelge 4.2'deki hipotetik veri seti kullanılmıştır. Burada;

t . örneklem sırasında çekilen n_t adet bireyden m_t tanesi daha önce yakalanmış ve işaretlenmiş, u_t tanesi de ilk defa t . örneklemde yakalanmıştır.

k : çalışmadaki örneklem sayısı ($t=1, \dots, k$)

m_t : t örnekleminde yakalanan işaretli birey sayısı

u_t : t örnekleminde yakalanan işaretli olmayan birey sayısı

n_t : t örnekleminde yakalanan toplam birey sayısı ($m_t + u_t$)

S_t : t örnekleminde sonra geri bırakılan birey sayısı (n_t - yaralanma, ölüm ya da işaretlenme silinmesi)

r_t : t örnekleminde yakalanıp R_t olarak geri bırakılan bireylerden sonraki örneklemlerde tekrar yakalananların sayısı ($t=1, \dots, (k-1)$)

z_t : t örnekleminde önce işaretlenmiş, t örnekleminde yakalanmamış ancak sonraki örneklemlerde yakalanmış bireylerin sayısını gösterir.

Çizelge 4.2. Jolly-Seber için yakalama sonuçları

t	m_t	u_t	n_t	S_t	r_t	z_t
1	-	50	50	50	30	-
2	14	41	55	51	25	16
3	14	24	38	38	17	17
4	25	20	45	42	6	6
5	25	14	39	39	-	-

t=1 yakalaması için m_1 ve z_1 olmadığından hesaplama t=2'den başlar.

t=2 yakalaması için,

$$\hat{\alpha}_2 = \frac{m_2 + 1}{n_2 + 1} = \frac{14 + 1}{55 + 1} = 0,268$$

$$\hat{M}_2 = \frac{(S_2 + 1)z_2}{r_2 + 1} + m_2 = \frac{(51 + 1)16}{25 + 1} + 14 = 46$$

$$\hat{N}_2 = \frac{\hat{M}_2}{\hat{\alpha}_2} = \frac{46}{0,268} \cong 172$$

t=3 yakalaması için,

$$\hat{\alpha}_3 = \frac{m_3 + 1}{n_3 + 1} = \frac{14 + 1}{38 + 1} = 0,385$$

$$\hat{M}_3 = \frac{(S_3 + 1)z_3}{r_3 + 1} + m_3 = \frac{(38 + 1)17}{17 + 1} + 14 = 50,8$$

$$\hat{N}_3 = \frac{\hat{M}_3}{\hat{\alpha}_3} = \frac{50,8}{0,385} \cong 132$$

t=4 yakalaması için,

$$\hat{\alpha}_4 = \frac{m_4 + 1}{n_4 + 1} = \frac{25 + 1}{45 + 1} = 0,565$$

$$\hat{M}_4 = \frac{(S_4 + 1)z_4}{r_4 + 1} + m_4 = \frac{(42 + 1)6}{6 + 1} + 25 = 61,9$$

$$\hat{N}_4 = \frac{\hat{M}_4}{\hat{a}_4} = \frac{61,9}{0,565} \cong 110$$

t=5 yakalaması için ise r_5 ve z_5 olmadığından ilgili hesaplamalar yapılamaz.

Çizelge 4.2'deki örneklemin yapıldığı popülasyonun genişliği hesaplanırken ilk yakalamadaki birey sayısı, ara yakalamalardaki tahmini birey sayısı ve son yakalamadaki işaretli olmayan birey sayısı toplanır.

$$n_1 + \hat{N}_2 + \hat{N}_3 + \hat{N}_4 + u_5 = 50 + 172 + 132 + 110 + 14 = 478 \text{ olarak tahmin edilir.}$$

4.1.3. Schnabel Tahmin Yöntemine Ait Uygulama

Schnabel tahmin yönteminin kullanılmasını göstermek amacıyla, Çizelge 4.3'deki hipotetik veri seti kullanılmıştır.

Çizelge 4.3. Schnabel için yakalama sonuçları

	Toplam (C)	İşaretli Yakalanan (R)	Yeni İşaretlenen (U)	M
1. Yakalama	39	0	39	0
2. Yakalama	43	8	35	39
3. Yakalama	44	15	29	74
4. Yakalama	44	19	25	103
5. Yakalama	41	24	17	128
6. Yakalama	45	26	19	145

$$M_6 = 145$$

$$\sum_{t=1}^6 C_t M_t = 39 \times 0 + 43 \times 39 + 44 \times 74 + 44 \times 103 + 41 \times 128 + 45 \times 145$$

$$\hat{N} = \frac{\sum_{t=1}^6 C_t M_t}{\sum_{t=1}^6 R_t} = \frac{21238}{92} = 231$$

$$\text{Var}\left(\frac{1}{\hat{N}}\right) = \frac{\sum_{t=1}^6 R_t}{\left(\sum_{t=1}^6 C_t M_t\right)^2} = \frac{92}{21238^2} = 2,03967 \times 10^{-7}$$

Yani Çizelge 4.3'deki örneklemin yapıldığı popülasyonun genişliği 231 olarak tahmin edilir.

4.1.4. Schumacher - Eschmeyer Yöntemine Ait Uygulama

Schumacher-Eschmeyer tahmin yönteminin kullanılmasını göstermek amacıyla, Çizelge 4.4'deki hipotetik veri seti kullanılmıştır.

Çizelge 4.4. Schumacher-Eschmeyer için yakalama sonuçları

	Toplam (C)	İşaretili Yakalanan (R)	Yeni İşaretlenen (U)	M
1. Yakalama	39	0	39	0
2. Yakalama	43	10	33	39
3. Yakalama	44	8	36	72
4. Yakalama	44	11	33	105
5. Yakalama	41	9	32	137
6. Yakalama	45	5	40	177

$$\hat{N} = \frac{\sum_{t=1}^6 C_t M_t^2}{\sum_{t=1}^6 R_t M_t} = \frac{2957933}{4239} = 698$$

$$S^2 = \frac{\sum_{t=1}^6 \left(\frac{R_t^2}{C_t}\right) - \frac{\sum_{t=1}^6 (R_t M_t)^2}{\sum_{t=1}^6 (C_t M_t^2)}}{6-1} = \frac{9,06 - \frac{4239^2}{2957933}}{5} = 0,597$$

$$Var(\hat{N}) = N^2 \left[\frac{NS^2}{\sum_{t=1}^6 R_t M_t} \right] = 698^2 \left[\frac{698 \times 0,597}{4239} \right] = 47872,87$$

Güven aralığı:

$$\hat{N} \pm t_{5, \alpha/2} \sqrt{Var(\hat{N})} \Rightarrow [(698 - 2,571 \times 218,8) - (234 + 2,571 \times 218,8)] \\ \Rightarrow [135,26 - 1260,32]$$

Yani Çizelge 4.4'deki örneklemin yapıldığı popülasyonun genişliği, Schumacher-Eschmeyer yöntemiyle 698 olarak tahmin edilir.

4.2. Yakala Tekrar Yakala Yaklaşımı ile Tarama Testlerinin Doğruluk Tahmini

Strax, Venet, Shapiro ve Gross'un (1967) çalışmasına ait veri seti meme kanseri şüphesiyle hastaneye başvuran 20211 kadına ait test 1 (Fiziksel muayene) ve test 2 (mamografi) tarama sonuçlarını içermektedir. Burada, en az bir testten pozitif sonuç alanlara biyopsi yapılmıştır.

Çizelge 4.5. Meme kanserinden şüphelenilen 20211 hastaya ait fizik muayene ve mamografi sonuçları

		Mamografi	
		+	-
Fiziksel Muayene	+	23	168
	-	116	19904

Çizelge 4.6. Meme kanserinden şüphelenilen ve biyopsi yapılan hastaların fizik muayene ve mamografi sonuçları.

Biyopsi ile sağlıklı bulunanlar				Biyopsi ile hasta bulunanlar			
		↓				↓	
		Mamografi				Mamografi	
		+		-		+	
Fiziksel	+	13	144	Fiziksel	+	10	24
Muayene	-	95	n₀₀¹=?	Muayene	-	21	n₀₀²=?

Çizelge 4.5 ve Çizelge 4.6'deki n_{00}^1 ve n_{00}^2 değerlerini elde etmek için Lincoln-Petersen yöntemiyle popülasyon genişliği aşağıdaki gibi tahmin edilir.

$$\hat{n}_i = \frac{(x_{1+}^{(i)} + 1)(x_{+1}^{(i)} + 1)}{x_{11}^{(i)} + 1} - 1$$

$$\hat{n}_1 = \frac{(157 + 1)(108 + 1)}{13 + 1} - 1 = 1229 \quad \rightarrow \quad \hat{n}_{00}^{(1)} = 977$$

$$\hat{n}_2 = \frac{(34 + 1)(31 + 1)}{10 + 1} - 1 = 101 \quad \rightarrow \quad \hat{n}_{00}^{(2)} = 46$$

Tahmin edilen toplam $n_{00}^1 + n_{00}^2 = 977 + 46 = 1023$ olarak bulunmuştur, oysa gerçek toplamının 19904 olduğu bilinmektedir. Bu durumda tahmini toplam içindeki ağırlıklandırmalardan faydalanılarak gerçek $n_{00}^{(i)}$ 'ler hesaplanır.

$$w_i = \frac{\hat{n}_{00}^{(i)}}{\hat{n}_{00}}$$

$$w_1 = \frac{977}{1023} = 0,95 \quad \rightarrow \quad n_{00}^{(1)} = 0,95 * 19904 = 18909$$

$$w_2 = \frac{46}{1023} = 0,05 \quad \rightarrow \quad n_{00}^{(2)} = 0,05 * 19904 = 995$$

olarak bulunur.

4.3. Logaritmik Doğrusal Modeller Kullanılarak Prevalans Tahmini

Prevalans tahmininin logaritmik doğrusal modeller kullanılarak analizini ve sonuçların yorumlanışını göstermek amacıyla, yakalama oranı $p=0,30$, $n=50$ ve yakalama sayısı $k=3$ olacak şekilde Ek'deki makro ile veri simülasyonu yapılmıştır.

Çizelge 4.7. Logaritmik doğrusal model için simülasyon sonucu elde edilen veriler

1. Yakalama	2. Yakalama	3. Yakalama	Sayı
1	1	1	11
1	1	0	1
1	0	1	3
1	0	0	3
0	1	1	12
0	1	0	26
0	0	1	6
0	0	0	?

Bu veriler için yapılan olası tüm logaritmik doğrusal model sonuçları Çizelge 4.8'de verilmiştir.

Çizelge 4.8. Tüm yakalamalar için logaritmik doğrusal model sonuçları

Model	L ²	sd	p	AIC	BIC
M ⁽⁰⁾	15,0676	3	0,0017	9,0676	9,5323
M ⁽¹⁾	15,0465	2	0,0005	11,0465	11,3439
M ⁽²⁾	3,8676	2	0,1446	-0,1324	0,1175
M ⁽³⁾	14,7469	2	0,0006	10,7469	11,0692
M ⁽⁴⁾	0,5616	1	0,4536	-1,4384	-1,5845
M ⁽⁵⁾	14,7125	1	0,0001	12,7125	12,8799
M ⁽⁶⁾	3,8676	1	0,0492	1,8676	1,9925
M ⁽⁷⁾	0	0		0	
M ⁽⁸⁾	0	-			

Model seçimine geriye dönük değişken eleme yöntemiyle başlanmış ve başlangıç model olarak doymuş model ele alınmıştır. Modelden çıkarıldığından L²'de anlamlı artış yaratmayan terimler modelden tek tek çıkarılır. Yakala tekrar yakala verisinde hiçbir kaynakta tespit edilemeyenlere ait göze boş olduğundan tüm kaynakların etkileşimi logaritmik doğrusal modelde söz konusu değildir (46). Bu yüzden başlangıç modeli M⁽⁷⁾'dir.

İşlem 1:

Tüm ana etkilerin ve ikili etkileşimlerin yer aldığı M⁽⁷⁾ modeli başlangıç olmak üzere başlangıç modeli kendisinden bir eksik parametre içeren modellerle karşılaştırılır.

$$L^2 (M^{(7)}) - L^2 (M^{(6)}) = 3,8676 > X^2_{(1)}$$

$$L^2 (M^{(7)}) - L^2 (M^{(5)}) = 14,7125 > X^2_{(1)}$$

$$L^2 (M^{(7)}) - L^2 (M^{(4)}) = 0,5616 < X^2_{(1)}$$

Bu durumda λ^{12} teriminin istatistiksel olarak anlamsız olduğu ve M⁽⁷⁾ yerine M⁽⁴⁾ modelinin kullanılabileceği sonucu elde edilmiştir.

İşlem 2: Başlangıç modeli kendisinden iki eksik parametre içeren modellerle karşılaştırılır.

$$L^2 (M^{(7)}) - L^2 (M^{(3)}) = 14,7469 > X^2_{(2)}$$

$$L^2 (M^{(7)}) - L^2 (M^{(2)}) = 3,8676 < X^2_{(2)}$$

$$L^2 (M^{(7)}) - L^2 (M^{(1)}) = 15,0465 > X^2_{(2)}$$

İşlem sonucunda $M^{(1)}$ modelinin $M^{(7)}$ yerine kullanılabileceği sonucu elde edilmiştir.

İşlem 3:

$$L^2 (M^{(7)}) - L^2 (M^{(0)}) = 15,0676 < X^2_{(3)}$$

Çizelge 4.8 incelendiğinde, $M^{(2)}$ ve $M^{(4)}$ modellerinin her ikisi veriye uyumlu model olduğu görülmüş, ancak hangisinin daha iyi bir model olduğu araştırılırken AIC ve BIC hesaplamalarından faydalanılmıştır. Elde edilen sonuca göre daha düşük AIC ve BIC değerine sahip $M^{(4)}$ modelinin kullanılacağına karar verilmiştir.

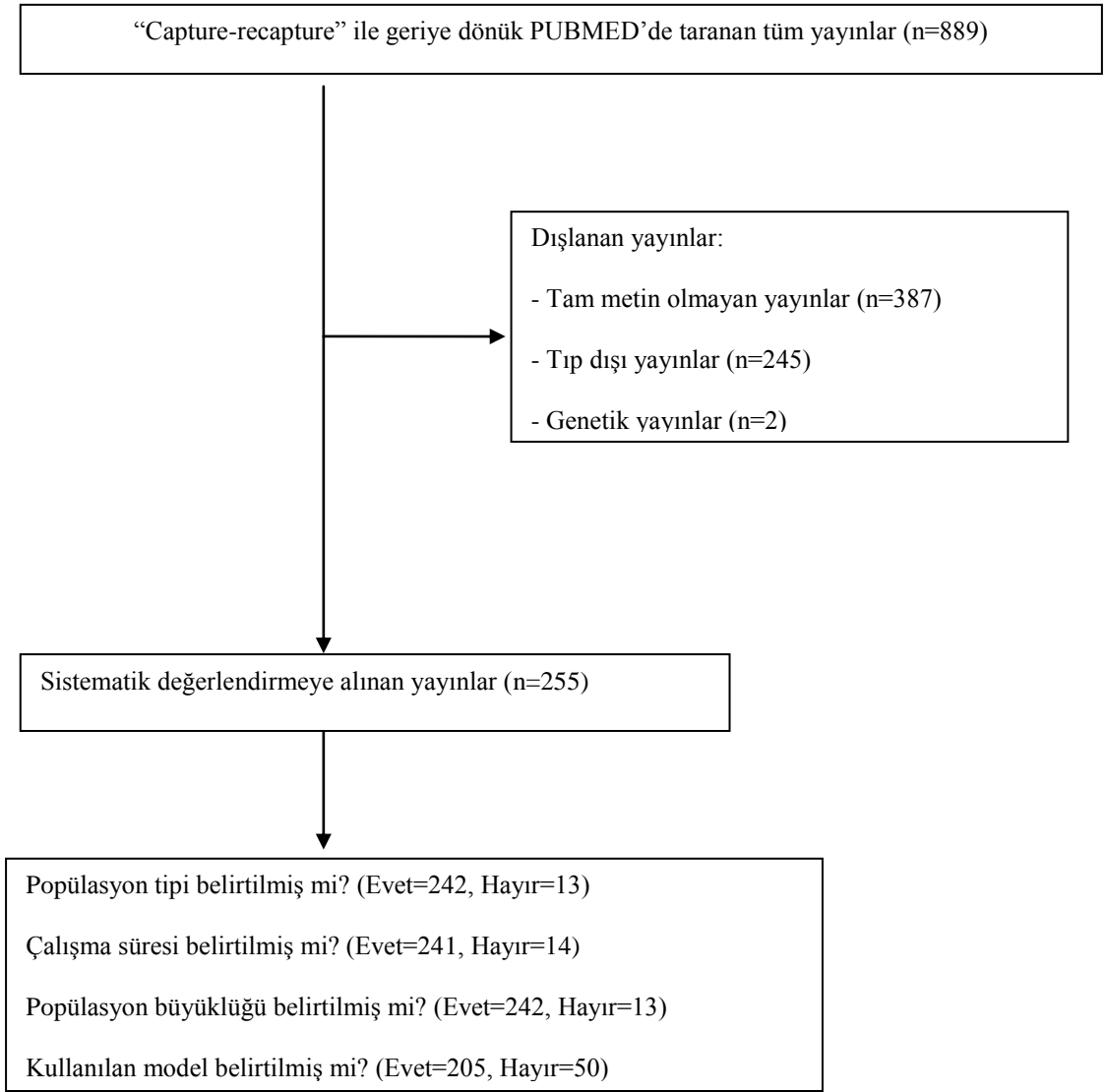
İlgili veri için kullanılacak model $\hat{y} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AC} + \lambda_{jkc}^{BC}$ olarak belirlenmiştir. Bu model kullanılarak elde edilen tahmini frekanslar Çizelge 4.9'daki gibidir. Böylece popülasyon genişliği $N=140$ olarak hesaplanır.

Çizelge 4.9. $\hat{y} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AC} + \lambda_{jkc}^{BC}$ modeli kullanılarak elde edilen tahmini frekanslar.

1. Yakalama	2. Yakalama	3. Yakalama	Tahmini Frekans
1	1	1	10,06
1	1	0	1
1	0	1	3,94
1	0	0	3
0	1	1	12,94
0	1	0	26
0	0	1	5,06
0	0	0	78

4.4. Literatür Taraması Bulguları

Çalışma kapsamına dahil edilmek üzere Mersin Üniversitesi internet erişim ağı üzerinden “capture recapture” anahtar tümcesiyle Pubmed veri tabanında geriye dönük sistemik olarak yapılan taramada Temmuz 1961 – Ağustos 2012 tarihleri arasında olduğu belirlenen 889 yayından tam metin, tıbbi ve uygulama olanlar detaylı olarak incelenmiştir. Tarama aşamaları Şekil 4.1'deki akış şeması ile verilmiştir.



Şekil 4.1. Literatür taraması için akış şeması

Değerlendirmeye alınan 889 yayından, ilk aşamada tam metin olmayan 387 yayın çalışma dışı bırakılmıştır. Daha sonra geriye kalan 502 tam metin yayından 245’i diğer türlerde (teorik çalışma, biyolojik çalışma, hayvan çalışması, derleme, gözden geçirme, editöre mektup) ve 2’si gen çalışması olmak üzere toplam 247’si çalışmadan dışlanmıştır. Dahil edilme kriterlerine uygun 255 yayın kalmıştır.

Bunlar popülasyon tipi (açık/kapalı), çalışma süresi, popülasyon büyüklüğü, yakalama (liste) sayısı ve kullanılan modeli belirtme durumlarına göre incelenmiştir.

Çalışmalarındaki hastalık veya tıbbi durumların, modellerin, popülasyon tipinin, tahminlenen istatistiklerin ve örnek genişliklerinin dağılımları Çizelge 4.10, Çizelge 4.11, Çizelge 4.12, Çizelge 4.13 ve Çizelge 4.14’de verilmiştir.

Çizelge 4.10. Çalışmalarda incelenen hastalık veya tıbbi durumların dağılımı.

Hastalık veya tıbbi durum	Sayı	Yüzde
Göz titremesi	1	0,4
Advers ilaç reaksiyonları	1	0,4
Aile içi şiddet	1	0,4
Akut flask paralizi	1	0,4
Akut serebellar ataksi	1	0,4
Alkolle ilişkili bozukluklar	1	0,4
Alt ekstremité amputasyon (LEA)	1	0,4
Amyotrofik lateral Skleroz	4	1,6
Anafilaksi	1	0,4
Ani bebek ölümü sendromu	1	0,4
Anne ölümü	1	0,4
Arteriovenöz malformasyon (AVM)	1	0,4
Ayak ve ağız hastalığı	1	0,4
B grubu streptokok enfeksiyonu	3	1,2
Bacak ülserleri	1	0,4
Bağımlılık	22	8,7
Boğmaca	2	0,8
Boğulma	1	0,4
Bruselloz	1	0,4
Bulaşıcı hastalık	2	0,8
Bulaşıcı olmayan hastalıklar	1	0,4
Buruli ülser	1	0,4
Canlı doğum	1	0,4
Churg-Strauss sendromu	1	0,4
Cinsel yolla bulaşan hastalıklar	1	0,4
Crohn hastalığı	1	0,4
Çocuk istismarı	1	0,4
Çocukluk konvülsif epileptikus durumları	1	0,4
Çölyak hastalığı	1	0,4
Demans	1	0,4
Dengue	1	0,4
Diyabet	6	2,4
Diyabetik ayak ülseri (DFU)	1	0,4
Doğum	1	0,4

Çizelge 4.10. (devam) Çalışmalarda incelenen hastalık veya tıbbi durumların dağılımı.

Doğum kusurları	2	0,8
Doğumsal kalça çıkığı	1	0,4
Down sendromu	3	1,2
Ekinokok hastalığı	1	0,4
Enfeksiyon hastalıkları	1	0,4
Epilepsi	1	0,4
Felç	2	0,8
Fetomaternal alloimmün trombositopeni	1	0,4
Gebelik kanaması	1	0,4
Gebelik sırasında sigara içimi	1	0,4
Gebelik sonlanımı	1	0,4
Genel tarama	1	0,4
Gestasyonel diyabet	1	0,4
Giardiasis	1	0,4
Gizli diyabet	1	0,4
Görsel bozukluklar ve oküler bozuklukları	1	0,4
Grip	2	0,8
Gut hastalığı	1	0,4
H1N1	2	0,8
Hemiplejik Migren	1	0,4
Hemofilus gribi	1	0,4
Hepatit	1	0,4
Hepatit A	1	0,4
Hepatit B	1	0,4
Hepatit C	3	1,2
HIV (AIDS)	15	5,9
Hipoglisemi	1	0,4
Hodgkin lenfoma	1	0,4
İmmünkompetan	1	0,4
İnfanıl spazm (IS)	1	0,4
İnflamatuvar barsak hastalığı	1	0,4
İnvajinasyon (besinin dokuya karışması)	1	0,4
İnvazif meningokok hastalığı	1	0,4
İnvaziv pnömokok hastalığı	2	0,8
Juvenil dermatomyozit	1	0,4
Kalça displazisi	1	0,4
Kalp malformasyonu	1	0,4
Kanser	13	5,1
Kas-iskelet sistemi hastalıkları	2	0,8
Kızamık	6	2,4
Konjenital hipotiroidi (CH)	1	0,4
Konjenital katarakt	1	0,4

Çizelge 4.10. (devam) Çalışmalarda incelenen hastalık veya tıbbi durumların dağılımı.

Konjenital malformasyonlar	1	0,4
Konjenital oküler anomalileri	1	0,4
Körlük	1	0,4
Kriptokokkozis	1	0,4
Kutanöz leishmaniasis (CL)	1	0,4
Langerhans hücreli histiyositoz (LCH)	1	0,4
Lejyoner hastalığı	3	1,2
Leptospirozis	1	0,4
Lösemi	1	0,4
Lupus nefriti	1	0,4
Melanom	1	0,4
Menenjit	1	0,4
Meningokok hastalığı	3	1,2
Mikroskopik polianjiitis	1	0,4
Motor nöron hastalıkları	1	0,4
Multipl skleroz	9	3,6
Nöral tüp bozukluğu	2	0,8
Omurilik yaralanması	2	0,8
Orak hücre hastalığı	2	0,8
Otistik spektrum bozukluğu	1	0,4
Otozomal resesif konjenital iktiyozis	1	0,4
Parkinson hastalığı	1	0,4
Pnömonok	3	1,2
Pnömonok menenjiti	1	0,4
Poliarteritis nodosa	1	0,4
Romatoid artrit	1	0,4
Rotavirüs	1	0,4
Salmonella enfeksiyonu	1	0,4
Serebral palsi	1	0,4
Sıtma	5	2
Sifiliz	1	0,4
Siliko	1	0,4
Sistemik skleroz	2	0,8
Splenektomi sepsis	1	0,4
Suçiçeği	1	0,4
Tetanoz	1	0,4
Tip 1 diyabet	23	9,1
Tip 2 diyabet	4	1,6
Tiroid kanseri	1	0,4
Travmatik beyin yaralanması	1	0,4
Tüberküloz	12	4,7
Varisella zoster virüsü (VZV)	1	0,4

Çizelge 4.10. (devam) Çalışmalarda incelenen hastalık veya tıbbi durumların dağılımı.

Wegener granulomatozu	2	0,8
Yaralanmalar	1	0,4
Toplam*	253	100

*Katlamalı

Çizelge 4.11. Çalışmalarda kullanılan modellerin dağılımı

Model	Sayı	Yüzde
Logaritmik Doğrusal	116	55
Lincoln-Petersen	92	44
Jolly-Seber	2	1
Toplam	210	100

*Katlamalı

Çizelge 4.12. Çalışmalarda incelenen popülasyon tiplerinin dağılımı

Popülasyon tipi	Sayı	Yüzde
Kapalı	242	99
Açık	2	1
Toplam	244	100

*Katlamalı

Çizelge 4.13. Çalışmalarda tahmin edilen istatistiklerin dağılımı.

Tahmin Edilen İstatistikler	Sayı	Yüzde
Doğruluk	1	0,4
Hastalık karakteristiği	1	0,4
Hastalık yükü	2	0,7
Hastaneye yatırma oranı	2	0,7
İnsidans	96	35,7
Mortalite	15	5,6
Ölüm hızı (death rate)	2	0,7
Popülasyon çeşitliliği	1	0,4
Popülasyon genişliği	68	25,3
Pozitif prediktif value (PPV)	1	0,4
Prevalans	71	26,4
Sensitivite	6	2,2
Spesifite	1	0,4
Tanımlayıcı istatistik	1	0,4
Yaşam oranı (survival rate)	1	0,4
Toplam*	269	100,0

* Katlamalı

Çizelge 4.14. Çalışmalarda incelenen örnek genişliklerinin dağılımı.

Örnek Genişliği	Sayı	Yüzde
<100	19	7,9
100-500	74	30,7
501-1000	35	14,5
1001-10000	81	33,6
>10000	32	13,3
Toplam	241	100

Bu çalışmanın amaçlarından birinin yakala tekrar yakala yöntemiyle prevalans tahmini olması sebebiyle 255 yayından prevalans hesabı yapan 71 yayın ayrıca incelenmiş ve Çizelge 4.15’de detaylı olarak verilmiştir.

Çizelge 4.15. Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp arařtırmalarının, çalıřılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.

Arařtırma Kaynak No	Çalıřma Süresi	Tahminlenen İstatistik	Hastalık/Tıbbi Durum	Popülasyon Tipi	Tahmin Yöntemi	Yakalama/Liste Sayısı	Örnek Geniřliđi
50	2001-2009	Prevalans, İnsidans	Multipl Skleroz (MS)	Kapalı	Lincoln-Petersen	2	95
51	2008	Prevalans, İnsidans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	3	4299
52	1985-1997	Prevalans, İnsidans	İnfantil Spazm (IS)	Kapalı	Lincoln-Petersen	2	69
53	2010	Prevalans	Hipoglisemi	Kapalı	Logaritmik Doğrusal	3	37898
54	1989-1991	Prevalans, İnsidans, Yařam ve Ölüm Karakteristiđi	Sistemik Skleroz (SSc)			5	1596
55	1998-2002	Prevalans, Ölüm Oranı	Amyotrofik Lateral Skleroz (ALS)	Kapalı	Logaritmik Doğrusal	3	58
56	2004-2005	Prevalans, Mortalite	Bađımlılık	Kapalı	Logaritmik Doğrusal	5	5334
57	2003-2004	Prevalans, İnsidans	Tip 1 Diyabet	Kapalı	Lincoln-Petersen	2	130
58	2001	Prevalans, İnsidans	Lupus Nefriti	Kapalı		5	208
59	1997-1998	Prevalans, İnsidans	Bacak Ülseri	Kapalı	Logaritmik Doğrusal	4	426
60	1997-2002	Prevalans	Wegener granülomatozisi (WG)	Kapalı	Logaritmik Doğrusal	3	58
61	1995-1996	Prevalans	Tip 2 Diyabet	Kapalı	Logaritmik Doğrusal	6	2585
62	1991	Prevalans	Tip 2 Diyabet	Kapalı	Lincoln-Petersen	2	290

Çizelge 4.15. (devam) Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp arařtırmalarının, çalıřılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.

63	2001	Prevalans	Sistemik Skleroz (SSc)	Kapalı	Logaritmik Doğrusal	4	119
64	1993-1995	Prevalans	Splenektomi Sepsis	Kapalı	Lincoln-Petersen	2	387
65	1993-1995	Prevalans	Gebelikte sigara içimi	Kapalı	Lincoln-Petersen	2	
66	2006-2008	Prevalans	Orak Hücre Hastalığı (SCD)	Kapalı	Logaritmik Doğrusal	3	374
67	2003-2007	Prevalans	Orak Hücre Hastalığı (SCD)	Kapalı	Lincoln-Petersen	2	107
68	1984-2003	Prevalans	Serebral Palsi	Kapalı	Logaritmik Doğrusal		2337
69	2000	Prevalans	Poliarteritis Nodosa, Mikroskopik Polianjitis, Wegener Granüloatozu, Churg-Strauss Sendromu	Kapalı	Logaritmik Doğrusal	3	121
70	1997-1999	Prevalans	Parkinson Hastalığı (PD)	Kapalı	Lincoln-Petersen	2	5062
71	2010	Prevalans	Otozomal Resesif Konjenital İktiyozis	Kapalı	Logaritmik Doğrusal	3	144
72	2003	Prevalans	Nistagmus (Göz Titremesi)	Kapalı	Logaritmik Doğrusal	3	396
73	6 hafta	Prevalans	Bulařıcı Olmayan Hastalıklar (NCDs)	Kapalı		3	2455
74	1968-1990	Prevalans	Nöral Tüp Bozukluęu (NTD)	Kapalı	Lincoln-Petersen	2	549
75	1988-1998	Prevalans	Nöral Tüp Bozukluęu (NTD)	Kapalı	Logaritmik Doğrusal	5	2312
76	1995-2001	Prevalans	Kas İskelet Bozuklukları (MSD)	Kapalı	Logaritmik Doğrusal	3	15988

Çizelge 4.15. (devam) Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp arařtırmalarının, çalıřılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.

77	2003-2008	Prevalans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	4	1528851
78	1998-2002	Prevalans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	3	321
79	1985-1996	Prevalans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	4	727
80	2006	Prevalans	Multipl Skleroz (MS)			5	2917
81	2006	Prevalans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	6	2917
82		Prevalans	Multipl Skleroz (MS)	Kapalı	Logaritmik Doğrusal	4	803
83	2008-2009	Prevalans	Körlük	Kapalı	Logaritmik Doğrusal	3	1771
84		Prevalans	Gizli Diyabet	Kapalı	Logaritmik Doğrusal	3	2585
85	1995-2000	Prevalans	Bağımlılık	Kapalı	Lincoln-Petersen	2	
86	2007	Prevalans	İnsan Bağışıklık Yetmezlik Virüsü (HIV)				395000
87	2001	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	8141
88	1977-2009	Prevalans	Hepatit C	Kapalı	Logaritmik Doğrusal	4	6935
89	1994-1997	Prevalans	Hemiplejik Migren	Kapalı	Logaritmik Doğrusal	3	195
90	1993-1994	Prevalans	Kalp Malformasyonu	Kapalı	Logaritmik Doğrusal	6	522
91	1997	Prevalans	Epilepsi	Kapalı	Lincoln-Petersen	2	3134

Çizelge 4.15. (devam) Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp arařtırmalarının, çalıřılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.

41	1989-1991	Prevalans	Down Sendromu	Kapalı	Logaritmik Doğrusal	5	198
92	1995-1996	Prevalans	Diyabet	Kapalı	Logaritmik Doğrusal	6	2585
93	2003	Prevalans	Diyabet	Kapalı	Logaritmik Doğrusal	3	34420
94	1999	Prevalans	Diyabet	Kapalı	Logaritmik Doğrusal	3	368
95	1998	Prevalans	Diyabet	Kapalı		2	9178
96		Prevalans	Diyabet	Kapalı	Logaritmik Doğrusal	6	7596
97	1998-2000	Prevalans	Diyabetik Ayak Ülseri (DFU)	Kapalı	Lincoln-Petersen	2	1500
98	1988-1998	Prevalans	Demans	Kapalı	Logaritmik Doğrusal	3	14769
99		Prevalans	Buruli Ülser	Kapalı	Lincoln-Petersen	2	103000
100	2005-2007	Prevalans	Bağımlılık	Kapalı	Lincoln-Petersen	2	880
101	2005-2006	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	6	5657
102	2005-2006	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	4	
103	2001-2002	Prevalans	Bağımlılık	Açık ve Kapalı	Jolly-Seber, Logaritmik Doğrusal	3	761
104	2000-2001	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	4117

Çizelge 4.15. (devam) Prevalans tahmini amacıyla 2012 yılına kadar sistematik olarak derlenen tıp arařtırmalarının, çalıřılan hastalık / tıbbi durum, popülasyon tipi, kullanılan tahmin yöntemi, yakalama sayısı ve örnek genişlikleri.

105	2000-2001	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	5	1448
106	1999-2002	Prevalans	Bağımlılık	Kapalı	Lincoln-Petersen	2	8396
107	1992-2002	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	4	
108	1992-1995	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	1832
109	1990-1994	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	4	855
110	2002	Prevalans	Bağımlılık	Kapalı	Lincoln-Petersen	2	2063
111	1999	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	4035
112	1996	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	3170
113	1993	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	4	2310
114	1993	Prevalans	Bağımlılık	Kapalı	Logaritmik Doğrusal	3	3207
115	2001	Prevalans	Otistik Spektrum Bozukluğu	Kapalı	Lincoln-Petersen	2	13661
116	1998	Prevalans	Arteriovenöz Malformasyonlar (AVM)	Kapalı	Logaritmik Doğrusal	3	93
117	1997	Prevalans	Alkolle İliřkili Bozukluklar	Kapalı	Logaritmik Doğrusal	4	349
118	2003-2004	Prevalans	Aile İçi Şiddet	Kapalı	Logaritmik Doğrusal	3	149000
119	32 hafta	Prevalans, Mortalite	Genel tarama	Kapalı	Logaritmik Doğrusal	4	18081

4. TARTIŞMA

Bu çalışmada biyoloji, sosyoloji, sağlık gibi geniş bir kullanım alanı olan yakala-tekrar yakala tahmin yönteminin tüm detaylarıyla incelenmesi ve sanıldığı aksine oldukça yalın ve anlaşılır olan hesaplama adımlarının teorik ve uygulamalı olarak gösterilmesi amaçlanmıştır. Bu amaçla yapılan uygulamalar, yöntemin sağlık alanında epidemiyolojik çalışmaların yanı sıra tanı testlerinin performansları değerlendirilirken karşılaşılan eksik veri durumunda da kullanımının oldukça kolay olduğunu göstermektedir.

Özellikle epidemiyolojik çalışmalarda, kayıtlar arasındaki tutarsızlık sebebiyle, hastalık prevalansı, doğum ve ölüm hızlarının tahminlerinin önemli bir sorun olduğu, fakat yakala tekrar yakala yöntemi kullanılarak bu sorunun kolaylıkla aşılabileceği görülmüştür.

Belirli bir toplumun kayıt merkezlerinden elde edilen hastalık istatistikleri, diğer toplumlarla veya farklı zamanlarla karşılaştırılabilir olmalıdır. Karşılaştırılabilirliğin temelinde ise kayıt işleminin evrensel standartlar ve kurallar çerçevesinde yapılması vardır. Hedef nüfustaki bireylerin hepsi liste kayıtlarında bulunuyor ise veri tamdır ve bunu kontrol etmek için kullanılan yöntemlerden biri de yakala tekrar yakala yöntemidir.(10)

Yakala tekrar yakala yönteminin, tanı testlerinin performanslarını değerlendirmeye ve gerçek hastalık prevalansını tahmin etmeye yönelik çalışmalarda da kullanımının oldukça kolay olduğu görülmektedir. Özellikle altın standardın kullanımının zor veya pahalı olduğu veya ancak belirli koşullar altında uygulandığı durumlarda tahmin aşamasında önemli bir sorunla karşılaşılmaktadır. Bu sorun bazı vakaların hastalık durumunun altın standart ile doğrulanmamış olmasıdır. Bu durumda, tanı testlerinin performansını değerlendirmek ancak yakala tekrar yakala yöntemiyle mümkündür.

Tüm bu değerlendirmelerden sonra özenle üstünde durulması gereken konu, yakala tekrar yakala yönteminin varsayımlarıdır. Tahmin edilmeye çalışılan parametre ne olursa olsun, ilk olarak dikkat edilmesi gereken popülasyon tipidir. Popülasyonun kapalı veya açık oluşuna, yakalama sayısına göre kullanılacak tahmin yönteminin değiştiği

bilinmektedir. Çalışmamızda, Lincoln-Petersen yönteminin sadece iki yakalama söz konusu iken kullanıldığı, ikiden çok kez yakalama durumunda ise popülasyonun açık veya kapalı oluşuna göre Jolly-Seber veya Schnabel ve Schmacher-Eschmeyer yöntemlerinin kullanıldığı teorik ve uygulamalı olarak gösterilmiştir. Bunun yanı sıra, popülasyon kapalı iken kullanılabilen Schnabel ve Schumacher-Eschmeyer yöntemlerinin arasında pozitif doğrusal ilişki varsayımı bakımından bir farklılık olduğu gösterilmiştir.

Bütün bu tahmin yöntemlerinin ortak varsayımı, ilk örnekleme yakalanma şansı ile bir sonrakilerde yakalanma şansının eşit ve bağımsız olmasıdır. Akut hastalıklarla ilgili listelerde ortak birim sayısının daha az olması beklenir. Ancak kronikleşmiş veya enfeksiyon hastalıkları gibi çok ciddi hastalıklarda kliniğe bir kez gelen bir hastaya verilen tedavi veya konulan tanının hastanın tekrar tekrar gelme olasılığını etkileyebileceği aşıkardır. Bu tür problemlerle başa çıkan istatistiksel yöntemler vardır. Bunlardan biri de logaritmik doğrusal modellerdir. Çalışmamızda, yakalama olasılığı $p=0,30$ olacak şekilde, üç yakalamalı veri seti üretilmiş ve logaritmik doğrusal modeller kullanılarak eksik veri tahmin edilmiştir. Burada yakalama olasılığının çok küçük veya çok büyük oluşu eksik verinin olduğundan fazla veya olduğundan az tahmin edilmesine sebep olduğundan, 0,30-0,70 aralığında bir değer kullanılmıştır.

Çalışmamızda, yakala tekrar yakala yöntemlerinin teorik ve uygulamalı gösterimlerinden sonra, literatürdeki yeri, sıklığı, özellikle tercih edildiği alanlar ve kullanılan modelleri incelemek amacıyla, günümüze kadar yapılmış olan çalışmalar sistematik değerlendirmeye alınmıştır. Yapılan sistematik değerlendirme sonucunda, “capture recapture” tümcesiyle taranan 889 çalışmadan 644’ünün sağlık alanına özel çalışma olduğu, ancak bunlardan 255’inin değerlendirilebilir olduğu tespit edilmiştir. Bu çalışmalar içinden seçilen 71 çalışmanın da prevalans tahmini amacıyla yapılmış olduğu belirlenmiştir. Bu çalışmalar incelendiğinde, 68 çalışmada popülasyon tipi kapalı olarak belirtilmiş iken 3 çalışmada belirtilmemiştir. Çalışmaların 65’inde kullanılan tahmin modeli belirtilmiş olup, büyük çoğunluğunu (48/65) logaritmik doğrusal modeller oluşturmaktadır. Jolly-Seber yönteminin kullanıldığı bir çalışmada ise bu yöntemin tercih sebebinin popülasyonun büyüme, mobilite ve mortalitesini de hesaplıyor olması olarak belirtilmiştir (103). Burada dikkat çekici olan, yakalama sayısı ikiden fazla olan çalışmalarda, Schnabel ve Schumacher-Eschmeyer yöntemlerinin hiç tercih edilmemiş

olmasıdır. Bunun sebebi ise, alıřmalardaki listeler arası bağımsızlık varsayımının bozulmasıdır.

SONUÇ VE ÖNERİLER

Gün geçtikçe kalabalıklaşan popülasyonlar pek çok araştırma için kanıt değeri yüksek sonuçları ifade etmektedir. Oysa artan kalabalıkla birlikte kontrolü zorlaşan gürültülü bir veri topluluğu oluşmaktadır. Sağlığın söz konusu olduğu tüm çalışmalar için veri kalitesi büyük önem taşımaktadır. Sağlık alanında kullanımı yakın geçmişte artan yakala tekrar yakala yönteminin incelendiği bu çalışmada iki ya da daha çok veri kaynağı kullanılarak popülasyon genişliği, prevalans, insidans, mortalite, sensitivite ve spesifite gibi hesaplamalar için daha doğru sonuçlar elde edilebileceği gösterilmeye çalışılmıştır.

Çalışmamızda, hipotetik ve benzetim tekniği ile üretilen veriler üstünde yakala tekrar yakala yöntemlerinin nasıl kullanılacağı gösterilmiştir. İki kaynak, liste ya da yakalama durumunda basit matematiksel yöntemlerle hızlı ve güvenilir sonuç alınabileceği ikiden çok kaynak, liste ya da yakalama söz konusu olduğunda ise listelerin bağımlılık araştırması için logaritmik doğrusal modellerden faydalanılabileceği görülmüştür.

Çalışmaya başlarken yakala tekrar yakala yönteminin kullanılacağı çalışmalara dahil edilecek verinin elde ediliş biçiminin retrospektif ya da prospektif olması merak edilen bir konuydu. Tarama sonucu elde edilen sağlıkla ilgili 255 çalışmanın yaklaşık %38'inde insidans, %28'inde prevalans ve %27'sinde mortalite hesabı yapıldığı görülmüştür. Bu sonuca göre retrospektif ve prospektif biçimlerde iki tür verinin de bu yöntemde kullanılabilir olduğu görülmüştür.

Yakala tekrar yakala yöntemlerinin sağlık alanına özel kullanım alanlarını, hem teorik hem de uygulamalı olarak gösteren ve ayrıca günümüze kadar yapılmış çalışmalarda kullanılış biçimlerini sistematik olarak değerlendirerek geçerliliklerini kontrol eden bu kadar kapsamlı bir çalışma olmadığı tespit edilmiştir.

Sistematik değerlendirme sonucunda, prevalans tahmini yapılan yakalama sayısı ikiden fazla olan çalışmalarda listeler arası bağımsızlık varsayımının bozulması nedeniyle, Schnabel ve Schumacher-Eschmeyer yöntemlerinin hiç kullanılmadığı tespit edilmiştir. Çalışmalarda kullanılan tahmin yöntemlerinin geçerliliği açısından bu sonuç önemlidir.

Sensitivite ve spesifite deęerleri düşük referans tanı testleri sonrasında altın standart ile deęerlendirilmesine karar verilen hasta sayısında yapılabilecek hataları önlemek adına yakala tekrar yakala yönteminin kullanımı oldukça pratiktir. Son 10 yılda yakala tekrar yakala yöntemlerinin kullanıldığı makale sayısı oldukça artmıştır (120).

Prevalans ve diagnostik doğruluk üstüne planlanan bu çalışmadan sonra benzer şekilde insidans ve popülasyonda büyüme, mobilite ve mortalite oranlarının daha doğru belirlenmesi için yeni çalışmalar planlanabilir.

Yakala tekrar yakala yönteminin kullanıldığı çalışmalarda elde edilecek sonuçların hangi sıraya göre verilmesi gerektiğiyle ilgili bir kontrol listesi (checklist) hazırlanabilir.

Literatür taraması sonucunda elde edilen ortak özellikli çalışmalara meta analizi yapılarak kanıt deęeri daha yüksek çalışmalar elde edilebilir.

KAYNAKLAR

1. **Van Hest R.** Capture-recapture Methods in Surveillance of Tuberculosis and Other Infectious Diseases, Erasmus tezi, Rotterdam, the Netherlands, **2007**.
2. **Morrison A, Stone DH.** Capture-recapture: a useful methodological tool for counting traffic related injuries?. *Inj Prev*, **2000**; 6: 299-304.
3. **Brittain S, Böhning D.** Estimators in capture–recapture studies with two sources. *Asta-Advances in Statistical Analysis*. **2009**; 93 (1): 23-47.
4. **Özmen Ahmet,** Örneklemeye, Erişim adresi: <http://w2.anadolu.edu.tr/aos/kitap/ioltp/2294/unite03.pdf>, Erişim tarihi: 05.10.2012
5. **Sümbüloğlu V, Sümbüloğlu K,** Klinik ve saha araştırmalarında Örneklem Yöntemleri ve Örneklem Büyüklüğü. 1.Basım, Ankara, **2005**.
6. **Yamane T.** Temel Örneklemeye Yöntemleri. 1.baskı, (İstanbul: Literatür Yayınları: 53, (Çev: Alptekin Esin vd.) **2006**.
7. **Keser İ.** Anketler Aracılığı ile Örneklemeye Dayalı Olarak Yapılan Araştırmalarda İzlenecek Adamlar ve Karşılaşılabilecek Sorunlar. *Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, **2006**; 8(4): 128-147.
8. **Budakoğlu İİ, Akgün HS.** Sağlık Alanında Veri Toplama: Yakalama-Yeniden Yakalama Yöntemi. *Kor Hek*, **2008**; 7(1): 75-80.
9. Türkiye ile Avrupa arasında hasta veri kayıtları paylaşımı; Erişim adresi: <http://www.medikalplus.com/hastane/331-turkiye-avrupa-arasinda-hasta-veri-kayitlari-paylasimi.html>, Erişim tarihi: 14.11.2012
10. **Yalcın Eser S.** Kanser Kayıtçılığı ve Kanser Kayıt Merkezleri. Ed. Tuncer M. Türkiye’de Kanser Kontrolü. T.C. Sağlık Bakanlığı Kanserle Savaş Dairesi Başkanlığı Yayınları, **2007**, 47-77.
11. **Strax P, Venet L, Shapiro S, Gross S.** Mammography and Clinical Examination in Mass Screening for Cancer of the Breast. *Cancer*, **1967**; 20: 2184–2188.
12. **Hook BE, Regal RR.** Capture Recapture Methods in Epidemiology: Methods and Limitations. *Epidemiologic Reviews*, **1995**; 13(2): 243-264.
13. **Chandrasekar C, Deming WE.** On a method of estimating birth and death rates and the extent of registration. *J Am Stat Assoc*, **1949**;44:101-15.

14. **Mirtaghizadeh H, Kesici T.** Popülasyon hacminin yakalama-tekrar yakalama yöntemi kullanılarak ters tahmin yöntemi ile tahmini. *YYÜ, Ziraat Fakültesi, Tarım bilimleri dergisi*, **2003**; 13(2): 113-118.
15. **Verma V, Betti G.** Investigation into the feasibility of a new survey instrument targeted at people moving across the EU, Final Report. **2008**; 136-149.
16. **Lohr S L.** Sampling Design and Analysis. 2nd Ed., Duxbury Pres, **1999**.
17. **Paneth N.** Invited commentary: the hidden population in perinatal epidemiology. *Am J Epidemiol* **2008**; 167(7): 793–796.
18. **Dorak MT.** Bias & Confounding. Ders notu. Erişim adresi: <http://www.dorak.info/epi/bc.html>
Erişim tarihi: 11.12.2012
19. Ekoloji Laboratuvar Kitabı, Ekolojide Kullanılan Temel Örneklemeye Yöntemleri, Bölüm 3, **2006**;
Erişim adresi: yunus.hacettepe.edu.tr/~cagasan/images/BOLUM%2003.pdf, Erişim tarihi:
12.12.20012
20. Family Health International – Philippines. Training on Size Estimation Methods for Populations at High Risk for HIV. **2005**.
21. **Gill GV, Ismail AA, Beeching NJ.** The use of capture-recapture techniques in determining the prevalence of type 2 diabetes. *Q J Med*, **2001**; 94: 341-346.
22. **Mingoti S A, Caiiffa WT.** A capture-recapture technique to estimate the size of the injecting drug user population attending syringe exchange programs: AJUDE-Brasil II Project, Cad. Saúde Pública. **2006**; 22(4).
23. **Chao A, Pan H Y, Chiang,S C.** The Petersen-Lincoln Estimator and Its Extension to Estimate the Size of a Shared Population. *Biometrical Journal*, **2008**; 50: 957-970.
24. **Pollock K H, Nichols J D, Brownie C, Hines. J E.** Statistical inference for capture-recapture experiments. *Wildlife Monographs*, **1990**; 107:1–97.
25. **Schwarz C J, Arnason A N.** Chapter 12 Jolly-Seber models in MARK. In Cooch E, White G. Eds. Program MARK - A Gentle Introduction 11th ed. **2012**.
26. **Ogle D.** FishR Vignette - Open Mark-Recapture Abundance Estimates. Ders notu:(27.11.2012)
Northland College.
27. **Jolly GM.** Explicit estimates from capture-recapture data with both death and immigration stochastic model. *Biometrika*, **1965**; 52:225-247.

28. **Manly BFJ.** Obtaining Confidence Limits on Parameters of the Jolly-Seber Model for Capture-Recapture Data. *Biometrics*, **1984**; 749-758.
29. **Schwarz C J.** The Jolly-seber model: More than just abundance. *Journal of Agricultural, Biological, and Environmental Statistics*, **2001**; 6(2): 195-205.
30. **Krebs CJ.** *Ecological Methodology*. 2nd Edition, Harper and Row, Publishers. New York. **1998**: 654
31. **Schnabel ZE.** The estimation of total fish populations of a lake. *Am Math Monthly*, **1938**; 45: 348-352.
32. **Schmacher FX, Eschmeyer RW.** The estimation of fish populations in lakes and ponds. *J Tennessee Acad Sci*, **1943**; 18: 228-249.
33. **Zar JH.** *Biostatistical Analysis*. 5th Edition, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, **2009**.
34. **Lu Y, Fan JQ.** *Advanced Medical Statistics*, Singapore: World Scientific Pres, **2003**.
35. **Hook EB, Regal RR.** Capture-Recapture Methods in Epidemiology: Methods and Limitations. *Epidemiol Rev*, **1995**; 17(2): 243-264.
36. **Van HT, Singhasivanon P, Kaewkungwal J, Suriyawongpaisal P, Khai LH.** Estimation of non-fatal road traffic injuries in Thai Nguyen, Vietnam using capture-recapture method. *Southeast Asian J Trop Med Public Health*, **2006**; 37(2): 405-411.
37. **Davas A.** *Surveyans*. Eriřim adresi: http://halksagligi.med.ege.edu.tr/seminerler/2006-07/Surveyans_AD.pdf. Eriřim tarihi: 21.09.2012.
38. **Böhning D, Patilea V.** A Capture-Recapture Approach for Screening Using Two Diagnostic Tests With Availability of Disease Status for the Test Positives Only. *Journal of the ASA*, **2008**; 103(481): 212-221.
39. *Encyclopedia of Biostatistics* [8 vols] 2nd ed - P. Armitage, T. Colton (eds). Wiley, **2005**.
40. **Yule GU.** On the association of attributes in statistics: with illustration from the material of the childhood society, &c. *Philosophical Transactions of the Royal Society of London, Series A*, **1900**; 194: 257-319.
41. **Orton H, Rickard R, Miller L.** Using Active Medical Record Review and Capture-Recapture Methods to Investigate the Prevalence of Down Syndrome among Live-Born Infants in Colorado. *Teratology*, **2001**; 64: 14-19.

42. **Jeansonne A.** Log-Linear Models, Erişim Adresi; www.education.umd.edu/EDMS/fac/.../LogLinearModels%20reading.pdf. Erişim tarihi: 15.01.2012
43. **Bartlett M S.** Contingency table interactions. *Journal of the Royal Statistical Society, Supplement* **1935**; 2: 248–252
44. **Johnson J B, Omland K S.** Model selection in ecology and evolution. *Trends in Ecology and Evolutio.* **2004**; 19:101–108.
45. **Ripplinger J, Sullivan J.** Does choice in model selection affect maximum likelihood analysis? *Systematic Biology*, **2008**; 57: 76-85.
46. **Cormack RM.** Log-Linear Models for Capture-Recapture. *Biometrics*, **1989**; 45(2): 395-413.
47. **Eser S, Yakut C, Özalan S, Özdemir R, Göksel T.** İzmir’de akciğer kanseri hızı gerçekten yüksek mi? Kalite kontrolü olarak bir adres kaydı doğrulama çalışması. *Türkiye Halk Sağlığı Dergisi*, **2009**; 7(2): 53-60.
48. **Şen C, Yeşilova A, Kaki B.** Van İlindeki Kanser Hastaları Populasyonunun Büyüklüğünün Tahmini. **2004**; 342-349.
49. **Becanım C.** Log-lineer modeller ve doktor- tıbbi satış mümessilleri ilişkileri üzerine bir uygulama. Yüksek lisans tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2006**.
50. **Alcalde-Cabero E, Almazán-Isla J, Sempere A, Pedro-Cuesta JD, Sá JD.** Capture-recapture as a potentially useful procedure for assessing prevalence of multiple sclerosis: methodologic exercise using Portuguese data. *Neuroepidemiology*, **2012**; 38(4): 209-216.
51. **Addsi HE, Debouverie M, Guillemin F.** Estimating the prevalence and incidence of multiple sclerosis in the Lorraine region, France, by the capture-recapture method. *Mult Scler*, **2012**; 18(9): 1244-1250.
52. **Chen C, Chen T, Lin H, Oon P, Wu H, Wang PJ, Chen TH, Liou HH.** Estimation of prevalence and incidence of infantile spasms in Taiwan using capture-recapture method. *Epilepsy Res*, **2004**; 58(1): 37-42.
53. **Nirantharakumar K, Marshall T, Hodson J, Narendan P, Deeks J, Coleman J, Ferner RE.** Hypoglycemia in non-diabetic in-patients: clinical or criminal? *PLoS One*, **2012**; 7(7):e40384
54. **Mayes MD, Lacey JV Jr, Beebe-Dimmer J, Gillespie BW, Cooper B, Laing TJ, Schottenfeld D.** Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. *Arthritis Rheum*, **2003**; 48(8):2246-2255.

55. **Turabelidze G, Zhu B, Malone J, Horowitz S, Weidinger J, Williamson D, Simoes E.** An epidemiologic investigation of amyotrophic lateral sclerosis in Jefferson County, Missouri. *Neurotoxicology*, **2008**; 29(1), 81-86.
56. **Hickman M, Coleman B, Hope V, Parry J, Telfer M, Twigger J, Irish C, Macleod J, Annett H.** Assessing IDU prevalence and health consequences (HCV, overdose and drug-related mortality) in a primary care trust: implications for public health action. *J Public Health*, **2009**; 31(3): 374-382.
57. **Bahillo M, Hermoso F, Ochoa C, Garcia-Fernandez J, Rodrigo J, Marugan JM, de la Torre S, Manzano F, Lema T, García-Velázquez J.** Incidence and prevalence of type 1 diabetes in children aged <15 yr in Castilla-Leon (Spain). *Pediatr Diabetes*, **2007**; 8(6): 369-373.
58. **Patel M, Clarke A, Bruce I.** The prevalence and incidence of biopsy-proven lupus nephritis in the UK: Evidence of an ethnic gradient. *Arthritis Rheum*, **2006**; 54(9): 2963-2969.
59. **Walker NK, Vandal AC, Holden JK, Rodgers A, Birchall N, Norton R, Triggs CM, MacMahon S.** Does capture-recapture analysis provide more reliable estimates of the incidence and prevalence of leg ulcers in the community. *Aust N Z J Public Health*, **2002**; 26(5): 451-455.
60. **Mohammad AJ, Jacobsson LT, Mahr AD, Sturfelt G, Segelmark M.** Prevalence of Wegener's granulomatosis, microscopic polyangiitis, polyarteritis nodosa and Churg-Strauss syndrome within a defined population in southern Sweden. *Rheumatology*, **2007**; 46(8): 1329-1337.
61. **Ismail AA, Beeching NJ, Gill GV, Bellis MA.** Capture-recapture-adjusted prevalence rates of type 2 diabetes are related to social deprivation. *QJM*, **1999**; 92(19): 707-710.
62. **Sekikawa A, Eguchi H, Tominaga M, Manaka H, Sasaki H, Chang YF, Kato T.** Evaluating the reported prevalence of type 2 diabetes mellitus by the Oguni diabetes registry using a two-sample method of capture-recapture. *Int J Epidemiol*, **1999**; 28(3): 498-501.
63. **Guern VL, Mahr A, Mouthon L, Jeanneret D, Carzon M, Guillevin L.** Prevalence of systemic sclerosis in a French multi-ethnic county. *Rheumatology*, **2004**; 43(9): 1129-1137.
64. **Sarangi J, Coleby M, Trivella M, Reilly S.** Prevention of post splenectomy sepsis: a population based approach. *J Public Health Med*, **1997**; 19(2): 208-212.
65. **Dietz PM, Adams MM, Kendrick JS, Mathis MP.** Completeness of ascertainment of prenatal smoking using birth certificates and confidential questionnaires: variations by maternal attributes and infant birth weight. *Am J Epidemiol*, **1998**; 148(11): 1048-1054.
66. **Ilozue C, Cipolotti R, Melo CA, Gurgel RQ, Cuevas LE.** Estimating the post-neonatal prevalence of sickle cell disease in a Brazilian population. *Trop Med Int Health*, **2010**; 5(10): 1125-1131.

67. **Peters M, Fijnvandraat K, Tweel XW, Garre FG, Giordano PC, Wouwe JP, Pereira RR, Verkerk PH.** One-third of the new paediatric patients with sickle cell disease in The Netherlands are immigrants and do not benefit from neonatal screening. *Arch Dis Child*, **2010**; 95(10): 822-825.
68. **Surman G, Silva AA, Kurinczuk J J.** Cerebral palsy registers and high-quality data: an evaluation of completeness of the 4Child register using capture-recapture techniques. *Child Care Health Dev*, **2012**; 38(1): 98-107.
69. **Mahr A, Guillevin L, Poissonnet M, Ayme S.** Prevalences of polyarteritis nodosa, microscopic polyangiitis, Wegener's granulomatosis, and Churg-Strauss syndrome in a French urban multiethnic population in 2000: a capture-recapture estimate. *Arthritis Rheum*, **2004**; 51(1): 92-99.
70. **Bertoni JM, Strickland D.** Parkinson's prevalence estimated by a state registry. *Mov Disord*, **2004**; 19(3): 318-323.
71. **Hernandez-Martin A, Garcia-Doval I, Aranegui B, Unamuno P, Rodriguez-Pazos L, Gonzales-Ensenat MA, Vicente A, Martín-Santiago A, Garcia-Bravo B, Feito M, Baselga E, Círia S, de Lucas R, Ginarte M, González-Sarmiento R, Torrelo A.** Prevalence of autosomal recessive congenital ichthyosis: a population-based study using the capture-recapture method in Spain. *J Am Acad Dermatol*, **2012**; 67(2): 240-244.
72. **Sarvananthan N, Surendran M, Roberts EO, Jain S, Thomas S, Shah N, Proudlock FA, Thompson JR, McLean RJ, Degg C, Woodruff G, Gottlob I.** The prevalence of nystagmus: the Leicestershire nystagmus survey. *Invest Ophthalmol*, **2009**; 50(11): 5201-5206.
73. **Gill GV, Scott B, Beeching NJ, Wilkinson D, Ismail AA.** Enumeration of non-communicable disease in rural South Africa by electronic data linkage and capture-recapture techniques. *Trop Med Int Health*, **2001**; 6(6): 435-441.
74. **Hey K, O'Donnell M, Murphy M, Jones N, Botting B.** Use of local neural tube defect registers to interpret national trends. *Arch Dis Child Fetal Neonatal Ed*, **1994**; 71(3): 198-202.
75. **Bruin KM, Heijden PG, Buitendijk SE, Ouden AL.** Periconceptional folic acid use and the prevalence of neural tube defects in The Netherlands. *Eur J Obstet Gynecol Reprod Biol*, **2003**; 108(1): 33-39.
76. **Morse T, Dillon C, Kenta-Bibi E, Diva U, Warren N, Grey M.** Trends in work-related musculoskeletal disorder reports by year, type, and industrial sector: a capture-recapture analysis. *Am J Ind Med*, **2005**; 48(1): 40-49.
77. **Fernandez O, Fernandez V, Guerrero M, Leon A, Lopez-Madrona JC, Alonso A, Bustamante R, Tamayo JA, Romero F, Bravo M, Luque G, García L, Sanchis G, San Roman C, Romero M, Papais-Alvarenga M, de Ramon E.** Multiple sclerosis prevalence in Malaga, Southern Spain estimated by the capture-recapture method. *Mult Scler*, **2012**; 18(3): 372-376.

78. **Turabelidze G, Schootman M, Zhu BP, Malone JL, Horowitz S, Weidinger J, Williamson D, Simoes E.** Multiple sclerosis prevalence and possible lead exposure. *J Neurol Sci*, **2008**; 269(1-2): 158-162.
79. **Forbes RB, Swingler RJ.** Estimating the prevalence of multiple sclerosis in the United Kingdom by using capture-recapture methodology. *Am J Epidemiol*, **1999**; 149(11): 1016-1024.
80. **Richardson AK, Clarke G, Sabel CE, Pearson JF, Mason DF, Taylor BV.** Method for identifying eligible individuals for a prevalence survey in the absence of a disease register or population register. *Intern Med*, **2012**; 42(11): 1207-1212.
81. **Taylor BV, Pearson JF, Clarke G, Mason DF, Abernethy DA, Willoughby E, Sabel C.** MS prevalence in New Zealand, an ethnically and latitudinally diverse country. *Mult Scler*, **2010**; 16(12): 1422-1431.
82. **Cristiano E, Patrucco L, Rojas JI, Cacaress F, Carra A, Garcea O, Gold L, Tessler J, Kremenchutzky M.** Prevalence of multiple sclerosis in Buenos Aires, Argentina using the capture-recapture method. *Eur J Neurol*, **2009**; 16(2): 183-187.
83. **Crewe J, Morgan WH, Morlet N, Clark A, Lam G, Parsons R, Mukhtar A, Ng J, Crowley M, Semmens J.** Prevalence of blindness in Western Australia: a population study using capture and recapture techniques. *Br J Ophthalmol*, **2012**; 96(4): 478-481.
84. **Gill GV, Ismail AA, Beeching NJ, Macfarlane SB, Bellis MA.** Hidden diabetes in the UK: use of capture-recapture methods to estimate total prevalence of diabetes mellitus in an urban population. *J R Soc Med*, **2003**; 96(7): 328-332.
85. **Kraus L, Augustin R, Frischer M, Kümmler P, Uhl A, Wiessing L.** Estimating prevalence of problem drug use at national level in countries of the European Union and Norway. *Addiction*, **2003**; 98(4):471-485.
86. **Kruglov YV, Kobyshecha YV, Salyuk T, Varetska O, Shakarishvili A, Saldanha VP.** The most severe HIV epidemic in Europe: Ukraine's national HIV prevalence estimates for 2007. *Sex Transm Infect*, **2008**; 84 Suppl 1:i37-i41.
87. **Platt L, Hickman M, Rhodes T, Mikhailova L, Karavashkin V, Vlasov A, Tilling K, Hope V, Khutorksoy M, Renton A.** The prevalence of injecting drug use in a Russian city: implications for harm reduction and coverage. *Addiction*, **2004** ; 99(11):1430-1438.
88. **Christensen PB, Hay G, Jepsen P, Omland LH, Just SA, Krarup HB, Weis N, Obel N, Cowan S.** Hepatitis C prevalence in Denmark -an estimate based on multiple national registers. *BMC Infect Dis*, **2012**; 12:178.
89. **Lykke Thomsen L, Kirchmann Eriksen M, Faerch Romer S, Andersen I, Ostergaard E, Keiding N, Olesen J, Russell MB.** An epidemiological survey of hemiplegic migraine. *Cephalalgia*. **2002**; 22(5):361-375.

90. **Smeeton NC, Rona RJ, Sharland G, Botting BJ, Barnett A, Dundas R.** Estimating the prevalence of malformation of the heart in the first year of life using capture-recapture methods. *Am J Epidemiol*, **1999**; 150(7):778-785.
91. **Debrock C, Preux PM, Houinato D, Druet-Cabanac M, Kassa F, Adjien C, Avode G, Denis F, Boutros-Toni F, Dumas M.** Estimation of the prevalence of epilepsy in the Benin region of Zinvié using the capture-recapture method. *Int J Epidemiol*, 2000; 29(2):330-335.
92. **Ismail AA, Beeching NJ, Gill GV, Bellis MA.** How many data sources are needed to determine diabetes prevalence by capture-recapture? *Int J Epidemiol*, **2000**; 29(3):536-541.
93. **Gnavi R, Karaghiosoff L, Costa G, Merletti F, Bruno G.** Socio-economic differences in the prevalence of diabetes in Italy: the population-based Turin study. *Nutr Metab Cardiovasc Dis*, **2008**; 18(10):678-682.
94. **Oeltmann JE, Liese AD, Heinze HJ, Addy CL, Mayer-Davis EJ.** Prevalence of diagnosed diabetes among African-American and non-Hispanic white youth, 1999. *Diabetes Care*, **2003**; 26(9):2531-2535.
95. **Harvey JN, Craney L, Kelly D.** Estimation of the prevalence of diagnosed diabetes from primary care and secondary care source data: comparison of record linkage with capture-recapture analysis. *J Epidemiol Community Health*, **2002**; 56(1):18-23.
96. **Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM.** The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ*, **1997**; 315(7107):524-528.
97. **Ogbera AO, Fasanmade O, Ohwovoriole AE, Adediran O.** An assessment of the disease burden of foot ulcers in patients with diabetes mellitus attending a teaching hospital in Lagos, Nigeria. *Int J Low Extrem Wounds*. **2006**; 5(4):244-249.
98. **Sanderson M, Benjamin JT, Lane MJ, Cornman CB, Davis DR.** Application of capture-recapture methodology to estimate the prevalence of dementia in South Carolina. *Ann Epidemiol*, **2003**; 13(7):518-524.
99. **Porten K, Sailor K, Comte E, Njikap A, Sobry A, Sihom F, Meva'a A, Eyangoh S, Myatt M, Nackers F, Grais RF.** Prevalence of Buruli ulcer in Akonolinga health district, Cameroon: results of a cross sectional survey. *PLoS Negl Trop Dis*, **2009**; 3(6):e466.
100. **Astrauskienė A, Dobrovolskij V, Stukas R.** The prevalence of problem drug use in Lithuania. *Medicina (Kaunas)*, **2011**; 47(6):340-346.
101. **Vaissade L, Legleye S.** Capture-recapture estimates of the local prevalence of problem drug use in six French cities. *Eur J Public Health*, **2009**; 19(1):32-37.

102. **Hay G, Gannon M, MacDougall J, Eastwood C, Williams K, Millar T.** Capture--recapture and anchored prevalence estimation of injecting drug users in England: national and regional estimates. *Stat Methods Med Res*, **2009**; 18(4):323-339.
103. **Kimber J, Hickman M, Degenhardt L, Coulson T, van Beek I.** Estimating the size and dynamics of an injecting drug user population and implications for health service coverage: comparison of indirect prevalence estimation methods. *Addiction*, **2008**; 103(10):1604-1613.
104. **Hope VD, Hickman M, Tilling K.** Capturing crack cocaine use: estimating the prevalence of crack cocaine use in London using capture-recapture with covariates. *Addiction*, **2005**; 100(11):1701-1708.
105. **Hickman M, Higgins V, Hope V, Bellis M, Tilling K, Walker A, Henry J.** Injecting drug use in Brighton, Liverpool, and London: best estimates of prevalence and coverage of public health indicators. *J Epidemiol Community Health*, **2004**; 58(9):766-771.
106. **Chiang SC, Chen CY, Chang YY, Sun HJ, Chen WJ.** Prevalence of heroin and methamphetamine male users in the northern Taiwan, 1999-2002: capture-recapture estimates. *BMC Public Health*, **2007**; 7:292.
107. **Brady JE, Friedman SR, Cooper HL, Flom PL, Tempalski B, Gostnell K.** Estimating the prevalence of injection drug users in the U.S. and in large U.S. metropolitan areas from 1992 to 2002. *J Urban Health*, **2008**; 85(3):323-351.
108. **Hickman M, Cox S, Harvey J, Howes S, Farrell M, Frischer M, Stimson G, Taylor C, Tilling K.** Estimating the prevalence of problem drug use in inner London: a discussion of three capture-recapture studies. *Addiction*, **1999**; 94(11):1653-1662.
109. **Hay G, McKeganey N.** Estimating the prevalence of drug misuse in Dundee, Scotland: an application of capture-recapture methods. *J Epidemiol Community Health*, **1996**; 50(4):469-472.
110. **Holland R, Vivancos R, Maskrey V, Sadler J, Rumball D, Harvey I, Swift L.** The prevalence of problem drug misuse in a rural county of England. *J Public Health (Oxf)*, **2006**; 28(2):88-95
111. **Brugal MT, Domingo-Salvany A, Díaz de Quijano E, Torralba L.** Prevalence of problematic cocaine consumption in a city of southern Europe, using capture-recapture with a single list. *J Urban Health*, **2004**; 81(3):416-427.
112. **Comiskey CM, Barry JM.** A capture-recapture study of the prevalence and implications of opiate use in Dublin. *Eur J Public Health*, **2001**; 11(2):198-200.
113. **Brugal MT, Domingo-Salvany A, Maguire A, Caylà JA, Villalbí JR, Hartnoll R.** A small area analysis estimating the prevalence of addiction to opioids in Barcelona, 1993. *J Epidemiol Community Health*, **1999**; 53(8):488-494.

114. **Domingo-Salvany A, Hartnoll RL, Maguire A, Brugal MT, Albertín P, Caylà JA, Casabona J, Suelves JM.** Analytical considerations in the use of capture-recapture to estimate prevalence: case studies of the estimation of opiate use in the metropolitan area of Barcelona, Spain. *Am J Epidemiol*, **1998**; 148(8):732-740.
115. **Harrison MJ, O'Hare AE, Campbell H, Adamson A, McNeillage J.** Prevalence of autistic spectrum disorders in Lothian, Scotland: an estimate using the "capture-recapture" technique. *Arch Dis Child*, **2006**; 91(1):16-19.
116. **Al-Shahi R, Fang JS, Lewis SC, Warlow CP.** Prevalence of adults with brain arteriovenous malformations: a community based study in Scotland using capture-recapture analysis. *J Neurol Neurosurg Psychiatry*, **2002**; 73(5):547-551.
117. **Corrao G, Bagnardi V, Vittadini G, Favilli S.** Capture-recapture methods to size alcohol related problems in a population. *J Epidemiol Community Health*, **2000**; 54(8):603-610.
118. **Oosterlee A, Vink RM, Smit F.** Prevalence of family violence in adults and children: estimates using the capture-recapture method. *Eur J Public Health*, **2009**; 19(6):586-591
119. **Caleo GM, Sy AP, Balandine S, Polonsky J, Palma PP, Grais RF, Checchi F.** Sentinel site community surveillance of mortality and nutritional status in southwestern Central African Republic, 2010. *Popul Health Metr*, **2012**; 10(1):18
120. **Ikiz F.** Use of Capture-Recapture Techniques in Biological and Medical Sciences. *MedicRes World Congress on "Good Medical Research" 2012.*
121. **Peterson JT.** SAS macros for simulation. Erişim adresi: http://coopunit.forestry.uga.edu/open_source/SAS_simulation/multi_simul.pdf. Erişim tarihi: 06.12.2012

EK

Yakala Tekrar Yakala Dizayını İçin Simülasyon Makrosu (det_p= Yakalama oranı, n= Örnek genişliği, occas= Yakalama sayısı, simrun= Simülasyon sayısı) (121)

```
%macro capture(det_p=, PopN=, occas=, simrun=);
options nomprint;
%do sr=1 %to &simrun;
data temp; length caphist $&occas;
%do ind=1 %to &popN;
sum=0;
%do k=1 %to &occas;
%if %sysfunc(RANBIN(0, 1, &det_p)) eq 1 %then %do;
sum= sum + 1;
substr(caphist,&k,1)='1';
%end;
%else %do;
substr(caphist,&k,1)='0';
%end;
%end;
output;
%end;

proc freq data=temp noprint;
tables caphist /out=caphist; where sum ne 0;
run;
data caphist; set caphist;
put caphist count " ";
run;
%end;
%mend;
```

ÖZGEÇMİŞ

1987 yılında Mersin il merkezinde doğdu. İlk ve orta öğretimini Mersin’de tamamladı. 2005 yılında Mersin Ticaret ve Sanayi Odası Anadolu Lisesi’nden mezun oldu. Aynı yıl Muğla Üniversitesi İstatistik Bölümü’nde lisans eğitimine başladı. 2009 yılında lisans derecesini alarak Mersin Üniversitesi Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı’nda lisansüstü eğitime başladı. Eylül 2011’de aynı anabilim dalında araştırma görevlisi olarak çalışmaya başladı ve halen devam etmektedir. Yabancı dili İngilizce’dir.