



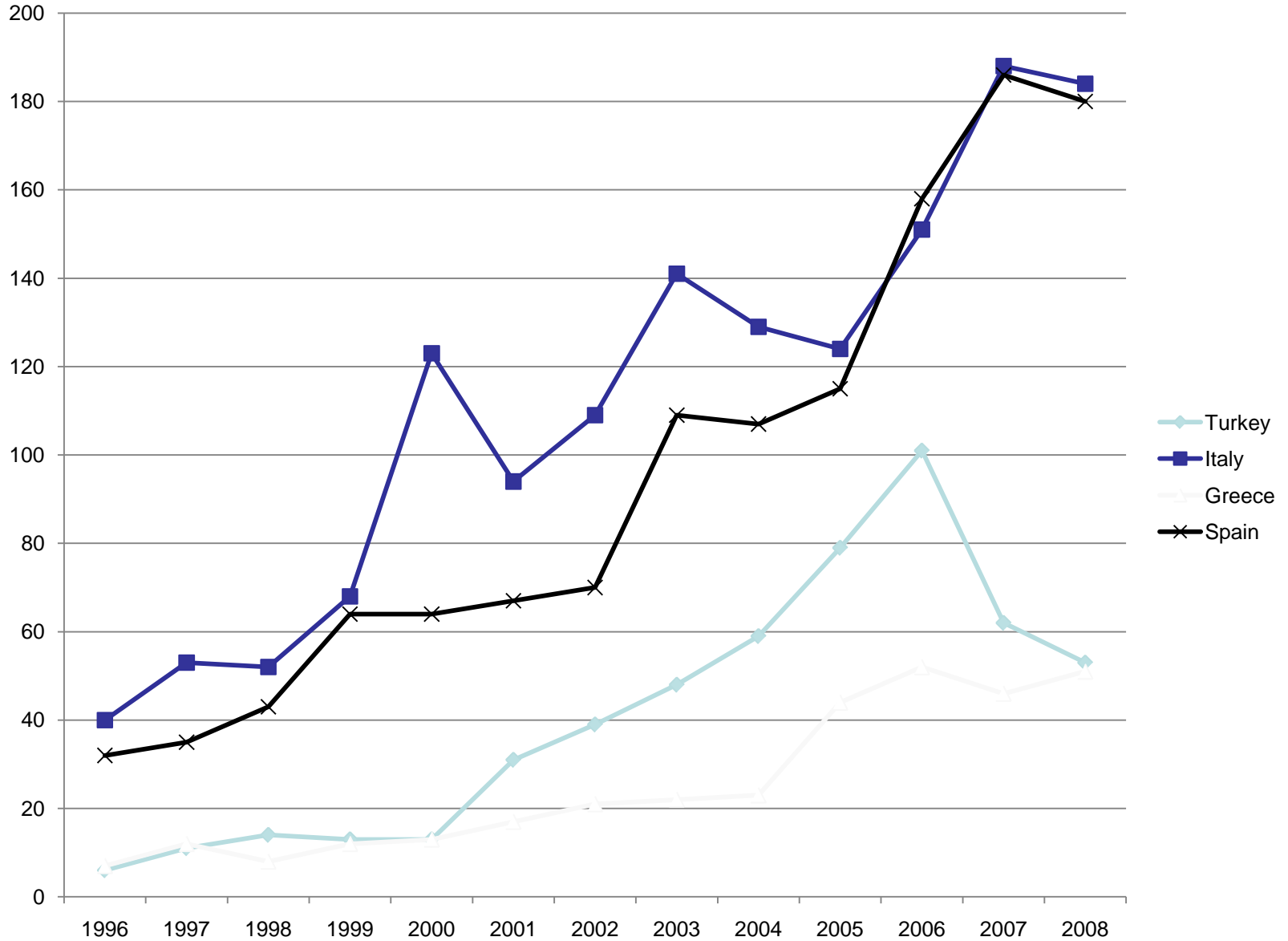
Arařtırma Tasarımları ve İstatistięe Giriř

Prof.Dr.Önder Ergönül

5 Temmuz 2011

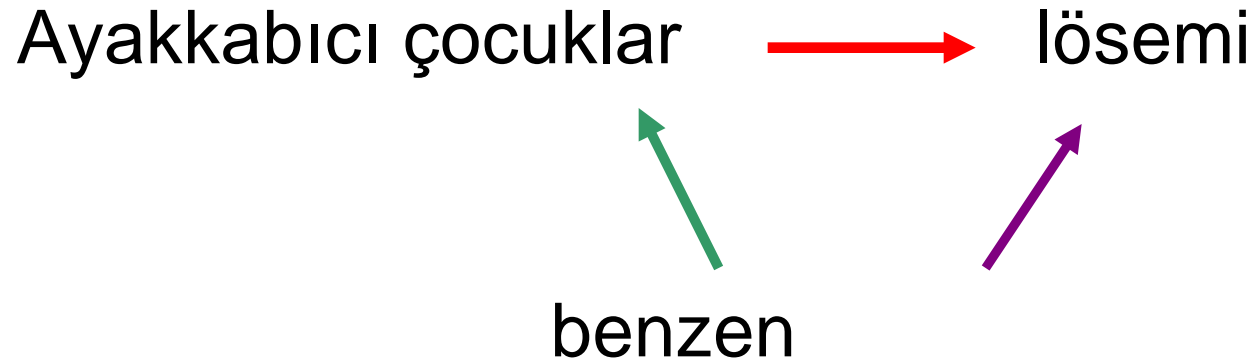
Koç Üniversitesi Tıp Fakóltesi

Enfeksiyon Hastalıklarında Yayınlar ve Türkiye' nin Durumu



Neden alıřma yapılır?

Bir sorunun özümü için gerekli bilgi toplamak ve bu sayede sorunu özmek ...



Muzaffer Aksoy, Blood, 1974

Orhan Pamuk neden yazıyor?

(Nobel konuşmasından 7 Kasım 2006)

1. İçimden geldiği için
2. Başkaları gibi normal bir iş yapamadığım için
3. Benim yazdığım gibi kitaplar yazılsın da okuyayım diye
4. Hepinize, herkese çok çok kızdığım için
5. Bir odada bütün gün oturup yazmak çok hoşuma gittiği için
6. Onu ancak değiştirerek gerçekliğe katlanabildiğim için
7. Ben, ötekiler, hepimiz, bizler İstanbul'da, Türkiye'de nasıl bir hayat yaşadık, yaşıyoruz, bütün dünya bilsin diye
8. Kâğıdın, kalemin, mürekkebin kokusunu sevdiğim için
9. Edebiyata, roman sanatına her şeyden çok inandığım için
10. Bir alışkanlık ve tutku olduğu için
11. Unutulmaktan korktuğum için
12. Getirdiği ün ve ilgiden hoşlandığım için
13. Yalnız kalmak için
14. Hepinize, herkese neden o kadar çok çok kızdığımı belki anlarım diye
15. Okunmaktan hoşlandığım için
16. Bir kere başladığım şu romanı, bu yazıyı, şu sayfayı artık bitireyim diye
17. Herkes benden bunu bekliyor diye
18. Kütüphanelerin ölümsüzlüğüne ve kitaplarımın raflarda duruşuna çocukça inandığım için
19. Hayat, dünya, her şey inanılmayacak kadar güzel ve şaşırtıcı olduğu için
20. Hayatın bütün bu güzelliğini ve zenginliğini kelimelere geçirmek zevkli olduğu için
21. Hikâye anlatmak için değil, hikâye kurmak için
22. Hep gidilecek bir yer varmış ve oraya tıpkı bir rüyadaki gibi bir türlü gidemiyormuşum duygusundan kurtulmak için
23. Bir türlü mutlu olamadığım için
24. Mutlu olmak için

Hekimlerin yazma nedenleri?

- Entellektüel stimulasyon
- Düşüncelerimizi paylaşmak
- Disiplin oluşturmak
- Araştırmalarımızı bildirmek
- Fikir ileri sürmek
- Tartışma başlatmak
- Bir başlığa sahiplenmek
- Meslekte ilerlemek
- Olgu bildirmek
- Kişisel ün sağlamak
- Unutulmamak (ölümsüzlük)
- Para kazanmak

Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.



Hekimlerin yazmama nedenleri?

- Zamansızlık
- Konusuzluk
- Çalışma arkadaşı yokluğu
- Sekreter yokluğu
- Yöntem bilmemek
- Mentor (danışman) yokluğu
- Motivasyon eksikliği
- Kendine güvensizlik
- Başlayamamak
- “Yazmaktan nefret etmek”



Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.

Çalışmalar neden yayınlanır?

- Sorunun “önemi”
- Yazarların ünü
- Denek sayısının çokluğu
- Beklenmedik girişim ve sonuçlar
- Günlük yayınların ilgisini çekecek mi?
- Sonuçlarda belirgin farklılıklar olması

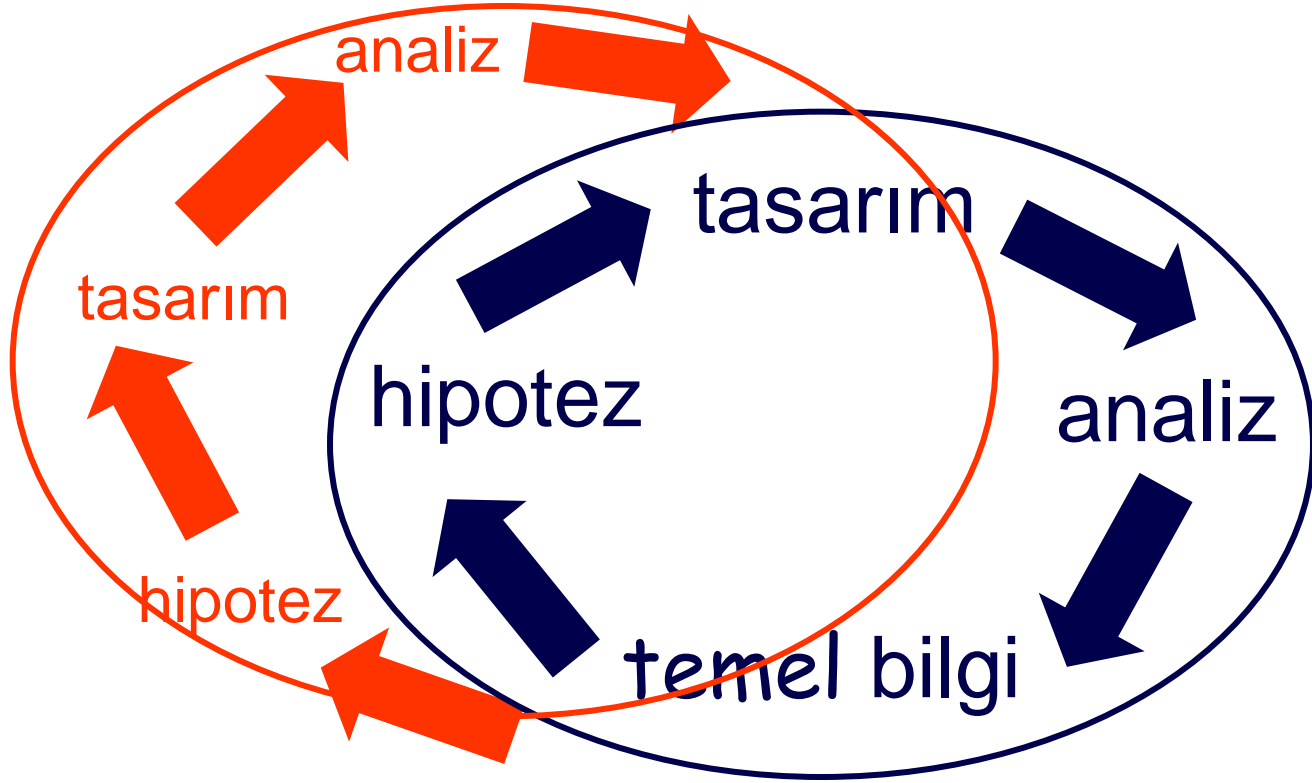


Çalışmalar neden yayına kabul edilmez?

- Önemli** bir bilimsel soruya yanıt vermemektedir
- Orijinal değildir
- Yazarın hipotezini araştırmamaktadır
- Tasarımı yanlıştır
- Örneklem sayısı azdır
- Kontrol kolu yoktur
- İstatistik analiz uygun değildir ya da yanlıştır
- Sonuçlar yanlış yorumlanmıştır
- Kötü yazılmıştır



Bilimsel üretim sarmalında eksiklerimiz nerede?



YAZIM

Bilim her konuda ilk sözü söyler ama hiçbir konuda son sözü söylemez
Victor Hugo

Araştırma Sonuçları

1. Hata / Yanlılık (*Bias*)

2. Karıştırıcı etken
(*Confounder*)

3. Şans faktörü

Tasarım

Analiz



Epidemiyoloji nedir?

“Hastalıkların belirleyicilerini inceleyen bilim dalıdır”

“İnsanları gruplamaktır”

“Ağaçlara bakarken ormanı görebilmektir”

“Neden-sonuç ilişkisi kurmaktır”

Epidemiyolojinin Sorunları

1. Tanımlayıcı

a) Ne?

b) Nerede?

c) Ne zaman?

d) Kim?

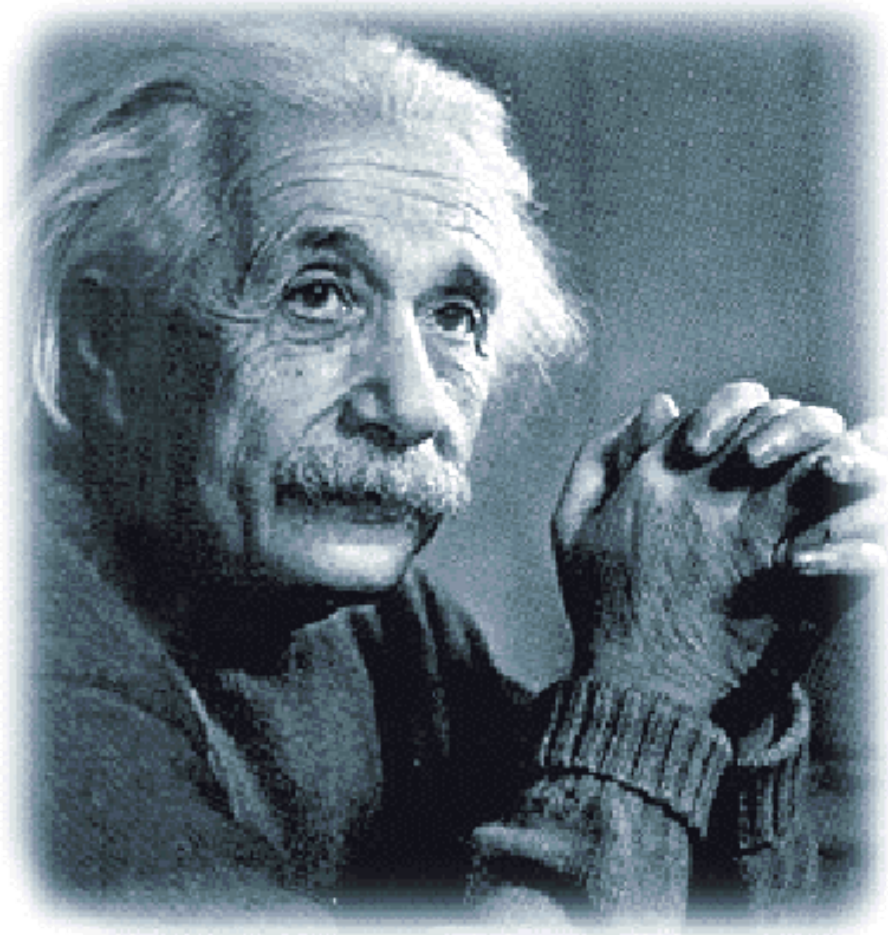
5N1K

2. Analitik

a) Neden?

b) Nasıl

3. Koruyucu önlemler



**“Basite indirgeyin,
kolayına kaçmayın...”**

Basit \neq Kolay

Klinik Arařtırmaların Sınıflandırılması

Kanıtın Kalitesi

- I. En az bir uygun tasarlanmış randomize klinik alıřma
Meta-analiz?
- II-1. Randomizasyon yapılmamıř iyi tasarlanmış kontrollü alıřma
- II-2. İyi tasarlanmış, tercihen birden ok merkezde yapılmıř kohort veya vaka-kontrol alıřma
- II-3. Giriřimli ya da giriřimsiz zaman serileri
- III. Saygın otoritelerin klinik deneyime dayanan grüşleri, tanımlayıcı alıřmalar, uzman komisyonların raporları

RKÇ veya
Meta-analiz

Randomizasyonsuz
kontrollü

Çok merkezli kohort veya
olgu-kontrol

Uzman görüşü, tanımlayıcı
çalışmalar, uzman raporları

Çalışma Tasarımları



Temel Tanımlar

İnsidans: Bir yıl içinde yeni hastaların nüfusa oranı
Örnek; Tifo, 100 000' de 35 (yılda),
35/100 000
Genellikle 100 000' de olarak ifade edilir

Kümülatif insidans:

Belirli bir zaman diliminde insidans

Zaman diliminin başlangıcındaki nüfus

Prevalans: Hasta olanların nüfusa oranı

Temel Tanımlar

Olgu fatalite (letalite) hızı:

Belirli bir hastalıktan ölenler

Hasta olanlar

Mortalite hızı:

Belirli bir hastalıktan ölenler

Nüfus

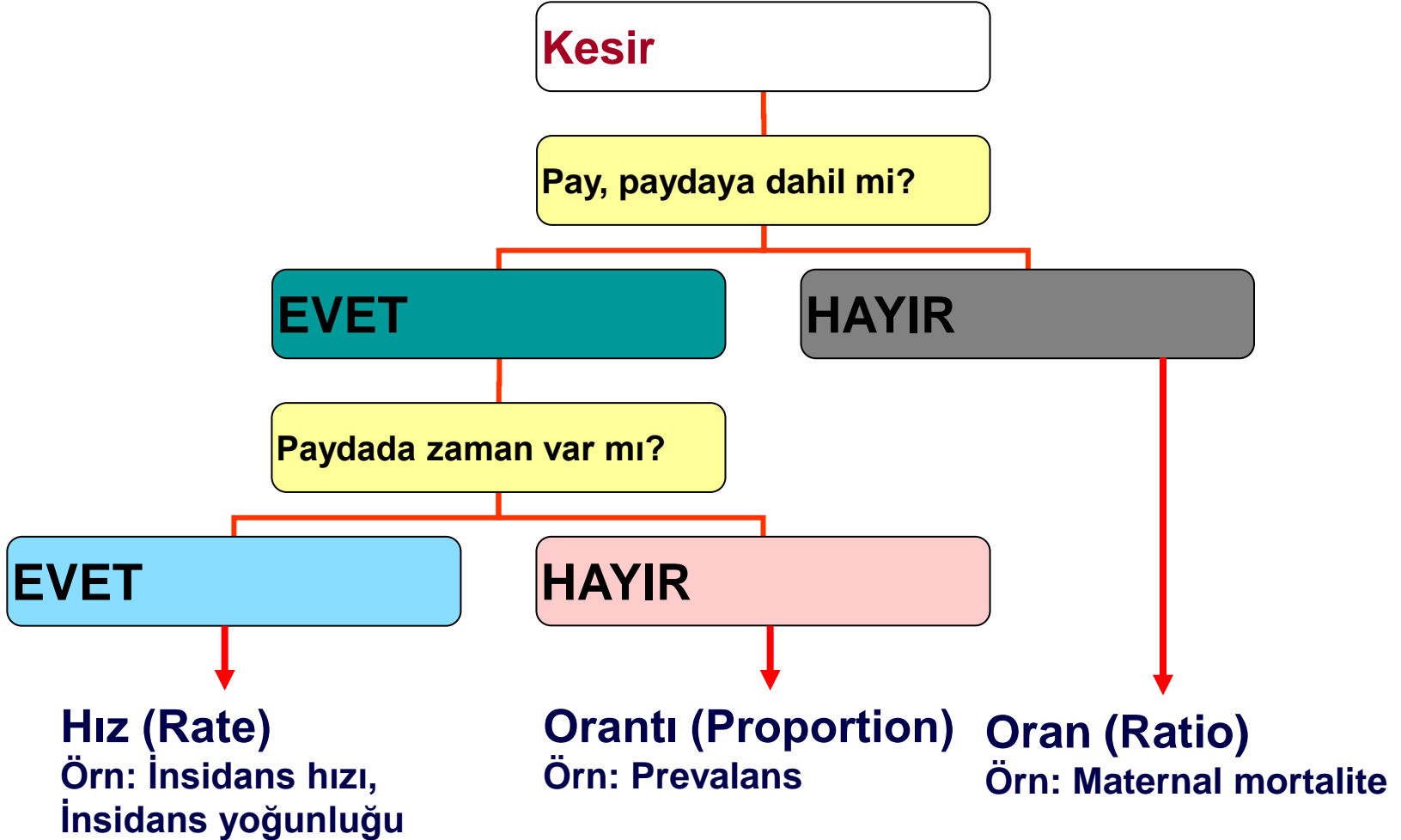
Örnek: Kuduzun **olgu fatalite hızı** yüksek, **mortalitesi** düşüktür.

Atak (hastalığa yakalanma) hızı:

Hasta olanlar

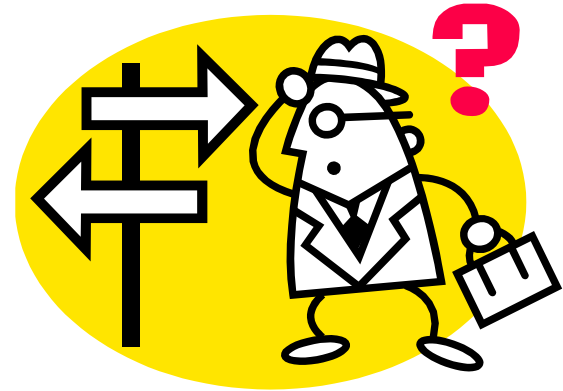
Risk altındaki nüfus

Temel Tanımlar



Çalıřma Evreleri

1. Ön hazırlık, hipotez kurma
2. Çalıřma tasarımı
3. Veri toplama
4. Verilerin deęerlendirilmesi ve analiz
5. Yazım



Occam'ın Usturası

Occam: 14 yy, İngiliz felsefeci ve din bilimci

“Birşeyleri açıklamak için ortaya atılan varsayımlar, ihtiyacın ötesinde çoğaltılmamalıdır.”

Taylor RB. The Clinician's Guide to Medical Writing Springer, 2005.

“Rakip varsayımlar arasında en doğru varsayım, en basit olanıdır.”

*Wynn CM, Wiggins AW. Yanlış Yönde Kuantum Sıçramalar.
TÜBİTAK, 2005.*

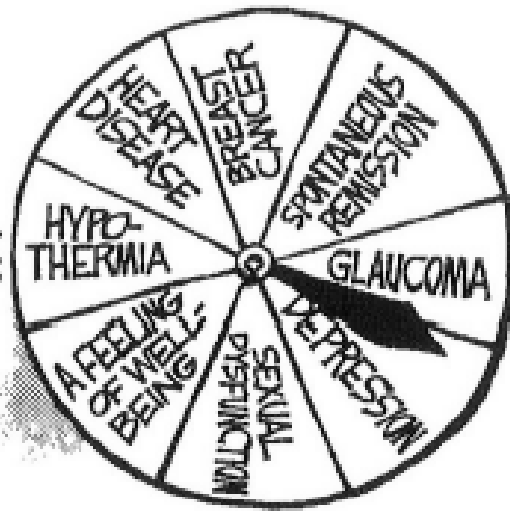
Today's Random Medical News

from the New England
Journal of
Panic-Inducing
Gobbledygook

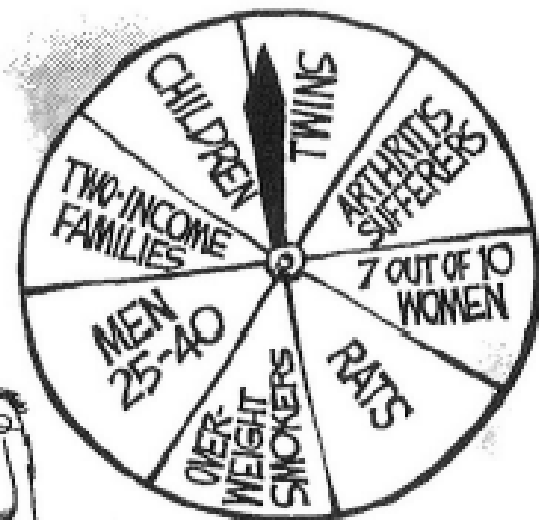
JIM BERGMAN © 2000 HART PUBLICATIONS



CAN CAUSE



IN



Çay ve Kanser

Cilt kanseri olan 1400 olgu vs 700 sağlıklı kontrol



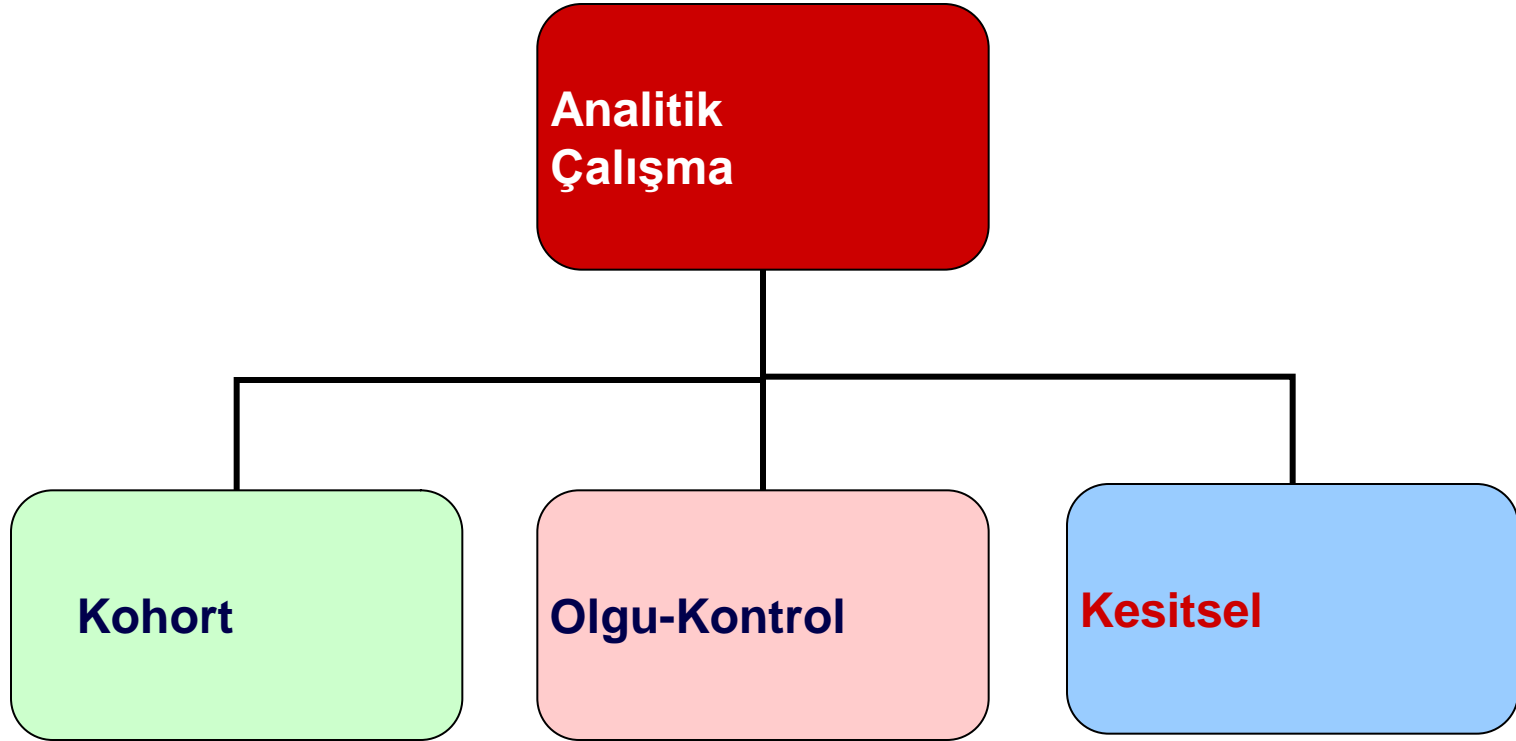
Günde iki fincan yeşil veya siyah çay içen deneklerin cilt kanserine yakalanma riskinin % 65 azalabildiği belirlendi.

Çaya limon katmak koruyuculuğu artırıyor.

(Eur J Cancer Prevention)

Araştırmacılar, çayın içinde bulunan "polifenol" maddesinin antioksidan özelliği sayesinde kanserden korunmaya yardımcı olduğunu belirttiler.

Girişimsel Olmayan/Gözlemsel Araştırma Tasarımları



Çalışma Tasarımları

Maruziyet
(exposure)

KOHORT

Sonuç
(outcome)

Maruziyet
(exposure)

Vaka-Kontrol

Sonuç
(outcome)

Kesitsel çalışma

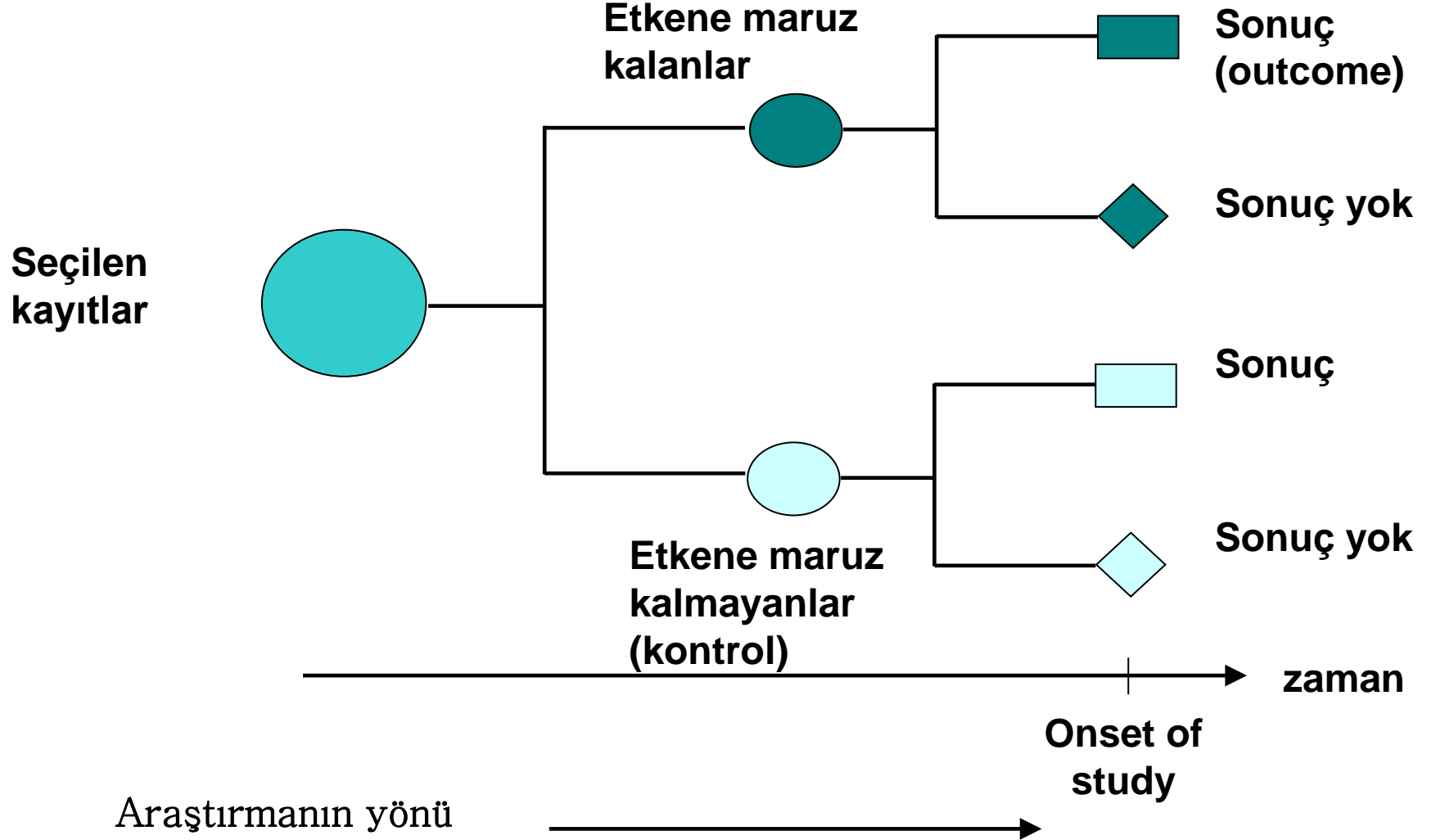
Maruziyet



Sonuç



Retrospektif Kohort



En İyİ Bilinen Kohort alıřmalar

- Framingham Kalp alıřması
- Hemřirelerin Saęlıęı alıřması
- Hekimlerin Saęlıęı alıřması

Kohort Çalışmalar: Framingham Heart Study



The screenshot shows a web browser window displaying the National Heart Lung and Blood Institute website. The browser's address bar shows the URL www.framinghamheartstudy.org. The website header includes the NHLBI logo, the text "Department of Health and Human Services • National Institutes of Health", and "National Heart Lung and Blood Institute People Science Health". Navigation links for "HOME", "SITE INDEX", and "CONTACT US" are visible. A search bar is present with "TIPS" and "ADVANCED SEARCH" options. The main content area features the Framingham Heart Study logo, which is a circular emblem with three overlapping hearts and the text "FRAMINGHAM HEART STUDY", "Three Generations of Dedication", and the years "1948", "1971", and "2002". Below the logo, a list of links is provided:

- [Design, Rationale, and Objectives](#)
- [List of Current Investigators](#)
- [Research Milestones](#)
- [Genetic Materials and Data Distribution](#)
- [Policy for Ancillary Studies](#)
- [CHD Risk Prediction Score Sheets](#)
- [Stroke Risk Predictor](#)
- [Bibliography of Published Papers](#)
- [Study Component Schedules](#)
- [Data Collection Forms](#)

The browser's status bar at the bottom indicates "Internet".

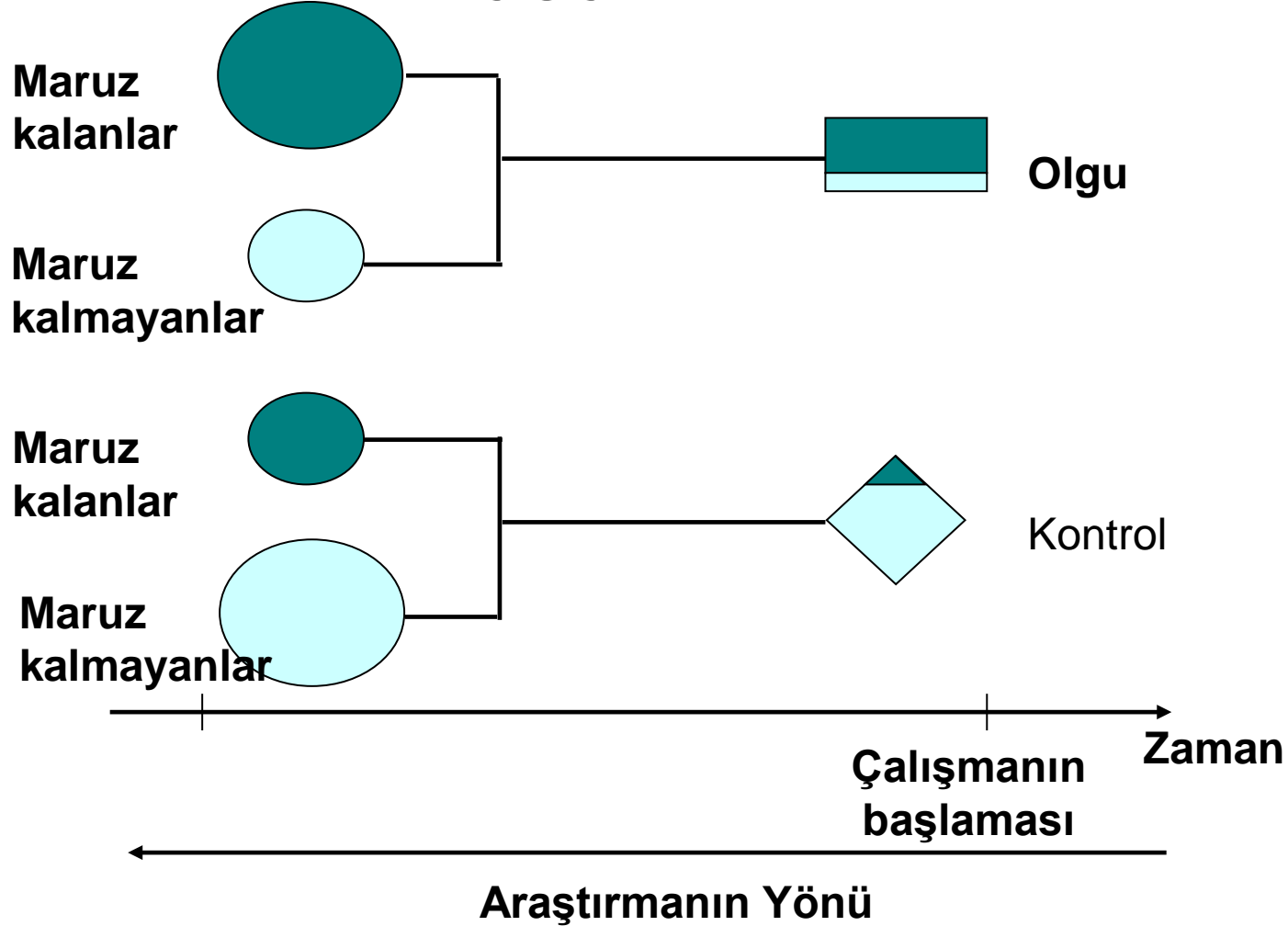
www.framinghamheartstudy.org bakılacak

Framingham Heart Study

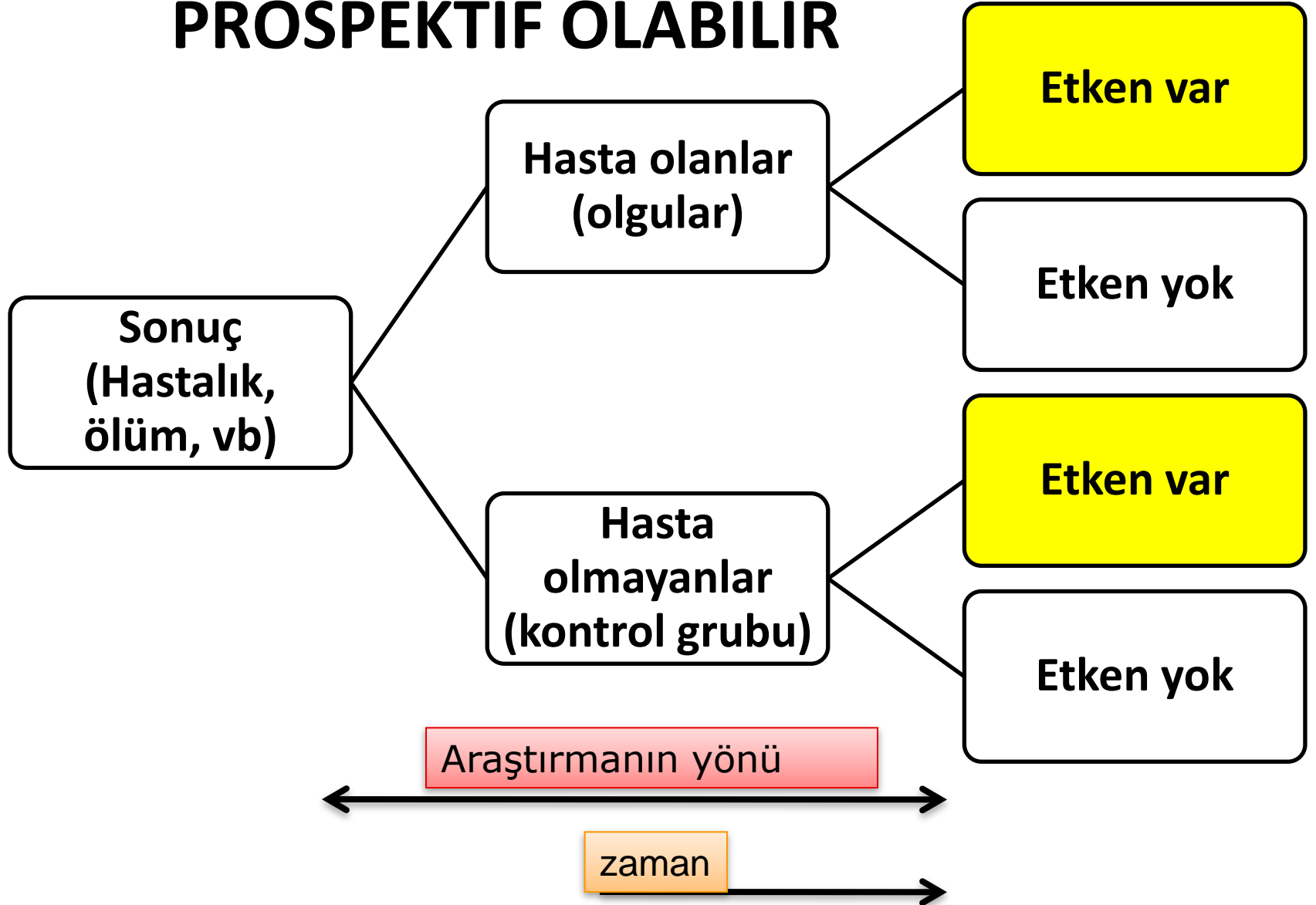
50 Years of Research Success

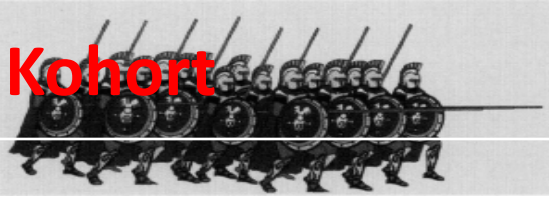
Olgu Kontrol Çalışma

Tasarımı



OLGU KONTROL ÇALIŞMALAR PROSPEKTİF OLABİLİR





Kohort

Vaka-Kontrol

1. Başlangıç noktası

Neden
(maruz kalma)

Sonuç
(vaka)

2. Çalışma süresi

Uzun

Kısa

3. Maliyet

Pahalı

Ucuz

4. Çalışmanın yönü

Etkenlerden
sonuca ya da
olgulara

Sonuçtan ya da
olgulardan
etkenlere

5. Gerçekleştirme

Zor

Kolay

Sıradışı bir durum!

Jaffe HWf, et al. National Case-Control Study of Kaposi 's Sarcoma and Pneumocystis carinii Pneumonia in Homosexual Men. Ann Int Med 1983; 99: 145-151.

1981 Haziran/Temmuz

Kaliforniya ve New York' da,

genç, sağlıklı homoseksüellerde

pneumocystis carinii pneumonia (PCP) ve

Kaposi sarcoma(KS) sıklığı artmaktadır.

Sıradışı bir durum!

Amaç: Genç homoseksüellerde Kaposi sarkomu ve Pneumocystis carinii pnömonisi görülmesini belirleyen risk faktörlerinin araştırılması

Yöntem: Olgu-Kontrol / 50 olgu ve 150 kontrol

Sonuç:

- Saunalara gidenler
- Sifiliz öyküsü olanlar
- İlaç kullananlar
- İn hale nitrit kullananlar
- İtalyanlar (HLA-DR5 antijeni ile ilişkili olarak)



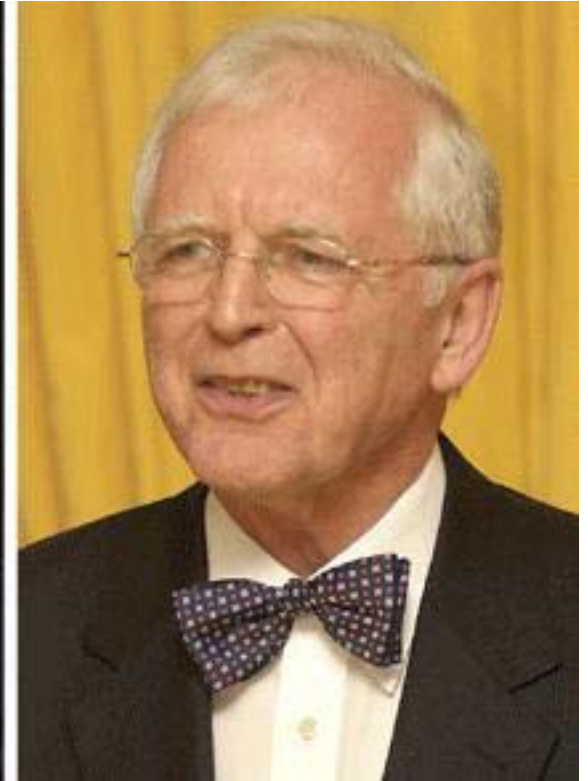
2008 Nobel Tıp Ödülü Sahipleri: HIV



Francoise Barre-Sinoussi



Luc Montagnier



Harald zur Hausen

Olgu Kontrol Çalışmalar

THE NEW ENGLAND JOURNAL of MEDICINE

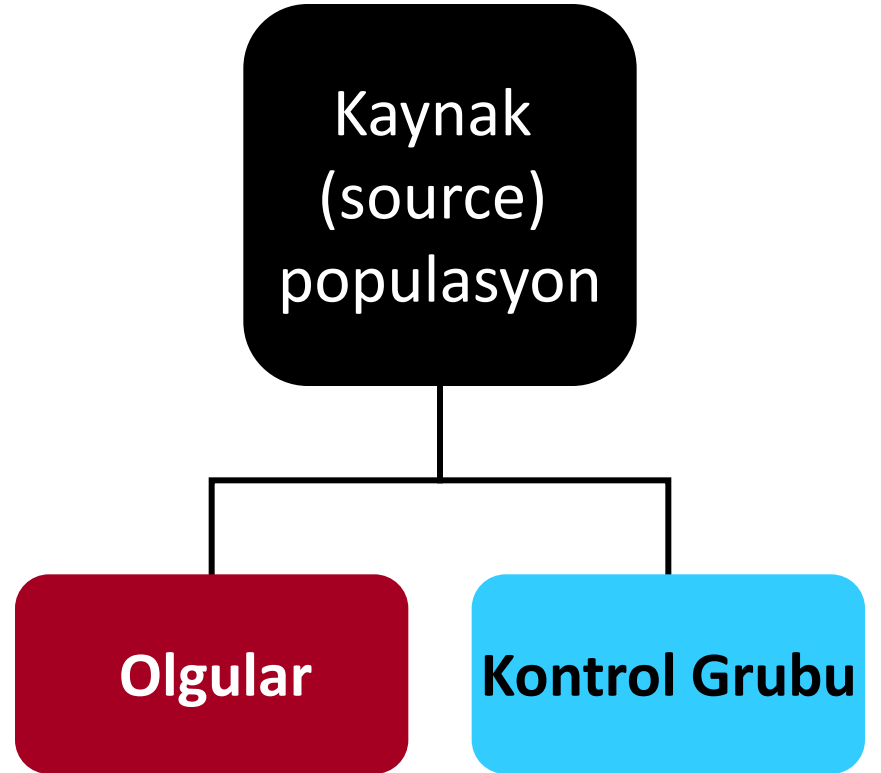
ORIGINAL ARTICLE

An Outbreak of Hepatitis A Associated with Green Onions

Variable	Patients (N=240)	Controls (N=134)	Odds Ratio (95% CI)
	<i>no. (%)</i>		
Age >12 yr	212 (88)	97 (72)	2.9 (1.7–5.0)
Menu items consumed			
Mild salsa†	218 (91)	45 (35)	19.6 (11.0–34.9)
Grilled chicken fajitas‡	29 (12)	8 (6)	2.2 (1.0–4.9)
Bar beverage§	49 (21)	16 (12)	1.9 (1.0–3.5)
Ground beef burrito	7 (3)	0	Indeterminate
Enchilada platter	7 (3)	0	Indeterminate
Ingredients consumed¶			
Green onions	234 (98)	76 (58)	33.3 (12.8–86.2)
Diced white onions**	237 (99)	95 (72)	30.8 (9.3–102.2)
Beans	113 (47)	35 (26)	2.5 (1.6–4.0)
Mixed cheese	181 (75)	76 (57)	2.3 (1.5–3.7)
Rice	124 (52)	46 (34)	2.0 (1.3–3.2)
Sour cream	92 (38)	33 (25)	1.9 (1.2–3.0)
Corn-cake mix	124 (52)	49 (37)	1.9 (1.2–2.9)

Çalışma Grubu

- Olgu kontrol çalışmalarında en büyük sorun, kontrol grubunun oluşturulmasıdır.
- Kontrol grubu, olgular ile aynı kaynak popülasyondan olmalıdır.



Olgu – Kontrol Çalışması

Table 2. Consumption of Selected Food Items by 240 Patients with Hepatitis A and 134 Controls Who Ate One Meal in Restaurant A between October 3 and October 6, 2003.*

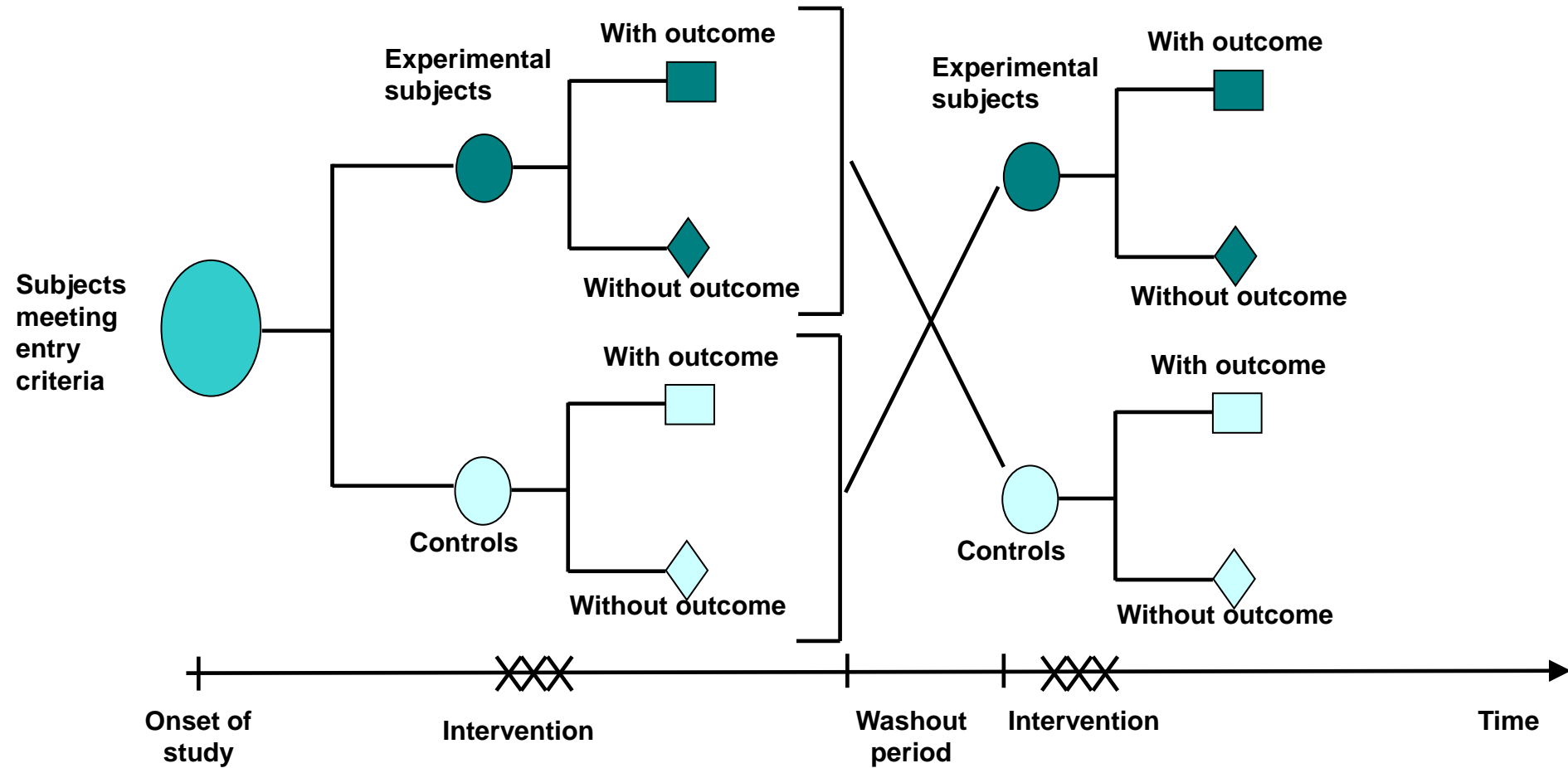
Variable	Patients (N=240)	Controls (N=134)	Odds Ratio (95% CI)
	<i>no. (%)</i>		
Age >12 yr	212 (88)	97 (72)	2.9 (1.7–5.0)
Menu items consumed			
Mild salsa†	218 (91)	45 (35)	19.6 (11.0–34.9)
Grilled chicken fajitas‡	29 (12)	8 (6)	2.2 (1.0–4.9)
Bar beverage§	49 (21)	16 (12)	1.9 (1.0–3.5)
Ground beef burrito	7 (3)	0	Indeterminate
Enchilada platter	7 (3)	0	Indeterminate
Ingredients consumed¶			
Green onions	234 (98)	76 (58)	
Diced white onions**	237 (99)	95 (72)	
Beans	113 (47)	35 (26)	
Mixed cheese	181 (75)	76 (57)	
Rice	124 (52)	46 (34)	
Sour cream	92 (38)	33 (25)	1.9 (1.2–3.0)
Corn-cake mix	124 (52)	49 (37)	1.9 (1.2–2.9)

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

An Outbreak of Hepatitis A
Associated with Green Onions

Cross-over study design



Potansiyel Yanlılık (bias, hata) Tipleri

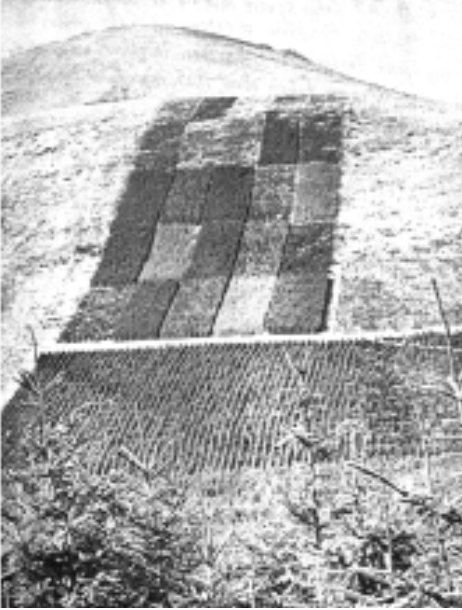
- Örnek seçerken (*selection bias*)
 - Kendi kendine seçim
 - Tanısal
- Bilgi toplarken
 - Yanlış gruplama (*misclassification bias*)
 - Hatırlayamama (*recall bias*)

Randomizasyon Nedir?

- Yanlıř gruptama (misklasifiasyon) engellemek için yapılır
- Karıřtırıcı deęiřkenlerle bařetmek için uygulanır

Randomizasyon

- İlk randomizasyon: Fisher, 1923
 - Amaç: Patates üretiminde farklı gübrelerin etkilerinin saptanması
 - Eski yöntem: Farklı tarlaların hasatlarının karşılaştırılması
 - Yeni yöntem:
 1. Tarlanın küçük bölümlere ayrılması
 2. Gübrelerin küçük bölümlere rasgele uygulanması

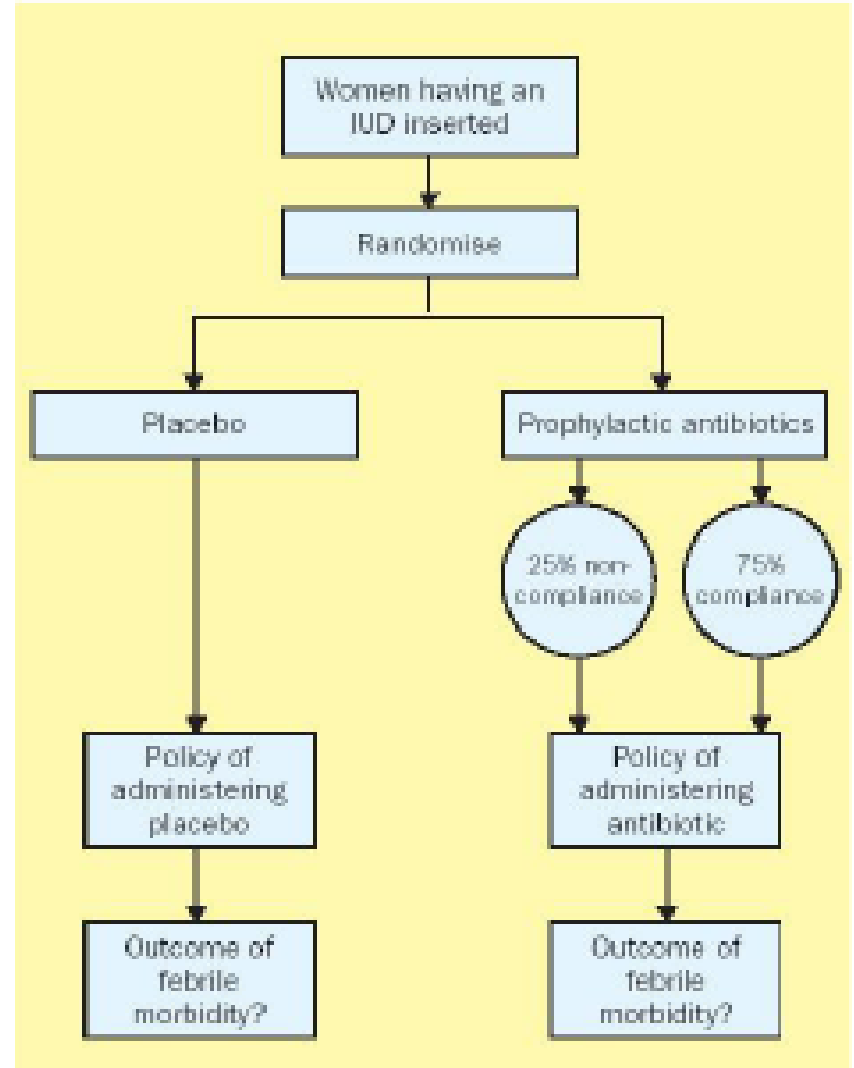


Akış Diyagramı

Tasarlararken

Analiz yaparken

Okuyucu için



En iyi bilinen kohort alıřmalar

- Framingham Kalp alıřması
- Hemřirelerin Saęlıęı alıřması
- Hekimlerin Saęlıęı alıřması



The screenshot shows the website for the Framingham Heart Study. The top navigation bar includes the National Heart Lung and Blood Institute logo, the text "Department of Health and Human Services • National Institutes of Health", and a search bar. Below the navigation bar, the text "Information for...Researchers" is displayed. The main content area features the website URL "www.framinghamheartstudy.org" and the title "Framingham Heart Study" in purple. Below the title, the text "50 Years of Research Success" is written in a purple serif font. To the right of the text is a circular logo for the Framingham Heart Study, which includes the text "FRAMINGHAM HEART STUDY", "Three Generations of Dedication", and the years "1948", "1971", and "2002". Below the logo, a list of links is provided, including "Design, Rationale, and Objectives", "List of Current Investigators", "Research Milestones", "Genetic Materials and Data Distribution", "Policy for Ancillary Studies", "CHD Risk Prediction Score Sheets", "Stroke Risk Predictor", "Bibliography of Published Papers", "Study Component Schedules", and "Data Collection Forms".

www.framinghamheartstudy.org

Framingham Heart Study

50 Years of Research Success

- [Design, Rationale, and Objectives](#)
- [List of Current Investigators](#)
- [Research Milestones](#)
- [Genetic Materials and Data Distribution](#)
- [Policy for Ancillary Studies](#)
- [CHD Risk Prediction Score Sheets](#)
- [Stroke Risk Predictor](#)
- [Bibliography of Published Papers](#)
- [Study Component Schedules](#)
- [Data Collection Forms](#)

Etken-Sonuç İlişkisinin Saptanması

Dört gözlü tablo=2x2 tablosu

	Sonuç VAR	Sonuç YOK
Etken VAR	a	b
Etken YOK	c	d

	Şizofreni VAR	Şizofreni YOK
Çocukken SSF enf. VAR	a	b
YOK	c	d

Etki Tahmini: *Klinik arařtırmalar;* *kohort tipi arařtırma*

□ Relatif risk

$$a/a+b : c/c+d$$

Maruz kalan bireylerin riski

Maruz kalmayan bireylerin riski

		DVT gelişti		DVT gelişmedi	
Heparin kullanımı		a	b		
Plasebo kullanımı		c	d		

Klinik arařtırmaların kanıt sınıflaması

Kanıtın Kalitesi

- I. En az bir uygun tasarlanmış randomize klinik çalışma; meta-analiz (?)
- II-1 Randomizasyon yapılmamış iyi tasarlanmış kontrollü çalışma
- II-2 İyi tasarlanmış, tercihen birden çok merkezde yapılmış kohort veya vaka-kontrol çalışma
- II-3 Girişimli ya da girişimsiz zaman serileri
- III. Saygın otoritelerin klinik deneyime dayanan görüşleri, tanımlayıcı çalışmalar, uzman komisyonların raporları

Neden metaanaliz yaparız?

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

JULY 28, 2005

VOL. 353 NO. 4

An Evaluation of *Echinacea angustifolia* in Experimental Rhinovirus Infections

Ronald B. Turner, M.D., Rudolf Bauer, Ph.D., Karin Woelkart, Thomas C. Hulsey, D.Sc., and J. David Gangemi, Ph.D.

CONCLUSIONS

The results of this study indicate that extracts of *E. angustifolia* root, either alone or in combination, do not have clinically significant effects on infection with a rhinovirus or on the clinical illness that results from it.

Fakat yayımlanmış veriler ekinezyanın soğuk algınlığının görülme sıklığını ve süresini azaltma konusunda yararlı olduğu yönündedir.

Metaanaliz gerçekten en değerli kanıt mıdır?

Evaluation of echinacea for the prevention and treatment of the common cold: a meta-analysis

Sachin A Shah, Stephen Sander, C Michael White, Mike Rinaldi, Craig I Coleman

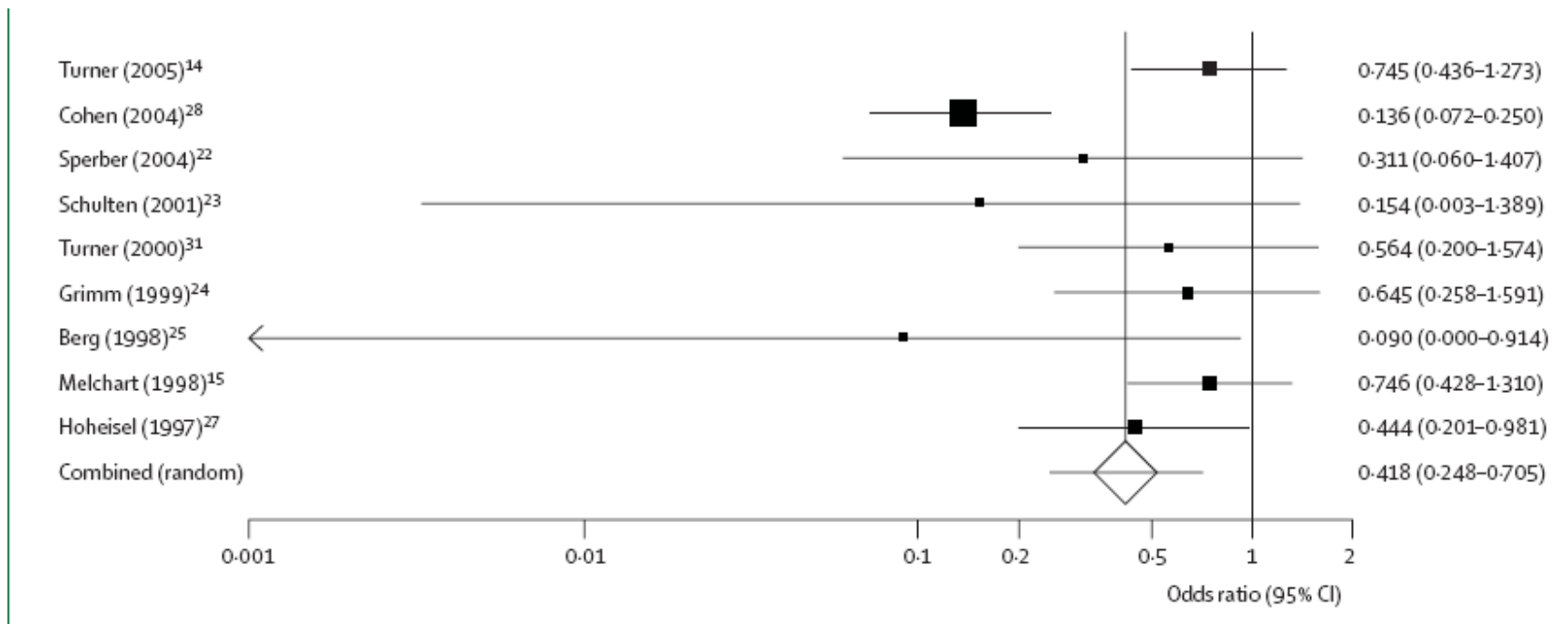
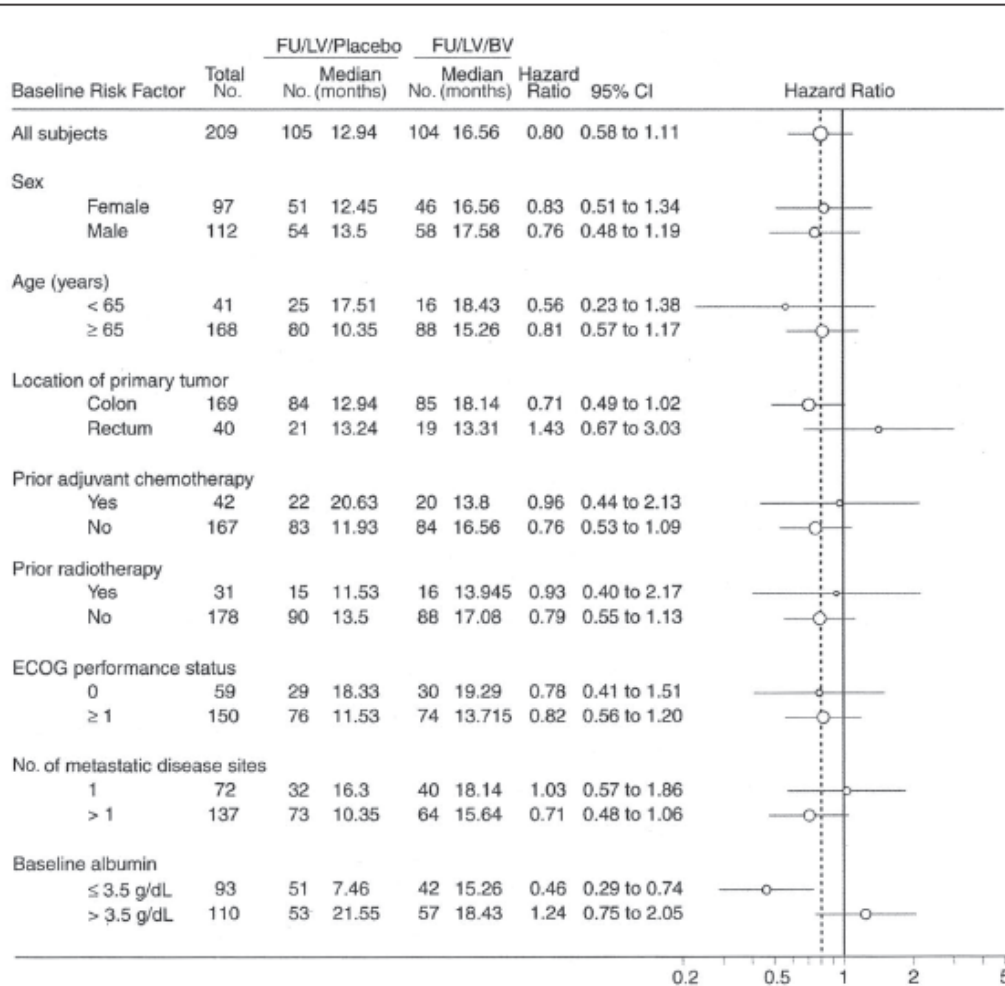


Figure 3: The effect of echinacea on incidence of common cold

Örnek



FU/LV = 5-fluorouracil/leucovorin; BV = bevacizumab; CI = confidence interval.

Median duration of survival is estimated from Káplan-Meier curves. Hazard ratio relative to

FU/LV/placebo is estimated by Cox regression. Unstratified hazard ratio is displayed. Median

baseline albumin was 3.6 g/dL.

Neden İstatistik Bilmeliyiz?

- İstatistikçilerle çalışabilmek (analiz)
- Rakamları sözle ifade edebilmek (sunuş)
- Hataları farketmek (okumak)

Rakamları Karşılaştırmayı Biliyor muyuz?

Alman kadınlar Türk erkekleri tercih ediyor
2007 yılında 3 bin 800 Alman kadın Türk erkekle evlendi...

Ülkede geçen yıl 1700 Alman kadınının İtalyanlarla ve 1200 kadının Amerikalılarla evlendiği ifade edildi.

Alman erkeklerinin ilk tercihinin ise Polonyalı kadınlar olduğu, 4 bin 200 Alman erkeğinin Polonyalı kadınlarla evlendiği bildirildi. Alman erkeklerinin ikinci tercihinin Türk kadınlarından yana olduğu, 1900 Alman erkeğinin Türk kadınlarıyla, 1700 Alman erkeğinin de Rus kadınlarıyla evlendiği kaydedildi.

Bir Bilimsel Yayının Bulgular Bölümünde 3 Ana Tablo Vardır

1. Demografik veriler
2. Tek deęişkenli analiz bulguları
3. Çok deęişkenli analiz bulguları

Sunum Planı

1. Veri Türleri
2. Merkezi Eğilim Ölçütleri
3. Merkezi Yayılım Ölçütleri
4. Örneklem Büyüklüğü

Değişken (variable)

- Bağımlı değişken, sonuç, outcome
- Bağımsız değişken: Cinsiyet, boy, ağırlık, kan basıncı, ilaç grubu, tedavi öncesi kan basıncı vb. hepsi birer değişkendir.
- Bilgisayar veritabanlarında;
 - Değişkenler, kolonda
 - Veriler, satırlarda gösterilir.

Veri Toplama

- Kağıt form
 - Her durumda gerekli, en azından şimdilik...
- Excell
- Stata
 - Veri editörü
- Access
 - Elektronik form oluşturma; birden çok sonuç alınmak istendiğinde

VERİ TOPLAMA: Tez Yazımında En Önemli Adımlardan Biridir

- Veriler kesişmemeli
- Seçenekler birbirini dışlamalı
- Her bir kolonun amacı ve anlamı olmalı
- Tüm veriler tek bir veritabanında toplanmalı



Stata Editor

Preserve Restore Sort << >> Hide Delete...

death[1] = 0

	death	incubation	prehospsym [~] s	incless8	gender	age	fever	lowext	upperext	
1	0	3	1	0	0	29	0	1	0	
2	0	2	1	0	0	62	0	0	1	
3	1	2	2	0	0	68	1	0	0	
4	1	2	3	0	1	70	0	1	0	
5	0	3	1	0	0	25	0	1	0	
6	0	2	1	0	0	48	0	1	0	
7	0	4	1	0	0	20	0	1	0	
8	0	2	2	0	0	66	0	1	0	
9	1	2	1	0	0	18	1	0	1	
10	1	2	1	0	0	68	0	1	0	
11	1	2	2	0	0	68	1	1	0	
12	1	2	2	0	0	57	1	1	0	
13	1	2	2	0	0	51	1	0	1	
14	0	2	1	0	0	22	0	1	0	
15	0	2	1	0	0	51	0	1	0	
16	1	2	1	0	1	45	1	0	0	
17	1	2	2	0	0	52	0	1	0	
18	0	2	1	0	0	38	0	0	1	
19	0	3	1	0	0	54	0	1	0	
20	0	3	1	0	0	30	1	1	0	
21	1	1	1	1	1	54	0	1	0	
22	1	1	1	1	1	46	1	1	0	



Demographics



reviewed

InitialInfectionDate:

IDconsultation

MRN:

Admit Date:

ID notes:

Account:

Discharge Date:

discharge status:

Age:

Location at time of Infectio:

disposition:

Gender:

Admitting Service:

primary diagnosis:

Comments

Readmission for same target diagnosis

Readmission within 60 days (Yes/No)

ABXhospital | micro | Mod | dischargeABX | radiology | clinical findings

datestamp	ID	account	AbxName	StartDate	EndDate	Route	Frequency	Dose	s
10/3 15:10:56	(AutoNumber)	0							0

Record: 1 of 1

Record: 1 of 1

Çalıřma Birimi Nedir?

- Hasta gn
- İlaç gn

Her bir satır neyi gsterir?

Veri Türleri

- Kategorik veriler
 - Nominal
 - Dikotom: kadın veya erkek
 - Kan grupları

Characteristic	Peginterferon Alfa-2a plus Placebo (N=271)	Peginterferon Alfa-2a plus Lamivudine (N=271)	Lamivudine (N=272)
Male sex — no. (%)	214 (79)	208 (77)	215 (79)
Race or ethnic group — no. (%)†			
White	24 (9)	23 (8)	32 (12)
Asian	237 (87)	236 (87)	232 (85)
Black	4 (1)	4 (1)	3 (1)
Other	6 (2)	8 (3)	5 (2)

Veri Türleri

- Kategorik veriler
 - Ordinal
 - Ciddiyet skalası
 - Derecelendirme

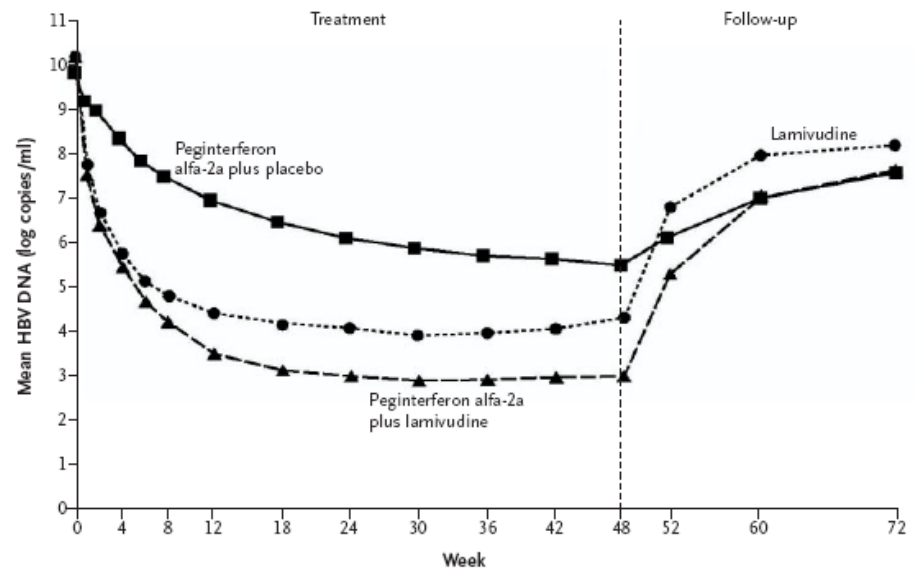
Table 1. (Continued.)

Variable	1 Yr of Trastuzumab (N=1694)	Observation (N=1693)
Histologic grade of tumor — no. (%)		
3 (Poorly differentiated)	1015 (59.9)	1012 (59.8)
2 (Moderately differentiated)	546 (32.2)	554 (32.7)
1 (Well differentiated)	45 (2.7)	42 (2.5)
Not assessed	75 (4.4)	76 (4.5)
Missing	13 (0.8)	9 (0.5)

Veri Türleri

- Devamlı: ölçülebilir değerler
 - Yaş
 - Kolesterol düzeyi

B



No. of Patients

Peginterferon alfa-2a plus placebo	271	269	265	265	262	263	259	258	255	248	254	241	248
Peginterferon alfa-2a plus lamivudine	271	268	268	267	268	263	259	254	255	249	254	244	254
Lamivudine	272	267	267	267	263	263	259	260	244	249	248	228	241

Figure 1. Rates of HBeAg Seroconversion (Panel A) and HBV DNA Levels (Panel B), from Baseline to Week 72.

HBeAg seroconversion was defined by the loss of HBeAg and the presence of anti-HBe antibody. Log to the base 10 was used. The information about the number of patients refers only to Panel B.

Veri Türleri

	Dikotom	Ordinal	Devamlı
	0	1	11
	1	2	24
	0	3	224
	1	5	45
	1	2	56
	0	4	57
	0	3	866
	1	3	34
	0	1	23

Veri Türleri

	Devamlı, Sürekli	Kategorik, Nominal
	17	1
	19	1
	21	1
	24	1
	27	2
	32	2
	33	2
	36	3
	41	3

Veri Türleri

	String, Sözel	Nümerik, Sayısal
	Hemşire	1
	Hemşire	1
	Hekim	2
	hemşire	1
	Hekim	2
	teknisyen	3
	hemşire	1
	hekim	2
	teknisyen	3

Meslek grubu	Nominal	Dummy Çiftçi	Dummy Hemşire
Çiftçi	1	1	0
Ev hanımı	2	0	0
Hekim	3	0	0
Hemşire	4	0	1
Polis	5	0	0

Eksik verilerle uğraşmanın yolları

1. Eksik verilerin olduğu değişken silinir.
2. Eksik verileri gösteren dikotom bir değişken oluşturulur.
3. Daha fazla veri elde etmek için çalışılır.
4. Eksik verilerin değerleri tahmin edilerek doldurulur.

Tanımlayıcı İstatistik merkezi eğilim ölçütleri

- Aritmetik ortalama
- Ortanca
- Mod
- Geometrik ortalama

Tanımlayıcı İstatistik merkezi eğilim ölçütleri

İki grup veri:

1, 2, 3, 4, 1000

1, 2, 3, 4, 5

Results

```
. sum var2, det
```

var2

Percentiles		Smallest			
1%	1	1			
5%	1	2			
10%	1	3		Obs	5
25%	2	4		Sum of Wgt.	5
50%	3			Mean	3
		Largest		Std. Dev.	1.581139
75%	4	2			
90%	5	3		Variance	2.5
95%	5	4		Skewness	0
99%	5	5		Kurtosis	1.7

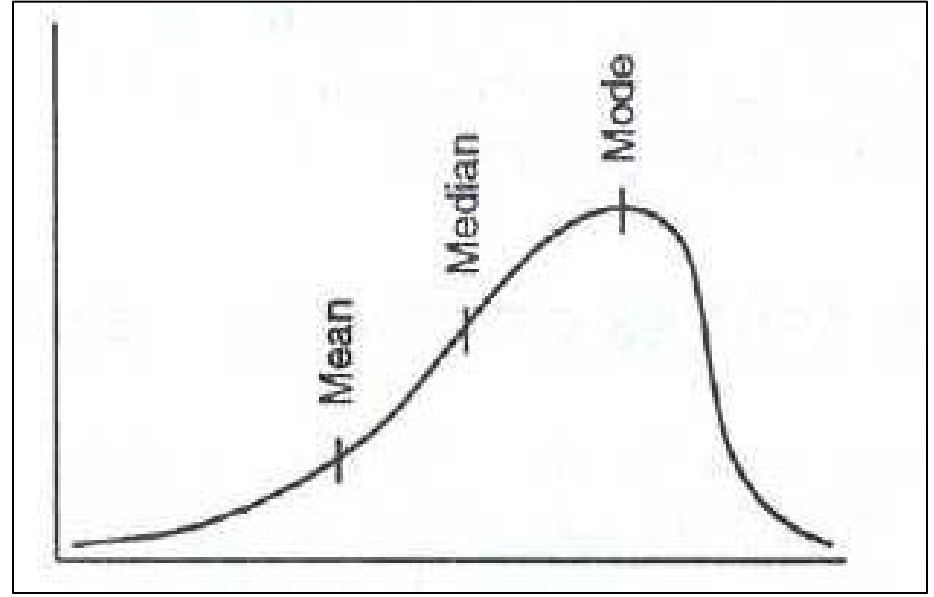
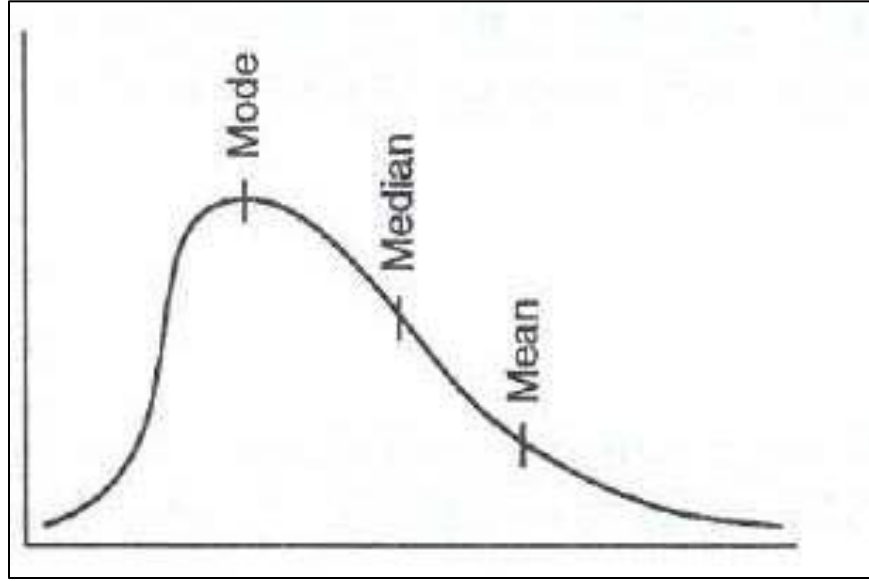


Results

```
. sum var1,det
```

Percentiles		Smallest		
1%	1	1		
5%	1	2		
10%	1	3	Obs	5
25%	2	4	Sum of Wgt.	5
50%	3		Mean	202
		Largest	Std. Dev.	446.097
75%	4	2		
90%	1000	3	Variance	199002.5
95%	1000	4	Skewness	1.499976
99%	1000	1000	Kurtosis	3.249969

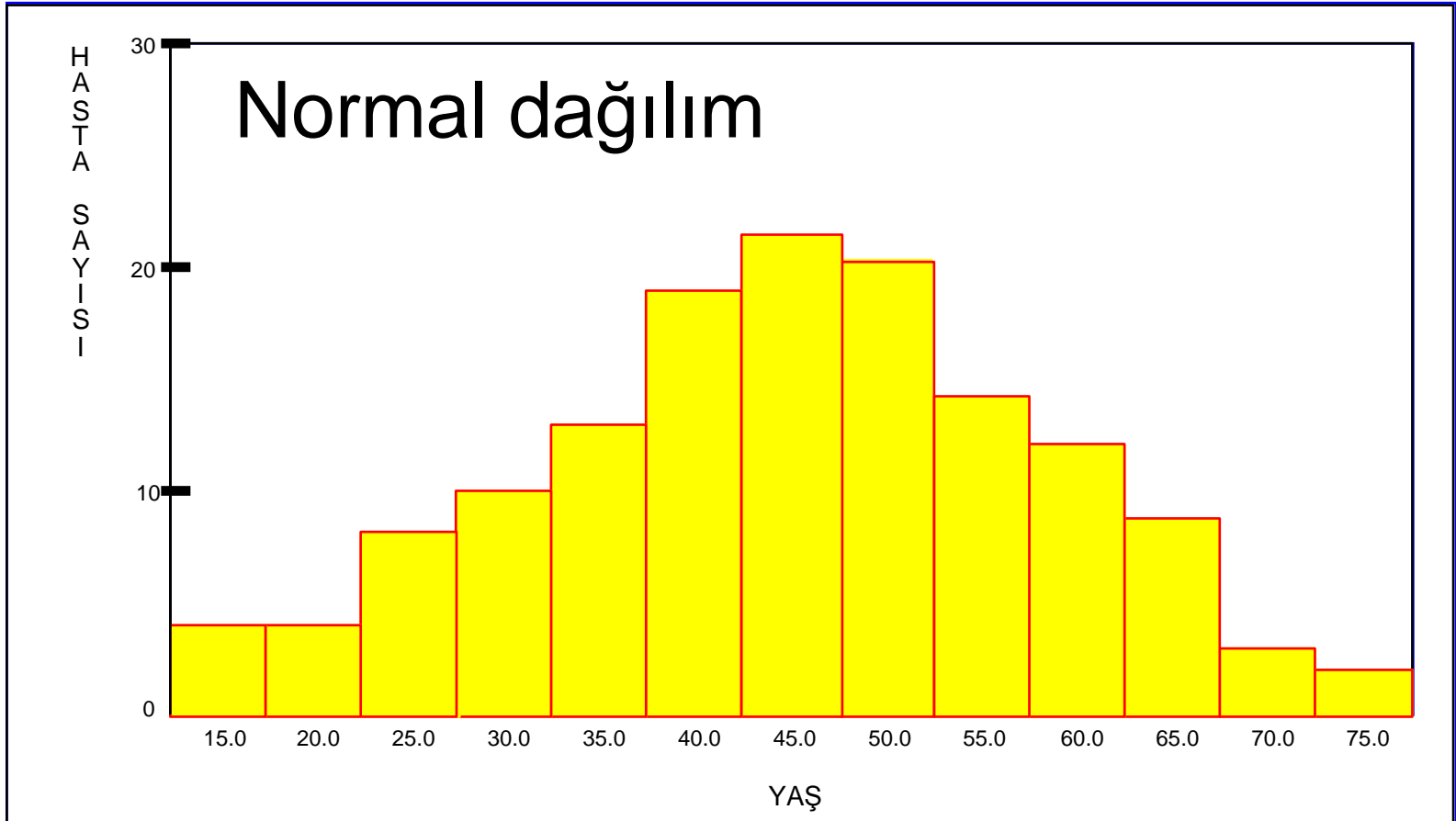
```
.
```

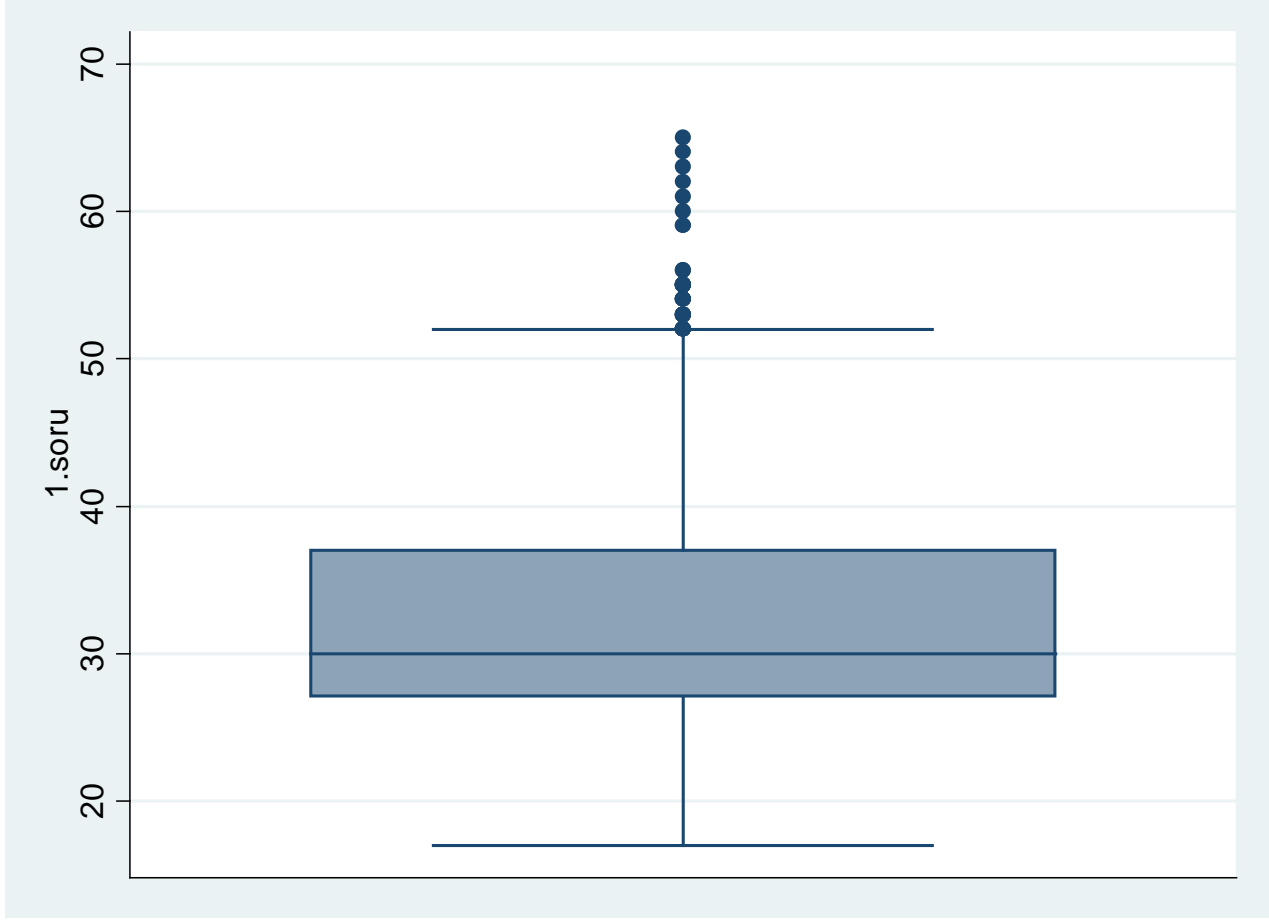



Tanımlayıcı İstatistik merkezi eğilim ölçütleri

- Görsel yöntem
 - Histogram
 - Gövde-yaprak grafiği
 - Boksör torbası grafiği
- İstatistik testler
 - Kolmogorov-Smirnov
 - Lilliefors
 - Shapiro wilk
- Varyasyon katsayısı (SD/ortalama)
 - $SD/ortalama \leq \%30$ ise, dağılım \approx normaldir.

Tanımlayıcı İstatistik merkezi eğilim ölçütleri





Tanımlayıcı İstatistik yayılım ölçütleri

Değer aralığı (*range*) : *min-max*

Standart sapma (*standard deviation*)

Persentil (*percentile*)

Çeyreklerarası aralık (*interquartile range*)

Tanımlayıcı İstatistik yayılım ölçütleri

Standart sapma

$$SD = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$

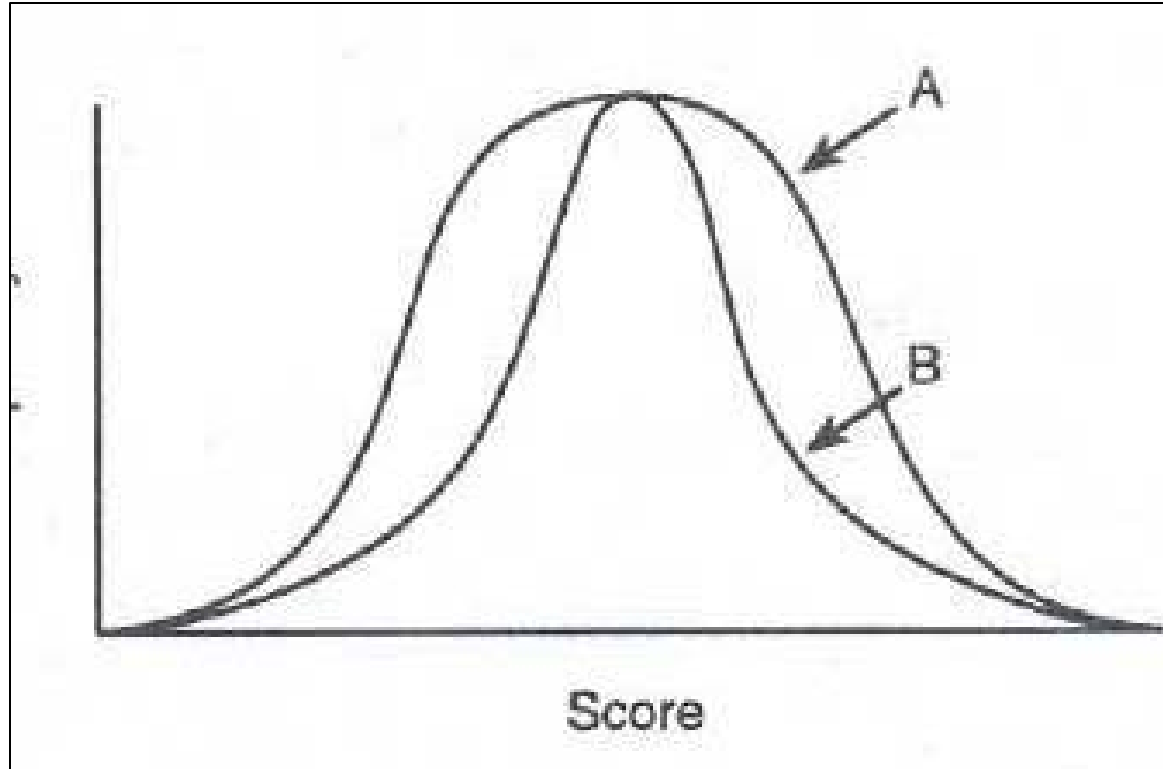
X=ortalama

n=denek sayısı

Tanımlayıcı İstatistik yayılım ölçütleri

- Standart sapma
 - SD, sd, Σ , σ (sigma)
 - Varyans' ın karekökü
 - Varyans' dan daha sık kullanılır
 - Tüm değerler eşitse SD sifıra eşittir.
 - Değerler arasında farklar arttıkça SD büyür.

Hangi verinin standart sapması daha büyük?





Results

```

. sum yas, det

                1.soru
-----
Percentiles      smallest
1%                19        17
5%                23        17
10%               24        18
25%               27        18
50%               30
75%               37
90%               43
95%               46
99%               54

                Largest
                62
                63
                64
                65

Obs                2372
Sum of Wgt.       2372

Mean              32.06661
Std. Dev.         7.468858

Variance          55.78384
Skewness          .9172034
Kurtosis          3.664814

```

Ortalamanın Standart Hatası

Standart sapmadan farklı bir ölçüt

Standart sapma ve örneklem sayısına bağlıdır

Standart sapma / $\sqrt{\text{örneklem sayısı}}$

Örneklem sayısı arttıkça standart hata küçülür

SS: verilerin yayılımını,

SH: ortalamanın belirsizliğini gösterir

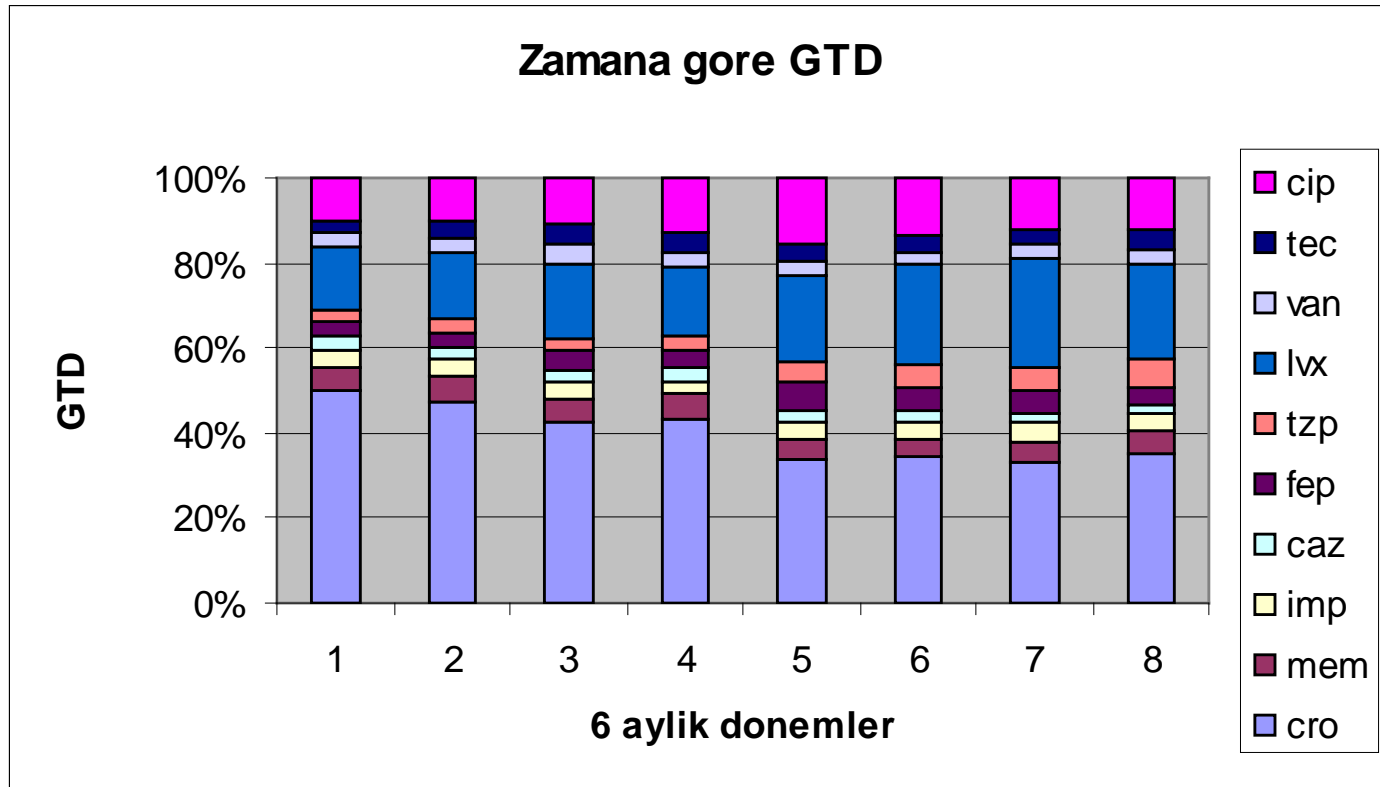
Tanımlayıcı İstatistik

- Frekans dağılımı ve kümülatif frekans

Yaş	n=	%	Kümülatif sıklık
5-14	15	17.6	17.6
15-24	19	22.3	39.9
25-34	21	24.8	64.7
35-44	30	35.3	100

Tanımlayıcı İstatistik

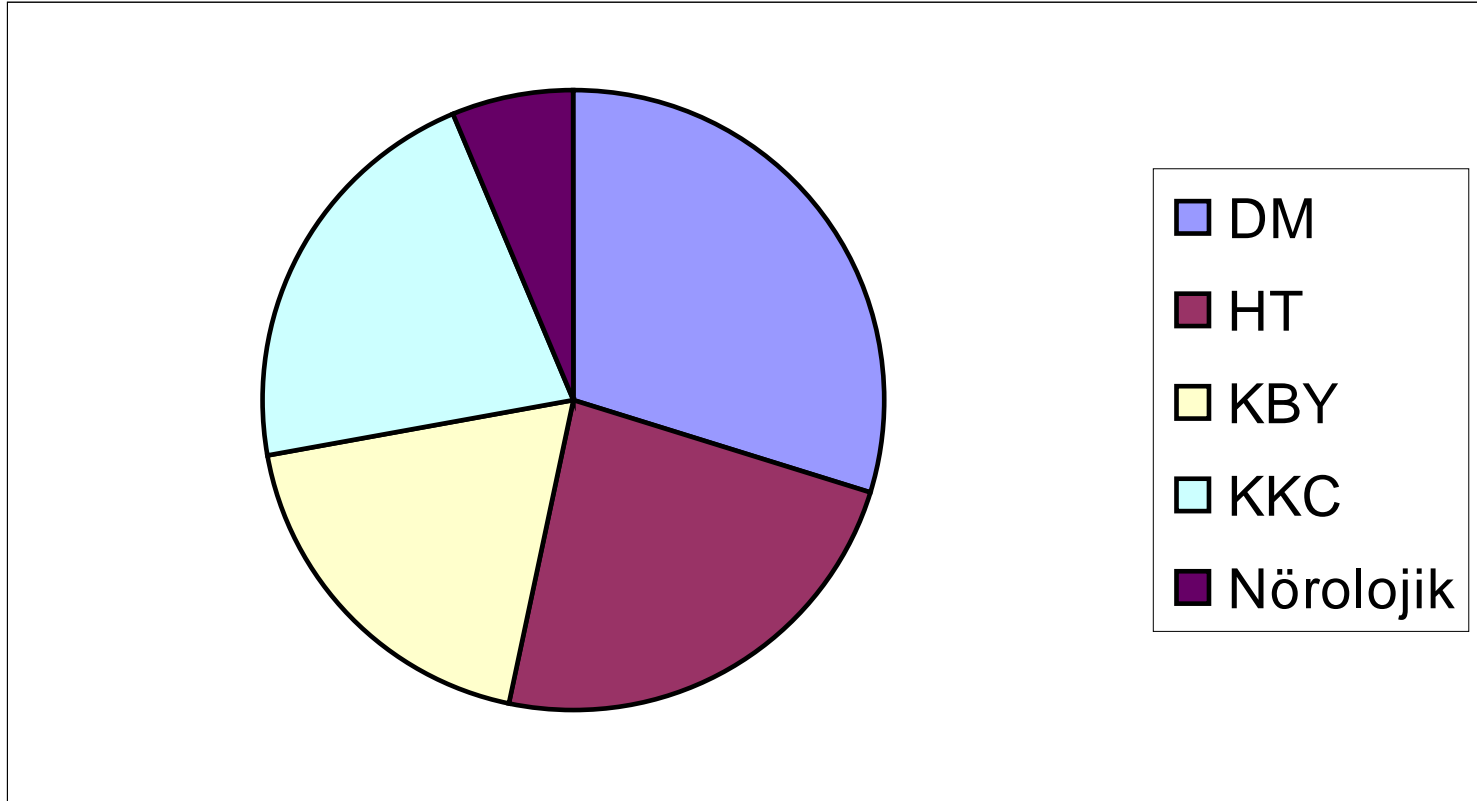
- Veri türleri ve grafik tipleri



Yıllar içinde DDD

Tanımlayıcı İstatistik

- Veri türleri ve grafik tipleri



Uyarı: Gruplar arasında kesişim olmamalı

Tek Değişkenli Analiz

İstatistik Test Seçimi

Parametrik teknikler

- Öncelikle tercih edilirler
- Verilerin normal dağılıma uyması gerek

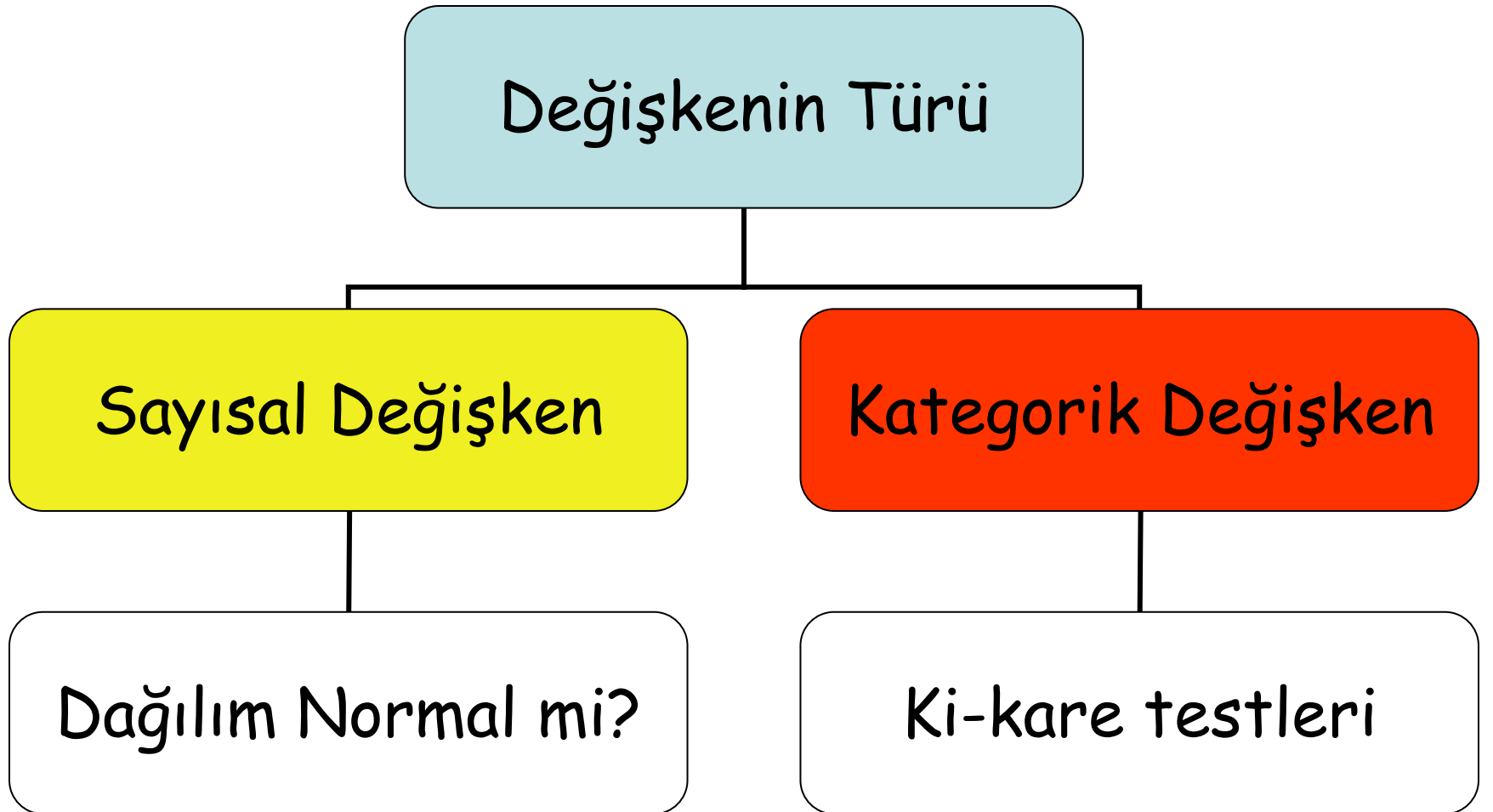
Nonparametrik teknikler

- Verilerin normal dağılım göstermediği, denek sayısının az olduğu ya da diğer parametrik test uygulayabilme koşullarından herhangi birinin sağlanamadığı durumlarda, parametrik testlerin eşdeğeri ya da alternatifi olan nonparametrik testler uygulanabilir.

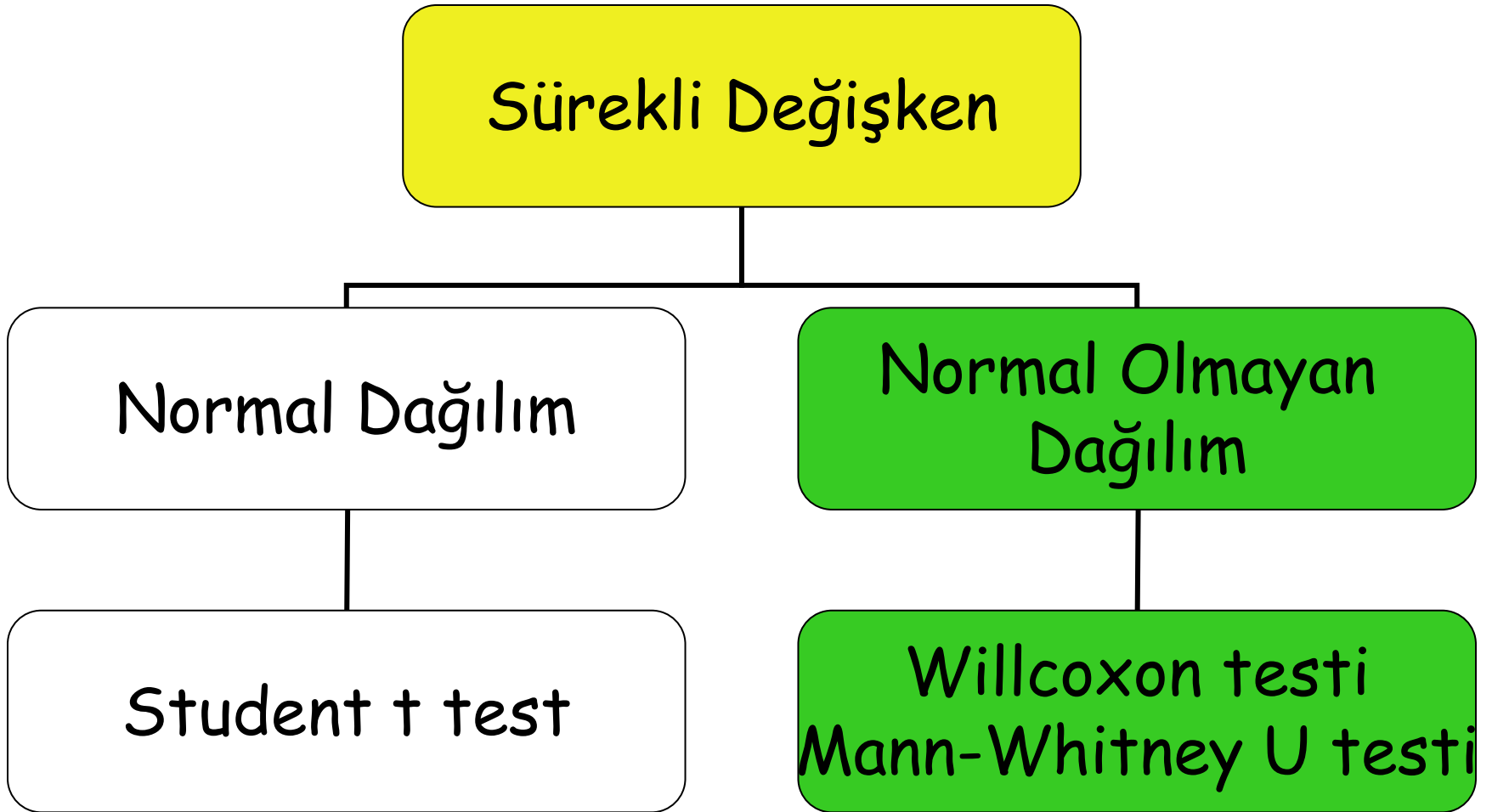
Sunumun Tabloları

Tablolar	Test	Örnek
Tablo 1 Tanımlayıcı istatistik	Ortalama, ortanca, standart sapma, vb.	Hastaların %51' i kadındır
Tablo 2 Farkları inceleyen analitik istatistik	İki grup karşılaştırması için t-testleri ve ki-kare (çapraz tablo) testleri	Ciddi seyirli hastalarda AST düzeyi daha yüksektir
Tablo 3 İlişki ve kestirim inceleyen analitik istatistik	Bağıntı analizi ve regresyon analizi	Çok değişkenli analiz sonucunda, AST düzeyi 300' ün üzerinde olanlarda olmayanlara göre ölüm oranı 3 kat fazladır (odds oranı).

Tek deęişkenli analiz: İki Grubun Karşılaştırılması



İki Grubun Karşılaştırılması



İki Grubun Karşılaştırılması

Kategorik Değişkenler

```
graph TD; A[Kategorik Değişkenler] --> B[Normal Dağılım]; A --> C[Normal Olmayan Dağılım]; B --> D[Ki Kare Veya Fisher Kesinleme]; C --> E[Kruskall Wallis];
```

Normal Dağılım

Ki Kare
Veya Fisher Kesinleme

Normal Olmayan
Dağılım

Kruskall Wallis

Etken-sonuç İlişkisi

	Şizofreni VAR	Şizofreni YOK
Çocukken SSF enf. VAR	a	b
YOK	c	d

İlişki (association)

Relatif risk (kohort)

$$a/a+b : c/c+d$$

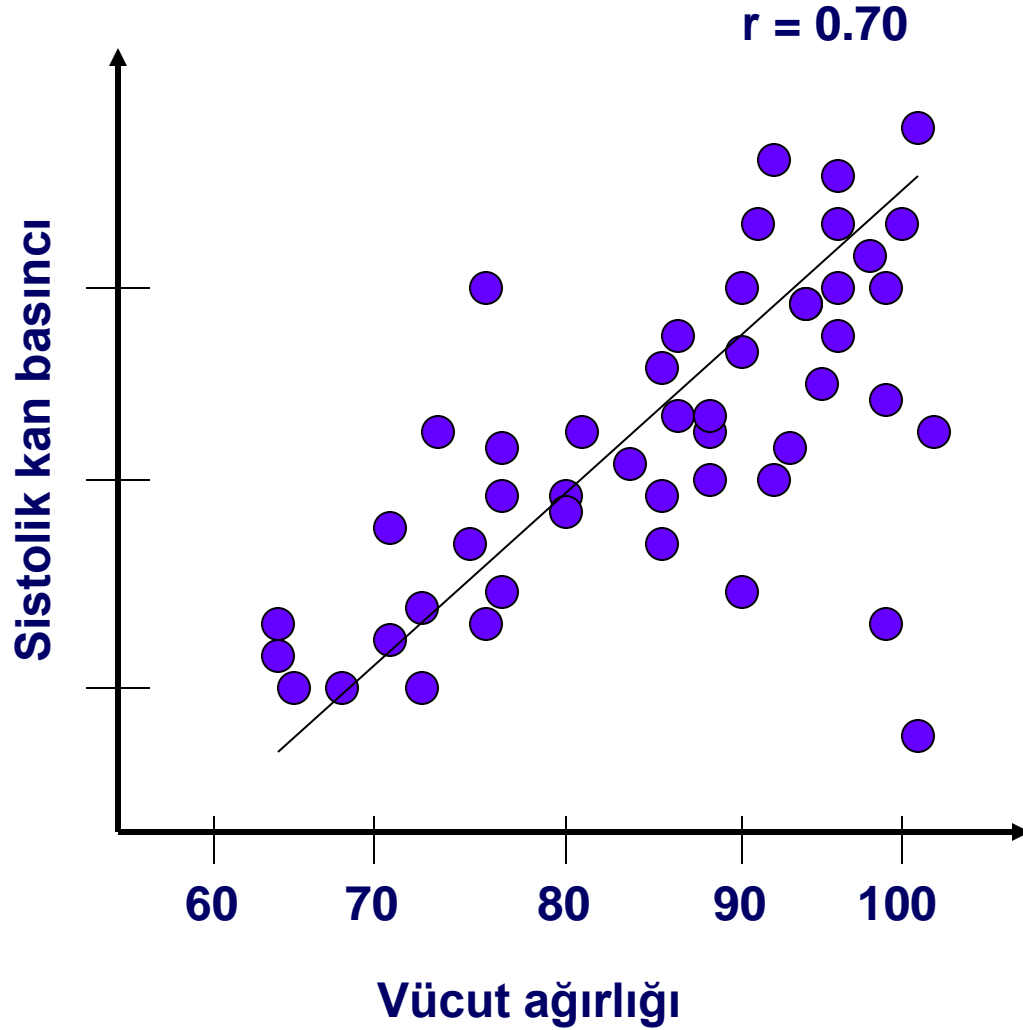
Odds oranı (olgu-kontrol)

$$a/a+c : b/b+d$$

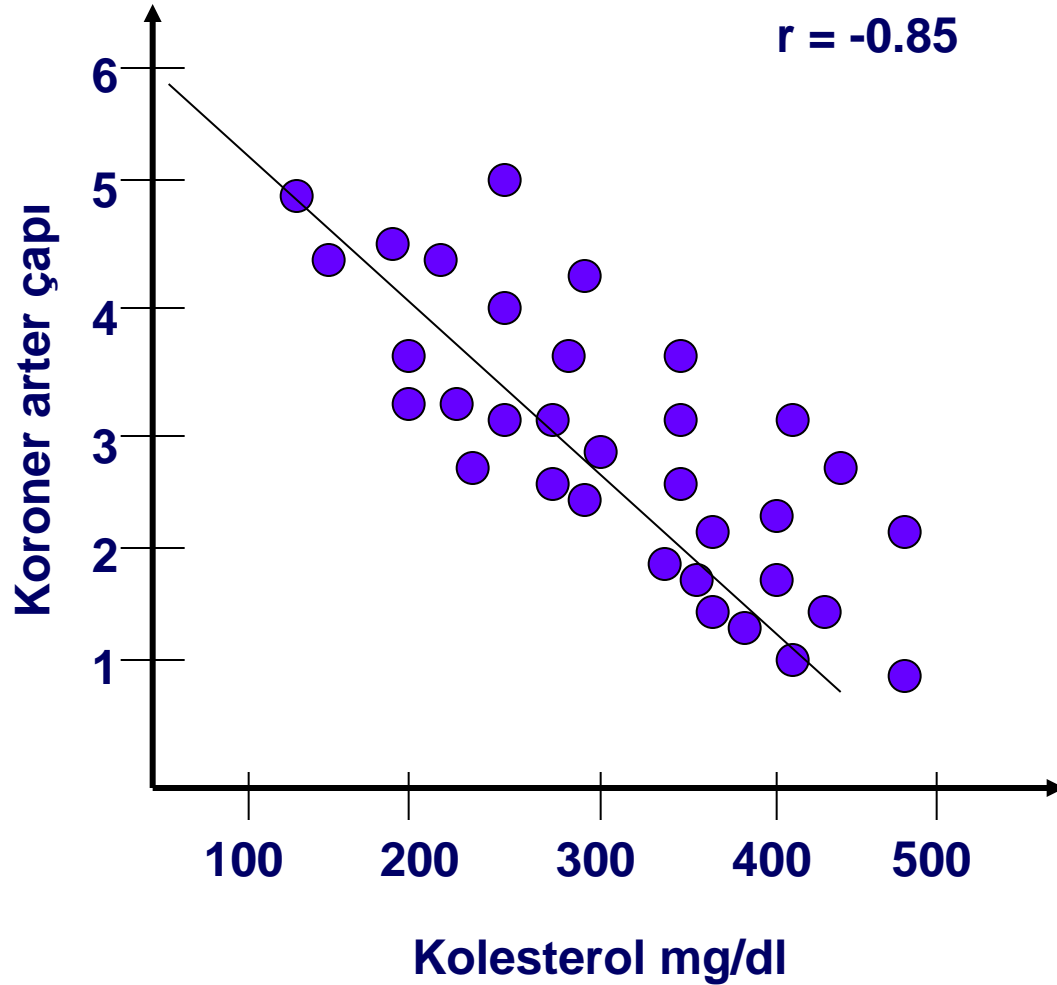
Odds ve Odds Oranı Nedir?

- Bir ilacın yan etki gelişme oranı $1/100$ ise,
odds= $1/99= 0.0101$
- Doğumda erkek çocuk olma ihtimali $\frac{1}{2}$,
odds= $1/1=1$
- İki odds' un birbirlerine oranı=odds oranı

Bağıntı analizi: Korelasyon



Bağıntı analizi: Korelasyon



Korelasyon

- Korelasyon kat sayısı -1 ve $+1$ arasında değişir
 - $0-0.25$ Hiç ilişki yok/zayıf ilişki
 - $0.25-0.50$ Zayıf - orta derece ilişki
 - $0.50-0.75$ İyi derece ilişki
 - $0.75-1.00$ Çok iyi derecede ilişki
- Spearman vs Pearson KK
- P-değeri
- Aşırı değerler problemi

Çok Değişkenli Analiz

Tabakalama yapılması gereken birden çok
değişken varsa,

ÇOK DEĞİŞKENLİ ANALİZ

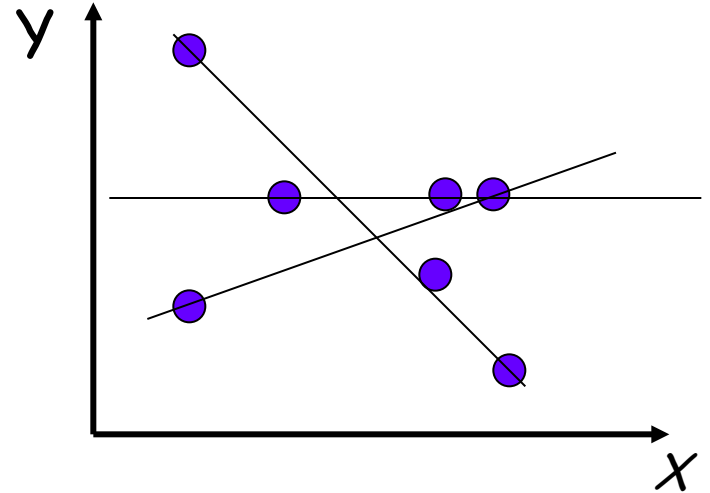
yapılır

Regresyon

Ana elementler

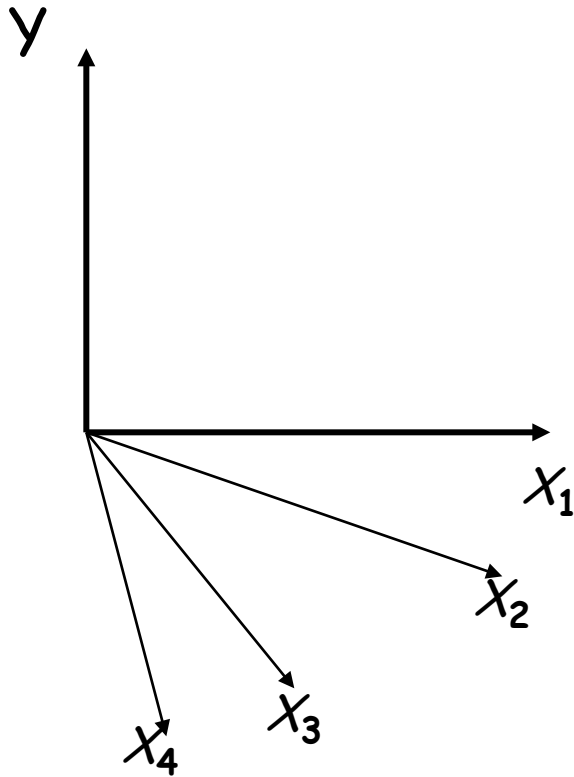
- Prediktör veya bağımsız değişken (x)
- Yanıt, sonuç (*outcome*) veya bağımlı değişken (Y)

$$Y = a + bx$$



Çoklu Regresyon: Multiple regression

- Her zaman prediktör değişkenler birden fazladır.

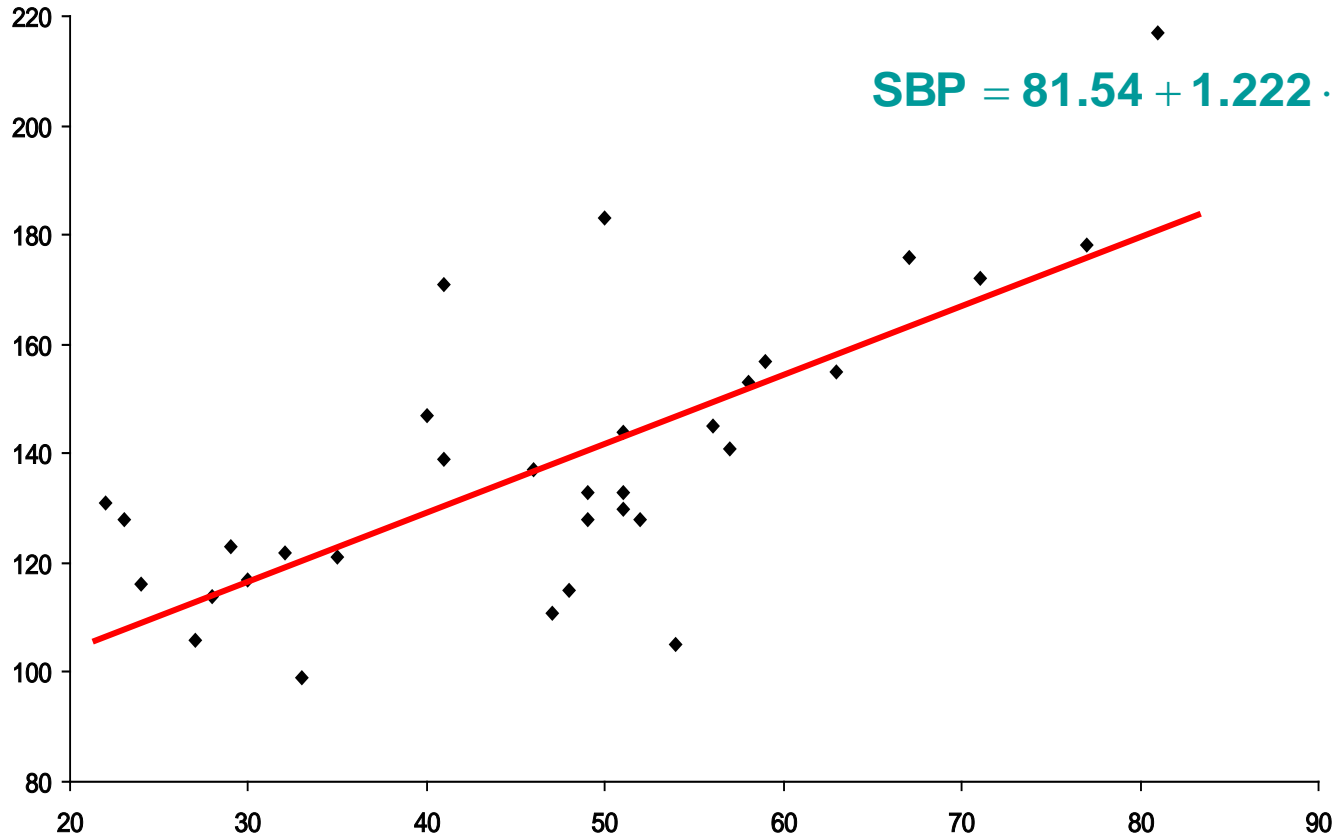


$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$$

Çok Değişkenli Analiz

Regresyon türü	Bağımlı değişken (outcome, sonuç)
Linear	Sürekli
Logistic	Dikotom
Cox	Dikotom
Poisson	Nadir

SBP (mm Hg)



SBP = 81.54 + 1.222 · Age

Age (years)



Review

```

oci 30 20 40 50
tab asi grip, chi
csi 2 20 17 64, fisher
csi 2 20 17 64, ex
csi 2 20 17 64, chi
csi 2 20 17 64
csi 2 40 17 64
csi 2 40 17 64, ex
tab grip asi
tab grip asi, chi
tab asisonrasigrip asi, c
tab asisonrasigrip asi,
tab asisonrasigrip asi,
logistic asisonrasigrip .
logistic asisonrasigrip .
logistic asisonrasigrip .
    
```

Variables

```

asisonrasigrip
antip
no
il
yas
cins
isyeri
meslek
hekim
bolum
cocuk
cocukyas1
cocukyas2
cocukyas3
cocukyas4
kres
hanehalki
asi
evetneden
yanetki
hayir1
hayir2
hayir3
hayir4
hayir5
    
```

Results

```

Log likelihood = -1233.1106
Prob > chi2      = 0.0000
Pseudo R2       = 0.0086
    
```

asisonrasi~p	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interva]	
asi	1.561187	.2272917	3.06	0.002	1.173625	2.076732
cins	1.417571	.1523884	3.25	0.001	1.148262	1.750043

```

. logistic asisonrasigrip asi cins yas

Logistic regression
Log likelihood = -1231.1693
Number of obs   = 2355
LR chi2(3)      = 21.67
Prob > chi2     = 0.0001
Pseudo R2      = 0.0087
    
```

asisonrasi~p	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interva]	
asi	1.550026	.2267601	3.00	0.003	1.163625	2.064738
cins	1.412579	.1519495	3.21	0.001	1.144063	1.744117
yas	.992997	.0067826	-1.03	0.304	.9797919	1.00638

Command

```
logistic asisonrasigrip asi cins yas
```


Lojistik Regresyon

- DMA heparin vs plasebo
- Randomizasyon
- Diğer risk faktörleri: cinsiyet, koroner arter hastalığı, aspirin kullanımı

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$$

$$DVT = a + b_1(\text{heparin}) + b_2(\text{cinsiyet}) + b_3(KAH) + b_4(\text{aspirin})$$

0.5

1.5

3

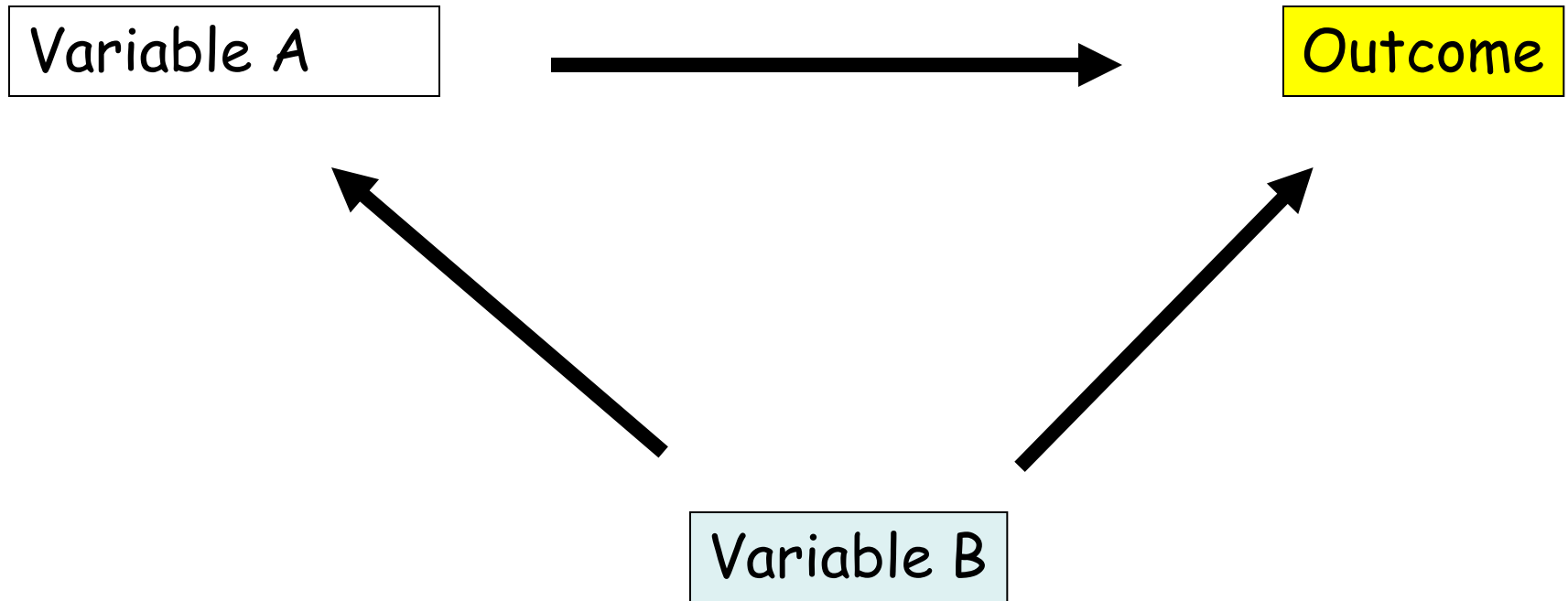
0.6

P-değeri kadar önemli

Saptanan etkinin/ilişkinin büyüklüğü ve kesinliği

- Etkiyi gösteren kestirim
 - Odds oranı, Relatif risk, Hazard ratio
- Güven aralıkları

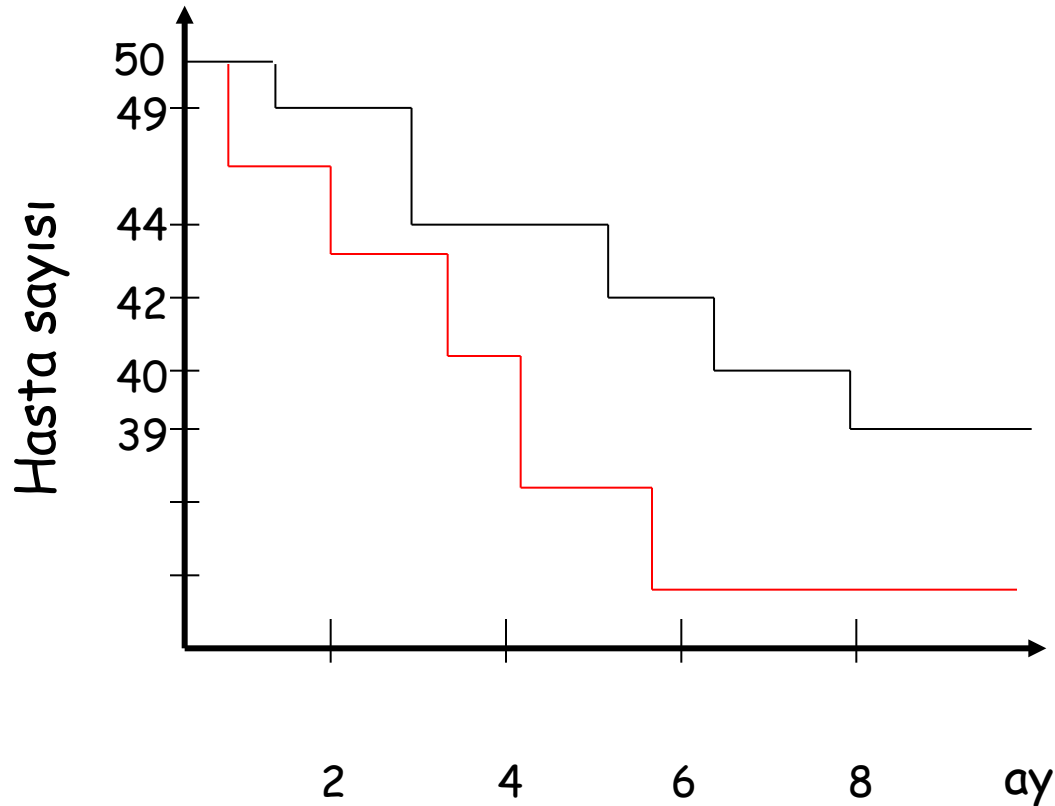
Confounder=Karıştırıcı Değişken



Sağkalım Analizi: Survival Analysis (Time-to-event analysis)

Sonuç değişkeni = Bir “olaya” kadar geçen süre

“Olay” = ölüm, iyileşme, relaps...



Sağkalım Analizi: Cox regresyonu

Cox regression

Log likelihood = -6.2444702

Number of obs = 10
LR chi2(4) = 0.97
Prob > chi2 = 0.9141
Pseudo R2 = 0.0722

DVT	Hazard ratio	Std. Err.	z	P> z	[95% Conf Interval]	
Heparin	.50	.023	-2.81	0.003	.15	.72
Kadin	1.48	1.08	0.01	0.504	.095	23.17
KAH	3.06	.03	2.36	0.009	1.34	12.37
aspirin	0.58	0.08	0.31	0.622	.46	1.03