

ESANTIONUL DE STUDIU CULEGEREA DATELOR



REZUMATUL EPISODULUI PRECEDENT...

IPOTEZA

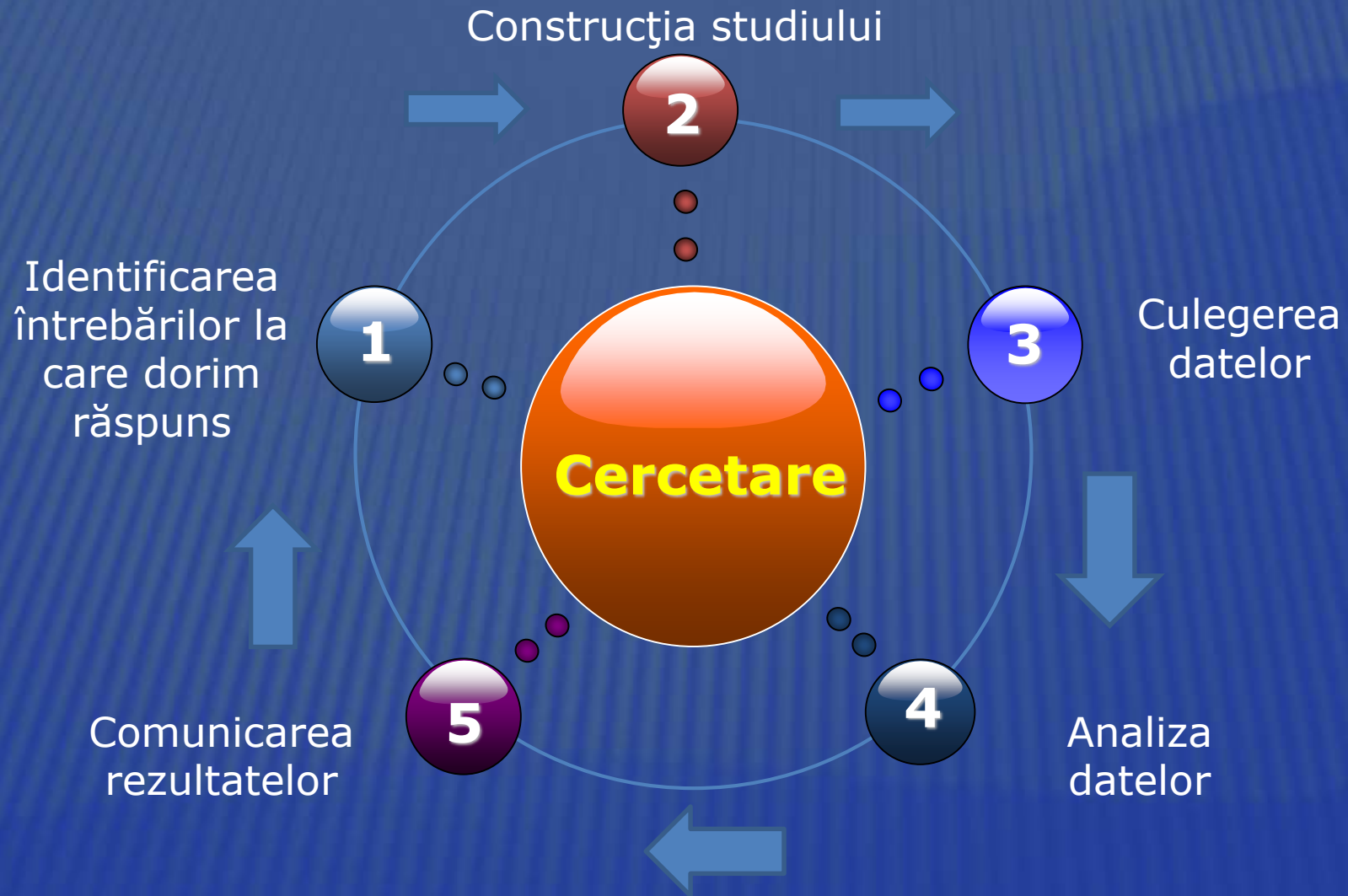
- × Cercetarea trebuie să înceapă cu o întrebare
- × Este partea creativă și personală a cercetării

STUDIUL BIBLIOGRAFIC

STUDIUL Regasirea informatiei

- × *Motoare de indexare*

DIAGRAMA CERCETARII



OBIECTIVE

- × Alegerea eșantionului de studiu
- × Calculul dimensiunii unui eșantion
- × Clasificarea studiilor după modalitatea de intervenție
- × Clasificarea studiilor după modalitatea de culegere a datelor



ALEGEREA EȘANTIONULUI DE STUDIU

- × Obiective propuse
- × Tipul studiului
- × Modalitățile de alegere a subiecților
- × Timpul disponibil
- × Resurse financiare și umane disponibile
- × Procedura folosită
- × Accesul la date

POPULAȚIA CUPRINSĂ ÎN STUDIU:

Un caz, o serie de cazuri



Exhaustiv



Eșantionare





• Example: "Normal plasma cholesterol in an 88-year-old man who eats 25 eggs a day" [Kern J, NEJM 1991; 324:896-899]

CASE REPORT

An 88-year-old man who lived in a retirement community complained only of loneliness since his wife's death. He was an articulate, well-educated elderly man, healthy except for an extremely poor memory without other specific neurologic deficits. He had been given a diagnosis of Alzheimer's disease and was intermittently depressed. His general health had been excellent, without notable symptoms. He had mild constipation. His weight had been constant at 82 to 86 kg (height, 1.87 m). He had no history (according to the patient and his personal physician of 15 years) of heart disease, stroke, or kidney disease except for an episode of mild chest pain three years earlier. The only objective change at that time was transient depression of the ST segments and T waves in the lateral leads on his electrocardiogram. The patient had been treated for angina and had had no recurrence. There was no history of gallstones or of symptoms of biliary tract disease, but no cholecystography or ultrasound examination had been done recently. His physician's records showed numerous serum cholesterol measurements that ranged from 3.88 to 5.18 mmol per liter (150 to 200 mg per deciliter).

The patient had never smoked and never drank excessively. His father died of unknown causes at the age of 40, and his mother died at 76. One sister died at the age of 82, and another was alive at 86; their plasma lipid values were not available.

The patient's poor memory impaired the accuracy of the dietary history, but his consumption of 20 to 30 eggs a day was verified. Although he could not remember the duration of this eating pattern, his physician attested to its presence for 15 years; a friend, for even longer. He always soft-boiled the eggs and ate them throughout the day. He kept a careful record, egg by egg, of the number ingested each day. The nurse at the retirement home confirmed the daily delivery to him of approximately two dozen eggs. A psychiatrist and a clinical psychologist had characterized this unusual eating habit as compulsive behavior, based on complex psychological factors. Efforts to modify the behavior had been unsuccessful. The patient stated, "Eating these eggs ruins my life, but I can't help it."

CAZ SAU SERIE DE CAZURI

× Avantaje:

- + Descrierea unei situații neobișnuite
- + Excelent pentru generarea de noi ipoteze medicale

× Dezavantaje:

- + Reflectă fenomene de scurtă durată
- + Cazuri selectate de investigator (posibilitate de eroare)
- + Nu există termen de comparație (control)

POPULAȚIA CUPRINSĂ ÎN STUDIU - EXHAUSTIV

- × Toți subiecții populației țintă
- × Imposibil de realizat pentru că:
 - + costurile ar fi foarte ridicate iar durată studiului îndelungată - în studiile clinice
 - + duce la alterarea întregii populații statistice - în studiile biologice
 - × testarea eficienței unui vaccin
 - × testarea eficienței unui produs

POPULAȚIA DE STUDIU

- × Populația = mulțime de elemente care au anumite caractere comune
 - + Oameni
 - + Animale
 - + Obiecte
- × Ex.: studiem efectele unor tratamente asupra osteoartritei
 - + Populația = toată populația din lume care prezintă osteoartrită

POPULAȚIA CUPRINSĂ ÎN STUDIU → EȘANTION

- × Alegem din populația disponibilă un grup de subiecți → eșantionul
- × Calitățile eșantionului:
 - + **reprezentativ** prin **talie** și **componență** pentru populația țintă
 - rezultatele obținute pe eșantion pot fi extrapolate asupra întregii populații vizate prin metode ale statisticii inferențiale

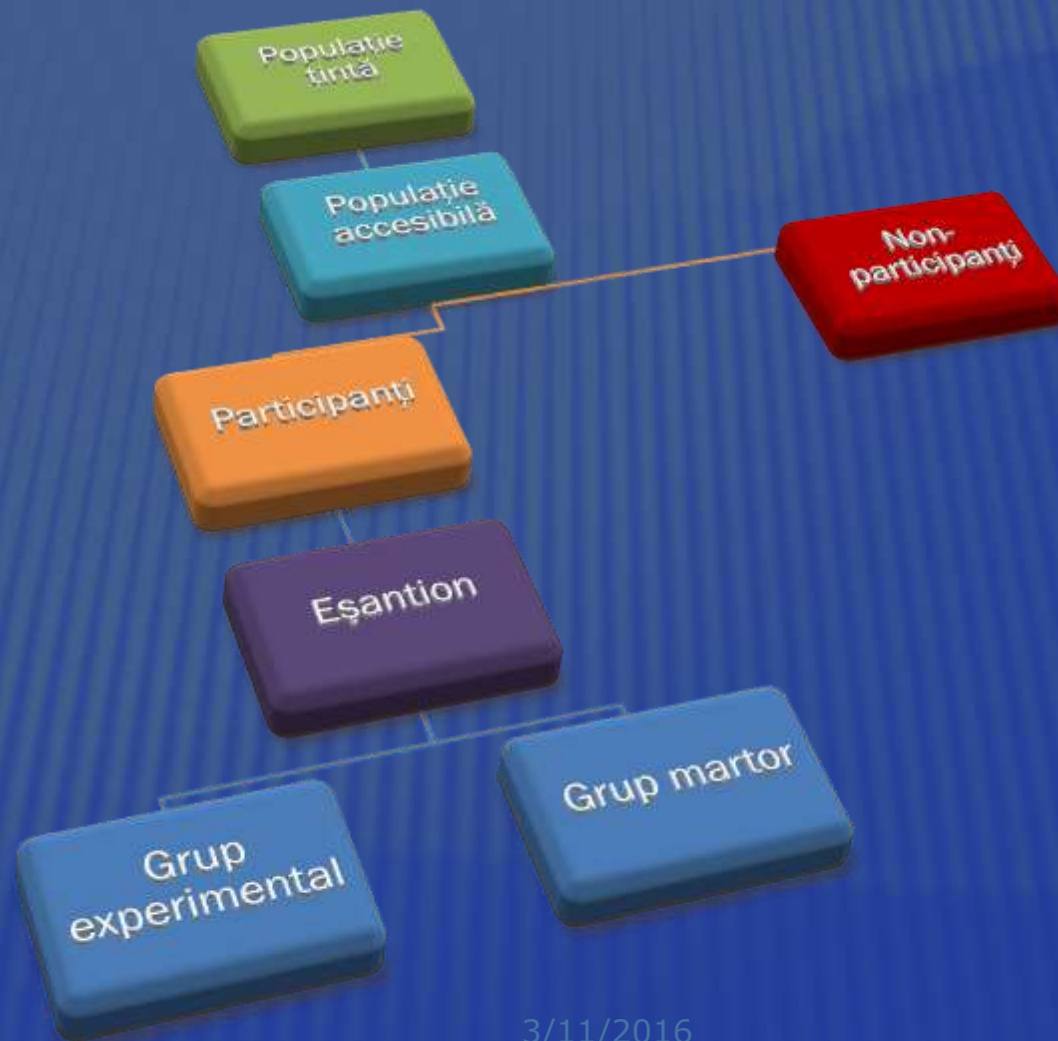
CRITERII DE INCLUDERE ȘI EXCLUDERE

× Definirea criteriilor de includere în studiu:
trebuie luate în considerare toate variațiile
caracteristice populației:

- + Clinice
- + Demografice
- + Factorii geografici

Importanța acestor factori în patologia de studiu

SELECȚIA EȘANTIONULUI



ETAPELE DEFINIRII EȘANTIONULUI

× Alegerea tipului de eșantionare:

1. Simplu randomizat
2. Sistematic
3. Stratificat
4. Cluster
5. Asignare aleatoare

× Alegerea modului de eșantionare:

1. Talia eșantionului
2. Pierduți din studiu (pierduți din vedere, non-răspunzători)

TIPURI DE EȘANTIONARE

1. Eșantion simplu randomizat:

- + Extragerea la întâmplare a elementelor din populația de studiu
- + Fiecare element are aceeași șansă de a fi inclus în eșantion
 - × **RANDBETWEEN**

TIPURI DE EȘANTIONARE

2. Eșantionare sistematică

- × Selectăm pentru a fi inclus în eșantion tot al k-lea element

$$k = \frac{\text{talie_populatie}}{\text{talie_esantion}}$$

- × Nu o folosim în cazul în care în structura de eșantionare ar putea să apară o periodicitate.



© AP

TIPURI DE EȘANTIONARE

3. Eșantionare stratificată

- + Populația este împărțită în grupuri omogene și mutual exclusive numite straturi → extragem aleator din straturi elementele eșantionului
- + Populația poate fi stratificată ținând cont de oricare variabilă de studiu: vârstă, sex, antecedente, greutate etc.
- + Straturile sunt reprezentate în eșantion în funcție de importanța lor în populație.

TIPURI DE EȘANTIONARE

4. Eșantionare cluster

- + Împărțim populația în clusteri
- + Selectăm aleator o submulțime de clusteri
 - × Evaluarea măsurilor de prevenire ale accidentului vascular cerebral în România → județe = clusteri → randomizare de clusteri
- × Criterii geografice
- × Studii epidemiologice

TIPURI DE EȘANTIONARE

5. Asignare aleatoare

- × Trialuri clinice randomizate (studii pentru compararea diferitelor tratamente)
 - × atribuirea aleatoare a pacienților în grupul martor sau în cel care primesc tratamentul = randomizare
 - pacienții din cele două grupe sunt similari singura diferență este de tratament

TIPURI DE EȘANTIONARE

× Eșantioane independente

- + Prelevarea unuia nu depinde de prelevarea celuilalt
- + Nu au (neapărat) același număr de unități

× Eșantioane dependente

- + Eșantioane perechi
- + Au aceeași talie iar unitățile unui eșantion sunt în corespondență bijectivă cu unitățile celuilalt



DETERMINAREA TALIEI EȘANTIONULUI

DETERMINAREA TALIEI EȘANTIOANELOR

- × Dacă pentru un risc dat α se dorește obținerea unei precizii date k pentru intervalul de încredere, adică acesta să aibă forma:

- × $[f - k, f + k]$

$$\left[f - Z_{\alpha} \cdot \sqrt{\frac{f(1-f)}{n}}, f + Z_{\alpha} \cdot \sqrt{\frac{f(1-f)}{n}} \right]$$

- × atunci ținând seama de faptul că intervalul de încredere se exprimă prin:

$$k = Z_{\alpha} \cdot \sqrt{\frac{f(1-f)}{n}}$$

$$n = \frac{Z_{\alpha}^2 \cdot f(1-f)}{k^2}$$

EXEMPLU

- Se dorește evaluarea prevalenței unei maladii într-o anumită populație. Se știe că în alte studii această prevalență a fost de ordinul a 30% ($=f$). Dacă se dorește o precizie a estimării prevalenței de $k=4\%$ la un risc $\alpha = 0.05$ (cărui îi corespunde $Z_{\alpha} = 1.96$) trebuie realizată această estimare utilizând un eșantion a cărei talie trebuie să fie superioară sau egală cu

$$n = \frac{Z_{\alpha}^2 \cdot f(1-f)}{k^2}$$

$$\frac{(1.96)^2 \times 0.3 \times 0.7}{(0.04)^2}$$

- $n = 505$.

DAR DACĂ NU AM O ESTIMARE ANTERIOARĂ

- Se ia în calcul cel mai prost scenariu posibil:
 - $f = 50\%$

DAR DACĂ NU AM O IDEE ASUPRA PRECIZIEI

- Convenția este să folosim:
 - $k = 5\%$

DETERMINAREA TALIEI EȘANTIOANELOR

- Dacă vreau să estimez o medie iau în calcul diferența între eșantion și populație:

$$E = \bar{x} - \mu = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$n = \left[z_{\frac{\alpha}{2}} \frac{\sigma}{E} \right]^2 \text{ pentru } \alpha = 0,05 \quad z_{\frac{\alpha}{2}} = 1,96$$

ȘI DE UNDE GĂSESC DEVIATIA?

- Alte studii
- Studiu pilot

$$n = \left[1,96 \frac{\sigma}{E} \right]^2$$

EXEMPLU

- Decanatul facultății vrea să verifice în ce măsură evaluările studenților pentru un cadru didactic sunt corecte (care este eșantionul minim necesar pentru ca estimatorul să fie corect statistic)
- Scala de măsură este 0-4 ca urmare amplitudinea este 4, nivelul de confidență cerut 95%, eroarea nu trebuie să depășească 0,25 puncte
- Din evaluarea răspunsurilor deviația standard a fost de 1,5
- $N_{0,25} = (1,96)^2 * (1,5)^2 / (0,25)^2 = 138$
- $N_{0,50} = (1,96)^2 * (1,5)^2 / (0,5)^2 = 34$
- $N_{1,00} = (1,96)^2 * (1,5)^2 / (1,0)^2 = 8$

EȘANTIONUL REPREZENTATIV

- × Un eșantion de o dimensiune statistic corectă asigură calcule corecte nu este neapărat reprezentativ
- × Eșantionul reprezentativ
 - + Corect extras
 - + De dimensiuni corecte



CLASIFICAREA STUDIILOR DUPĂ MODALITATEA DE INTERVENȚIE

STUDII

× Observaționale

- + Identificarea subiecților
- + Observarea și înregistrarea caracteristicilor



× Experimentale

- + Identificarea subiecților
- + Observarea și înregistrarea caracteristicilor
- + INTERVENȚIA
- + Observarea și înregistrarea caracteristicilor

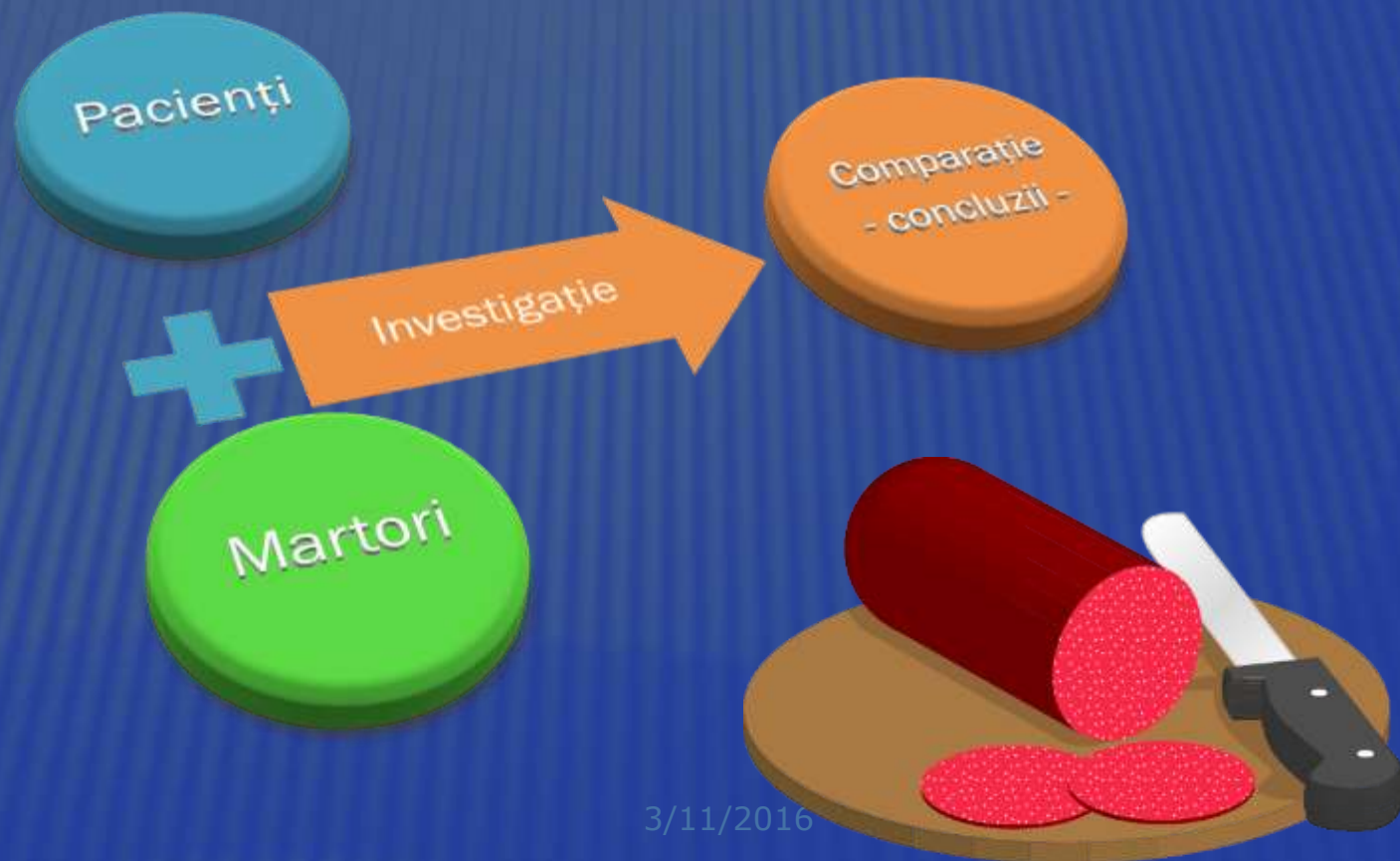


CLASIFICAREA STUDIILOR DUPA MODALITATEA DE CULEGERE A DATELOR



DURATA CULEGERII DATELOR

× Transversală



INFECTIOUS DISEASE

2 Out of 3 Medical Students Don't Know When to Wash Their Hands

By Alice Park [@aliceparkny](#) | Dec. 02, 2011 | 0

It's cold and flu season, which means you should be washing your hands — a lot. And that rule should apply to health care workers most of all, to protect not just themselves but their patients. Problem is, most doctors are confused about when they're supposed to hit the sink, a new study of medical students finds.

Researchers at Hannover Medical School in Germany surveyed 85 medical students who were about to enter their clinical training — when they deal with actual patients. The surveys were administered as part of a lecture class that students must pass before initiating patient contact.

The doctors-to-be were given seven scenarios, five of which required

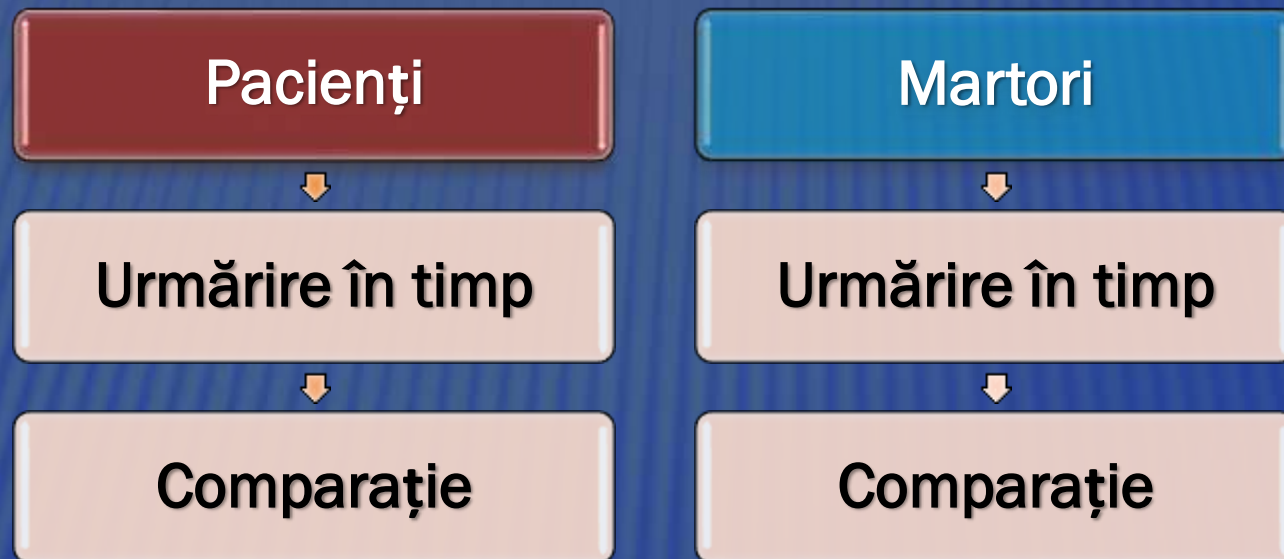


Getty Images

Follow [@TIMEHealth](#)

DURATA CULEGERII DATELOR

× Longitudinal (studii de cohortă)



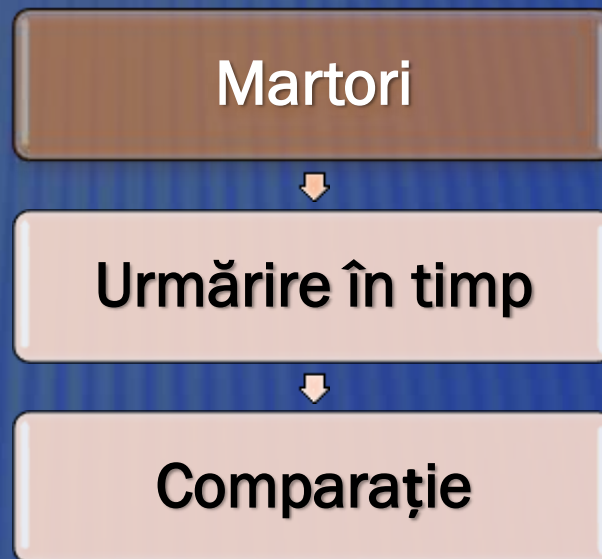
STUDII DE COHORTĂ

- × Permite culegerea informațiilor în mod repetat într-un interval de timp prestabilit pe un anumit grup de subiecți.
 - + **Prospectivă**: culegând date prestabilite la intervalele prevăzute în protocolul studiului, într-un timp determinat de la data debutului studiului.
 - + **Retrospectivă**: din documente medicale

STUDII DE COHORTĂ



Prospectiv



Retrospectiv

TIPURI DE CULEGERE A DATELOR

1. Tip eșantion reprezentativ
2. Tip expus – nonexpus
3. Tip “caz – martor”



EȘANTION REPREZENTATIV

- × Grupul de subiecți studiați este reprezentativ pentru populația țintă
- × Folosit în cazul:
 - + **studiului descriptiv** (descrierea unui fenomen de sănătate)
 - + **studiul analitic** (în care este obligatorie compararea a cel puțin două grupuri de subiecți).

EȘANTION REPREZENTATIV

- × numărul total al subiecților luați în studiu = N

	B^+	B^-	
FR^+	a	b	ne^+
FR^-	c	d	ne^-
	nB^+	nB^-	N

EȘANTION REPREZENTATIV

- × Permite efectuarea celor mai multe tipuri de calcule, neexistând îngrădiri de eșantionare:
 - + calcularea prevalenței bolii
 - + calcularea frecvenței expunerii
 - + măsurarea gradului de asociere între factorul incriminant (de risc sau prognostic) și boală.

EȘANTION REPREZENTATIV

× Limitele:

- + dificultățile de urmărire a unui număr mare de subiecți
 - × probleme de cost
 - × riscul scăderii interesului subiecților pentru studiu
 - × riscul numărului mare de pierduți din vedere.

EXPUS-NONEXPUS

- × Împarte subiecții luați în studiu în două categorii:
 - + expuși
 - + neexpuși
- × Se fixează factorul de expunere și se studiază apariția în timp a îmbolnăvirii
- × Expunere rară
- × Boala = factorul aleator

EXPUS-NONEXPUS

- × Numărul subiecților din cele două grupe poate fi **identic** sau **diferit**.

	B ⁺	B ⁻	
FR ⁺	a	b	<u>ne⁺</u>
FR ⁻	c	d	<u>ne⁻</u>
	-		

EXEMPLU



EXPUS-NONEXPUS

- × Calculele statistice se pot efectua în componenta de deasupra și dedesubtul liniei îngroșate a tabelului de contingență.
- × Se poate compara proporția de subiecți bolnavi dar nu și frecvența expunerii.

EXPUS-NONEXPUS

- × Avantaj:

- + posibilitatea alcătuirii de grupe mai mici de subiecți

- × Dezavantaj:

- + probleme de urmărire a subiecților

CAZ-MARTOR

- × predefinirea a doua grupe de subiecți:
 - + B+ = bolnavi = grupul caz
 - + B- = indemni de boală = grupul martor
 - + criteriul principal de incluziune sau excluziune într-un grup sau altul fiind prezența sau absența bolii

	B ⁺	B ⁻	
FR ⁺	a	b	
FR ⁻	c	d	
	<u>nB⁺</u>	<u>nB⁻</u>	

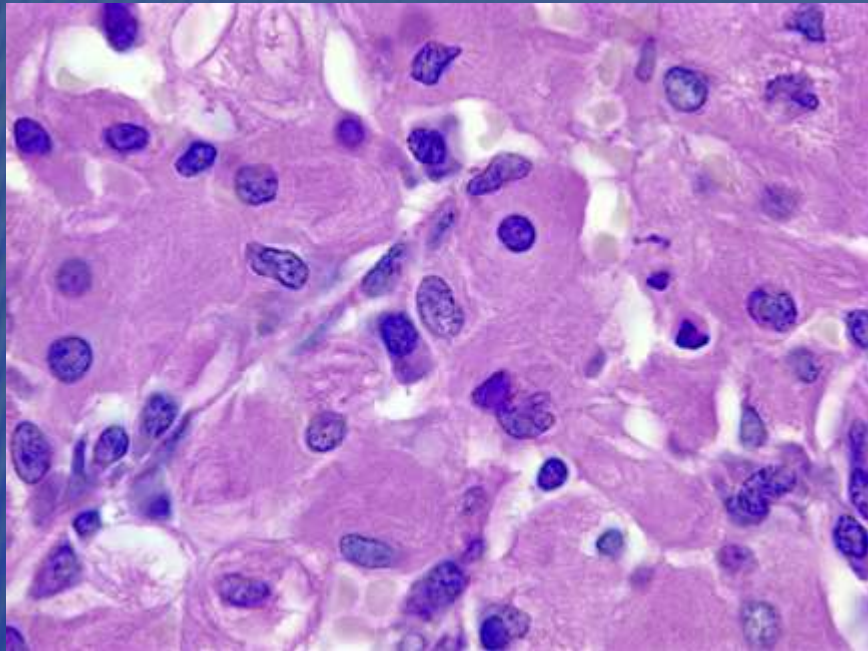
CAZ-MARTOR

- × boala = factorul controlat
- × expunerea = factorul aleator
- × $nB^+ \neq nB^-$
 - + se prestabilește boala și se studiază în general retrospectiv prezența expunerii (factorului de risc/prognostic) în apariția (sau evoluția) îmbolnăvirii

CAZ-MARTOR

- × Calcule și comparații se pot face numai între grupele prestabilite B+ și B-
- × Pot fi comparate proporțiile celor expuși dar nu poate fi calculată frecvența bolii.
- × modul de culegere “caz-martor” corespunde pentru studiul îmbolnăvirilor cu cazuri rare

EXEMPLU



CAZ-MARTOR

× Dezavantaje:

- + pot să apară aprecieri subiective atât din partea celui care culege datele (experimentator) cât și a subiectului urmărit (bolnavul)
- + apare riscul mare al erorilor sistematice de culegere

