The background of the slide is composed of several overlapping geometric shapes in various colors: red, orange, yellow, green, blue, and purple. These shapes create a dynamic, abstract pattern on the left side of the slide.

Studying medicine is
like true love..
You suffer, fail, learn,
fail again, fight and grow,
but you never want to let go.

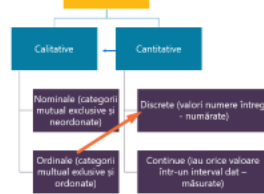
STATISTICĂ DESCRIPTIVĂ VARIABILE CALITATIVE

Sorana D. Bolboacă

Obiective educaționale

- Parametrii statistici de sumarizare a datelor calitative
- Sumarizarea tabelară: tabele de frecvență și contingență
- Sumarizarea grafică:
 - o variabilă
 - 2 variabile

Tipuri de variabile



Frecvența

Date brute → Date summarize

A			
1	TipInternare		
2	Programare		
3	Programare		
4	Programare		
5	Programare		
6	Transfer		
7	Programare		
8	Transfer		
9	Programare		
10	Urgenta		

TipInternare	Nr. (frecvență absolută)	% (frecvență relativă)
Programare	1312	77,7
Transfer	88	5,2
Urgentă	288	17,1
Total	1688	

Principii de sumarizare tabelară

Simple: de preferat 2/3 tabele mai mici în loc de unul încărcat	Informative prin ele însele	Abrevieri sau simboluri explicate la subsolul tabelului	Etichete de rând și coloană
Unități de măsură	Titlu: Ce? Când? Cum? Unde?	Linii și/sau coloane de sinteză (total)	Dacă datele nu sunt ale dvs. trebuie să se menționeze sursa lor într-o notă de subsol

Principii de sumarizare grafică



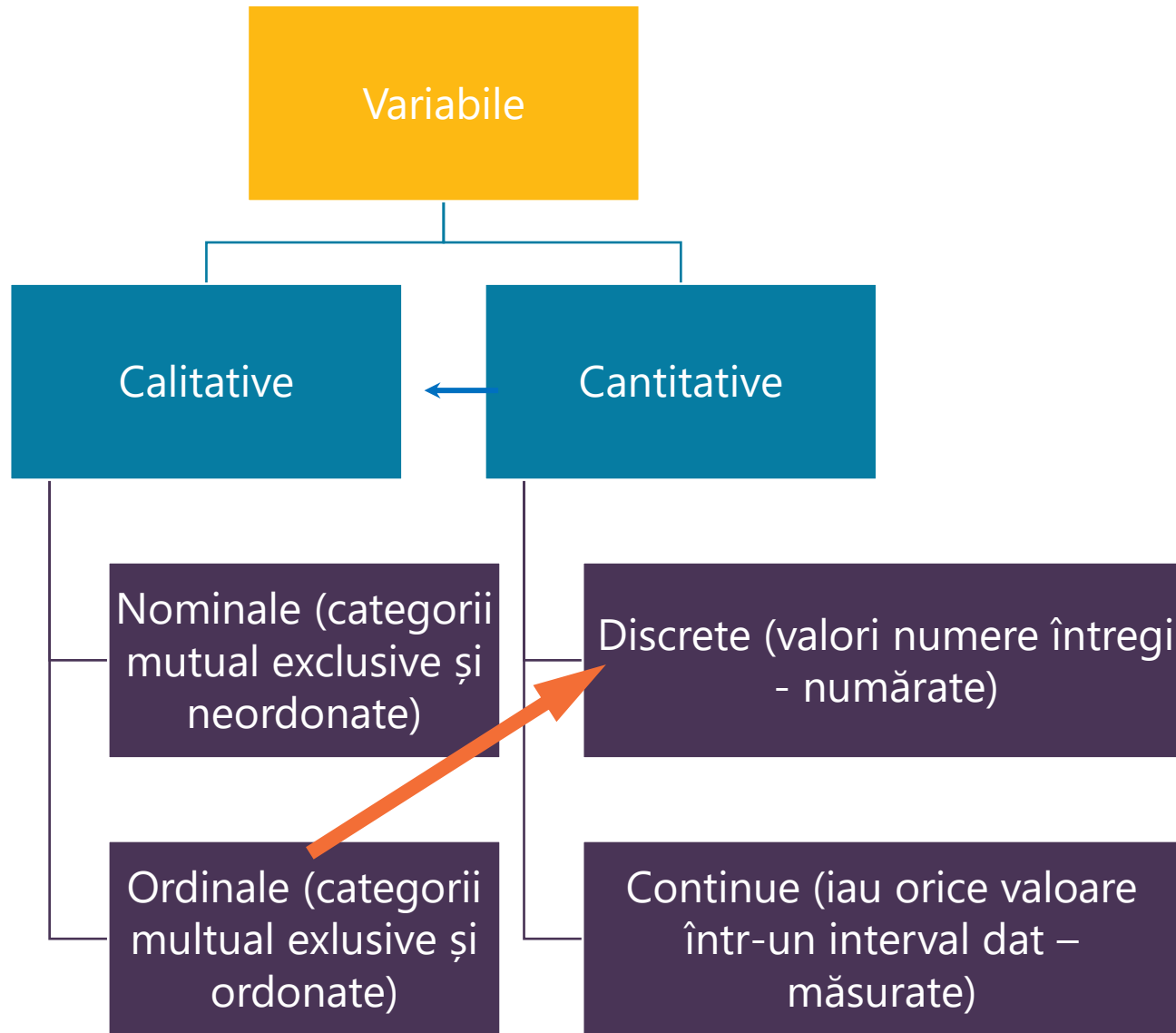
De reținut!

Variablelor calitative se sumarizează ca frecvențe absolute și relative	Raportăm % doar dacă n>30. În caz contrar folosim raportul (ex. 3/8)	Pentru variabilele dicotomice este suficientă raportarea sumară a unei valori posibile.	Tabelul și graficul trebuie să se explice prin ele însele
Nu există rețetă universală pentru reprezentarea tabelară / grafică.	Pe graficul sectorial punem o singură variabilă	Alegem cea reprezentare grafică care transmite cel mai corect informație	...

Obiective educaționale

- Parametrii statistici de sumarizare a datelor calitative
- Sumarizarea tabelară: tabele de frecvență și contingență
- Sumarizarea grafică:
 - o variabilă
 - 2 variabile





Frecvența

Date brute → Date sumarizate

	A
1	TipInternare
2	Programare
3	Programare
4	Programare
5	Programare
6	Transfer
7	Programare
8	Transfer
9	Programare
10	Urgenta

TipInternare	Nr. (frecvență absolută)	% (frecvență relativă)
Programare	1312	77,7
Transfer	88	5,2
Urgență	288	17,1
Total	1688	

Proporția

$$f = a:(a+b) \text{ sau } a/(a+b)$$

$$f \in (0, 1) \text{ sau } f \in (0\%, 100\%)$$

Exemple: Prevalența, Incidența, Se, Sp, VPP, VPN, A (indicatori de evaluarea a testelor diagnostice)

- Prevalența
- Incidența

La serviciul de urgență al unui spital județean au fost consultați 1200 într-o anumită perioadă de timp. 420 au fost internați, 200 fiind de gen feminin. Care este proporția subiecților de sex masculin dintre subiecții spitalizați?

$$220/420 * 100 = 52\%$$

Proporția

	Cancer mamar +	Cancer mamar -	Total
Alăptare la sân timp de 1 an +	70	130	200
Alăptare la sân timp de 1 an -	240	260	500
Total	310	390	700

Proporția femeilor care au alăptat la sân timp de 1 an în grupul celor indemni de cancer mamar = $130/390 = 33,34\%$

Proporția femeilor care au alăptat la sân timp de 1 an în grupul celor cu cancer mamar = $70/310 = 22,58\%$

Proporția femeilor cu cancer mamar în grupul celor care au alăptat timp de 1 an = $70/200 = 35,00\%$

Proporția femeilor cu cancer mamar în grupul celor care nu au alăptat timp de 1 an = $240/500 = 48,00\%$



Rata

$[0, \infty]$

- Riscul relativ
- Rata șansei

Rata = riscul de apariție a unui eveniment în timp (ex. secundă/minut/oră/zi/săptămână/lună/an).

Ex.: rata de morbiditate, rata de atac, rata de mortalitate, rata de natalitate, riscul atribuabil

	Cancer mamar+	Cancer mamar-	Total
Alăptare = Da	20	80	100
Alăptare =Nu	21	79	100
Total	41	159	200

- Riscul de cancer mamar la femeile care au alăptat = $20/100 = 0,20$
- Riscul de cancer mamar la femeile care nu au alăptat = $21/100 = 0,21$
- Riscul relativ = $0,20/0,21 = 0,95 \rightarrow$ Femeile care au alăptat au un risc de a dezvolta cancer mamar cu 5% mai mic comparativ cu cele care nu au alăptat.

Raportul

Se aplică doar în cazul numerelor raționale pozitive a și b unde $b \neq 0$
 $a:b$ sau a/b

- Două din 10 persoane consultate într-o zi de un medic de familie au presiunea arterială sistolică (PAS) mai mare decât valoarea normală. Care este valoarea raportului (PAS normal)/(PAS patologic)?

$$\text{(PAS normal)/(PAS patologic)} = 8/2 = 4$$

→ Avem un subiect cu valori patologice ale PAS la fiecare 4 indivizi cu valori PAS normale



Raportul

	Cancer mamar +	Cancer mamar -	Total
Alăptare la sân timp de 1 an +	70	130	200
Alăptare la sân timp de 1 an -	240	260	500
Total	310	390	700

- Raportul celor cu și fără cancer mamar = $310/390 = 0,79$
- Raportul celor fără și cu cancer mamar = $390/310 = 1,26$
- Raportul dintre cei care au alăptat la sân timp de un an și cei care nu au alăptat la sân timp de 1 an = $200/500 = 0,40$
- Raportul dintre cei care nu au alăptat la sân timp de un an și cei care au alăptat la sân timp de 1 an = $500/200 = 2,50$

Principii de sumarizare tabelară

Simple: de preferat
2/3 tabele mai mici
în loc de unul
încărcat

Informative prin ele
însele

Abrevieri sau
simboluri explicate
la subsolul tabelului

Etichete de rând și
coloană

Unități de măsură

Titlul: Ce? Când?
Cum? Unde?

Linii și/sau coloane
de sinteză (total)

Dacă datele nu sunt
ale dvs. trebuie să se
menționeze sursa
lor într-o notă de
subsol



1 variabilă calitativă

Tabelul de frecvență

Se ordonează datele/clasele crescător (variabilă cantitativă)

Se determină frecvența fiecărei valori/clase

Se includ valorile distincte și frecvențele într-un tabel pe diferite coloane:

- Frecvența absolută (numărul de cazuri care îndeplinesc criteriul)
- Frecvența relativă = raportul dintre frecvența absolută și volumul eșantionului/populației (simbol = n/N). Valorile se pot prezenta și procentual.
- Frecvențe absolute/relative cumulate crescător/descrescător

COUNTIF(array,criteria)

COUNTIFS(array1,criteria1, array2,criteria2, etc.)

PIVOT TABLE

Tabelul de contingență

PIVOT TABLE

Date brute

	File	Home	Insert	Page Lay
	PivotTable	Recommended	Table	Pictu
		PivotTables		
	Tables			
	K21			
	A	B	C	D
1	sex	dislipindemie	alcool	fumat
2	F	nu	nu	nu
3	M	da	nu	nu
4	M	nu	nu	nu
5	F	nu	nu	nu
6	M	nu	nu	nu
7	M	nu	nu	nu
8	M	nu	nu	da
9	M	nu	nu	nu
10	M	nu	da	da

Tabel de contingență 2×2

3	Count of dislipindemie	dislipindemie		
4	sex	da	nu	Grand Total
5	F	1	26	27
6	M	2	33	35
7	Grand Total	3	59	62

Tabelul de frecvență: 1 variabilă calitativă

Suma frecvențelor absolute ale tuturor valorilor seriei care sunt mai mici sau egale cu x

Scor durere	Nr. cazuri	%	Frecvența absolută cumulată crescător	Frecvența absolută cumulată descrescător
0	0	0	=0	=100-0=100
1	0	0	=0+0	=100-0=100
2	0	0	=0+0	=100-0=100
3	0	0	=0+0	=100-0=100
4	0	0	=0+0	=100-0=100
5	1	1	=1+0=1	=100-1=99
6	7	7	=7+1=8	=99-7=92
7	14	14	=14+8=22	=92-14=78
8	29	29	=29+22=51	=78-29=49
9	29	29	=29+51=80	=49-29=20
10	20	20	=20+80=100	=20-20=0
Total	100	100		

Tabelul de contingență: 2/x variabile

2×2

2 variabile calitative dicotomiale:
genul (F/M) & TBC (+/-)

	TBC+	TBC-	Total
Feminin	2	10	12
Masculin	24	54	78
Total	26	64	90

r×c

r = numărul de rânduri
c = numărul de coloane

6×4

	AB	BH	BN	CJ
Hepatită A	166	171	16	50
Hepatită B	13	14	9	9
Hepatită C	1	25	4	7
Alte tipuri de Hepatită	0	8	6	0
Hepatită cronică	0	0	12	9
Purtători cronici de HBsAg	21	53	14	2

2 variabile calitative nominale:
Tipul hepatiei (A / B / ...) &
Reședința (județul: AB / BH / ...)

Date brute

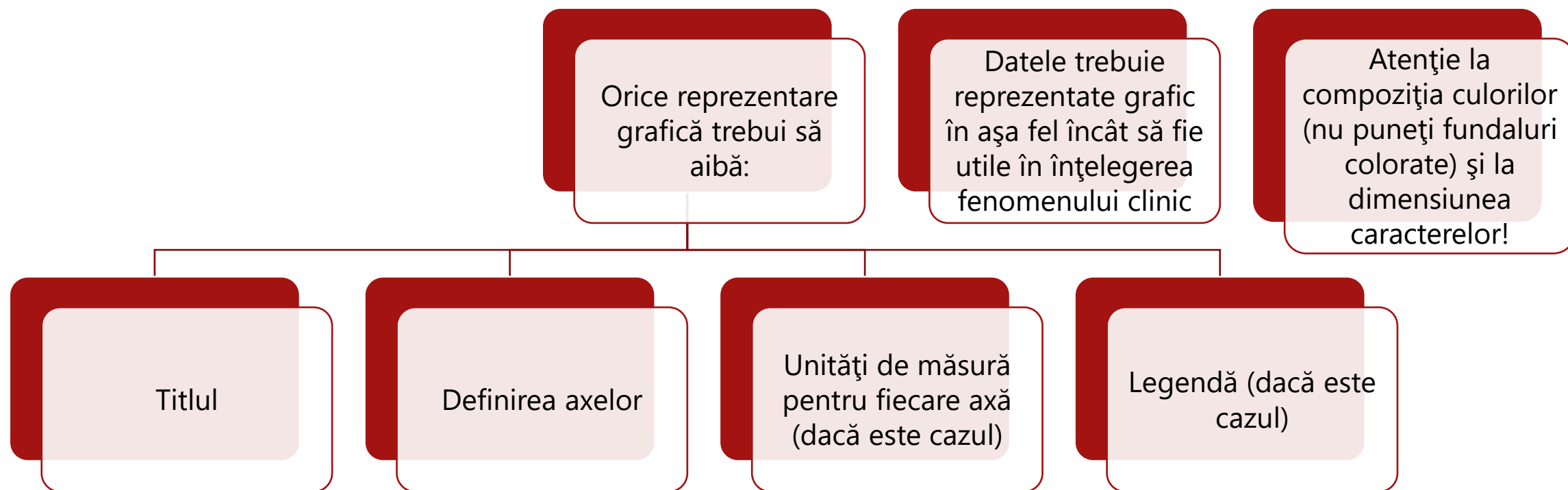
ID	TipHepatită	Județ
0001	A	AB
0002	B	AB



REPREZENTARE TABELARĂ

- Informații concise, afișare eficientă.
- Permit prezentarea nivelului de detaliu și de precizie dorit.
- Numerotați tabelele consecutiv în ordinea în care sunt citate în text.
- Fiecare coloană și rând trebuie să aibă o denumire și să prezinte unitățile de măsură.
- Dacă utilizați abrevieri pentru denumirea coloanelor sau a rândurilor prezentați la subsolul tabelului definiția abrevierilor.
- Dacă utilizați date care nu sunt ale dvs., fie că aceste date sunt publicate sau nu, trebuie să obțineți permisiunea de folosire și să specificați sursa acestora.

Principii de sumarizare grafică



Tipuri de grafice

Plăcintă (Pie)/Gogoașa (*Sunburst*)

O variabilă calitativă
(nominală sau ordinală)-
Pie ...

Coloane/Bare

1/k variabile calitative
(nominală sau ordinală)

Linie (*Line*) / Arie (*Area*)

1/k variabile calitative &
modificări în timp



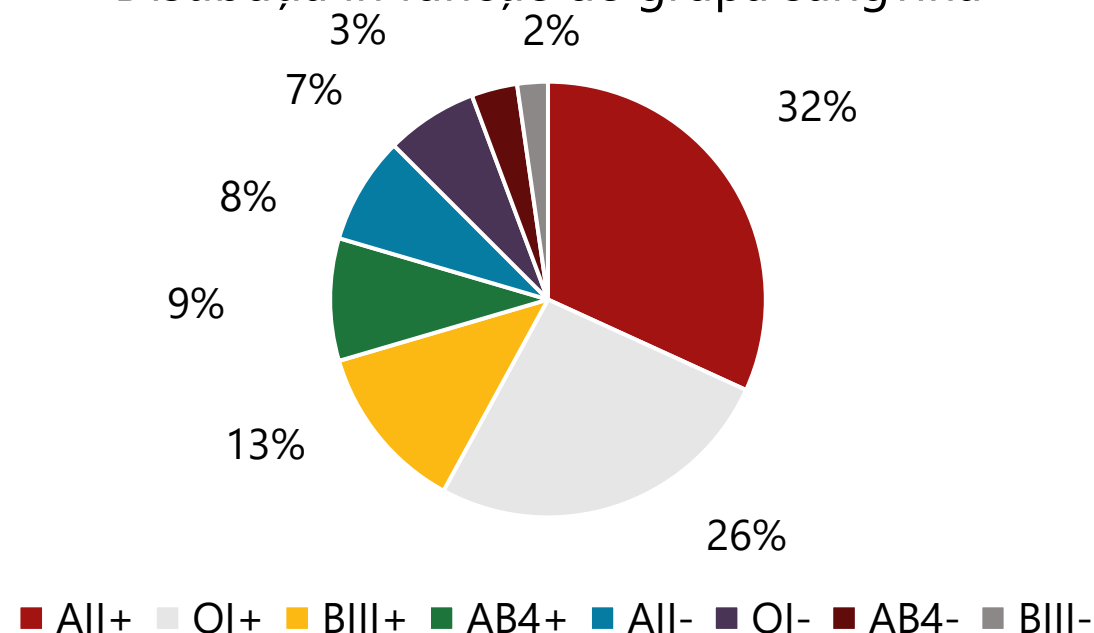
Graficul sectorial

E necesară sumarizarea tabelară înainte de realizarea graficului

Date sumarizate -Tabel de frecvență

AM	AN	AO	AP
Grupa sangvină		Row Labels	Count of Grupa sangvină
All+		AB4-	3
All+		AB4+	8
All-		All-	7
All+		All+	28
All+		BIII-	2
All+		BIII+	11
All+		OI-	6
AB4-		OI+	23
All+		Grand Total	88
BIII+			

Distribuția în funcție de grupa sangvină



Graficul sectorial

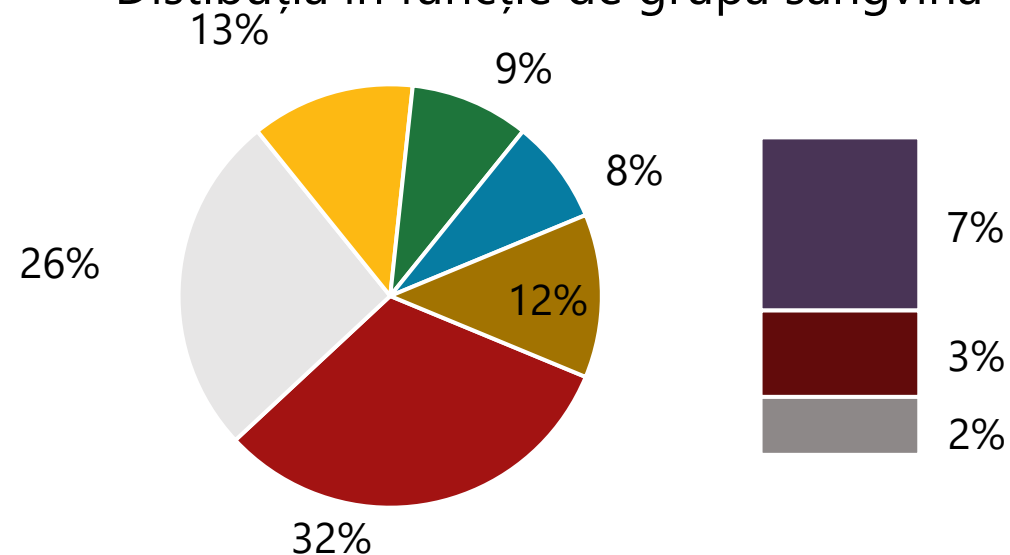
E necesară sumarizarea tabelară
înainte de realizarea graficului

Date sumarizate -Tabel de frecvență

AM
Grupa sangvină
All+
All+
All-
All+
All+
All+
All+
AB4-
All+
BIII+

Grupa sangvină	nr. subiecți
All+	28
OI+	23
BIII+	11
AB4+	8
All-	7
OI-	6
AB4-	3
BIII-	2

Distibuția în funcție de grupa sangvină



■ All+ ■ OI+ ■ BIII+ ■ AB4+ ■ All- ■ OI- ■ AB4- ■ BIII-

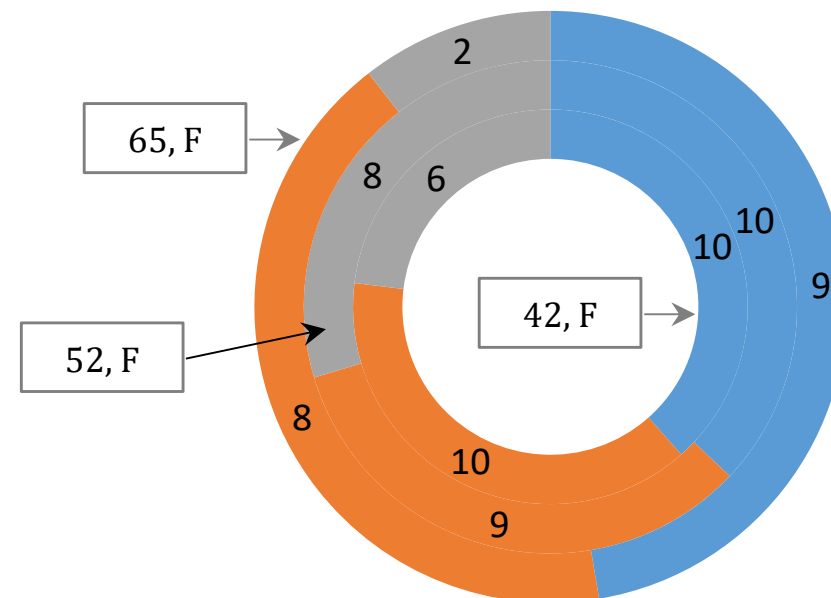
Gogoașa

Date brute

- Efectul fizioterapiei asupra durerii la pacienții cu artroză. Scorul de durere măsurat pe scala vizuală (0 = fără durere, 10 = durere extrem de puternică)

Vârsta (ani)	Genul	Scor durere		
		Inițial	Post-FT1	Post-FT2
42	F	10	10	6
52	F	10	9	8
65	F	9	8	2

Scor durere în evoluție

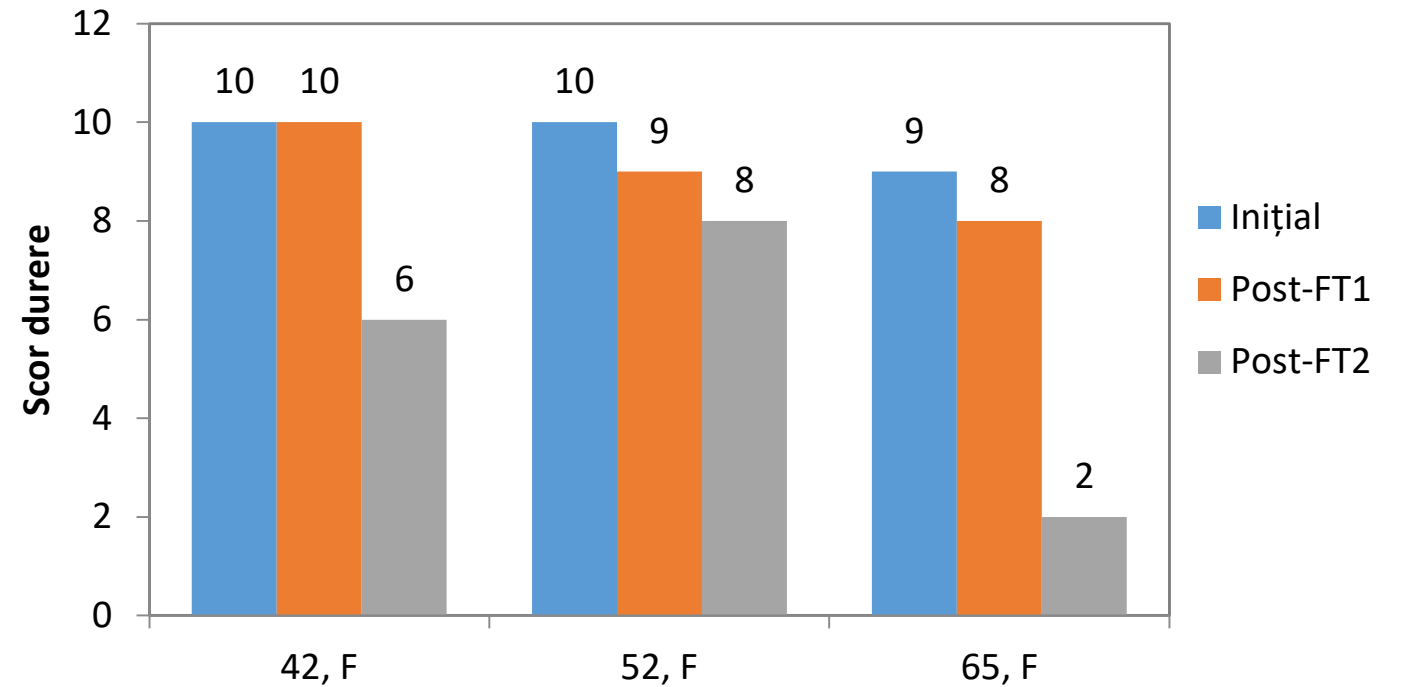


■ Inițial ■ Post-FT1 ■ Post-FT2

Bare/Coloane

Vârsta (ani)	Genul	Scor durere		
		Inițial	Post-FT1	Post-FT2
42	F	10	10	6
52	F	10	9	8
65	F	9	8	2

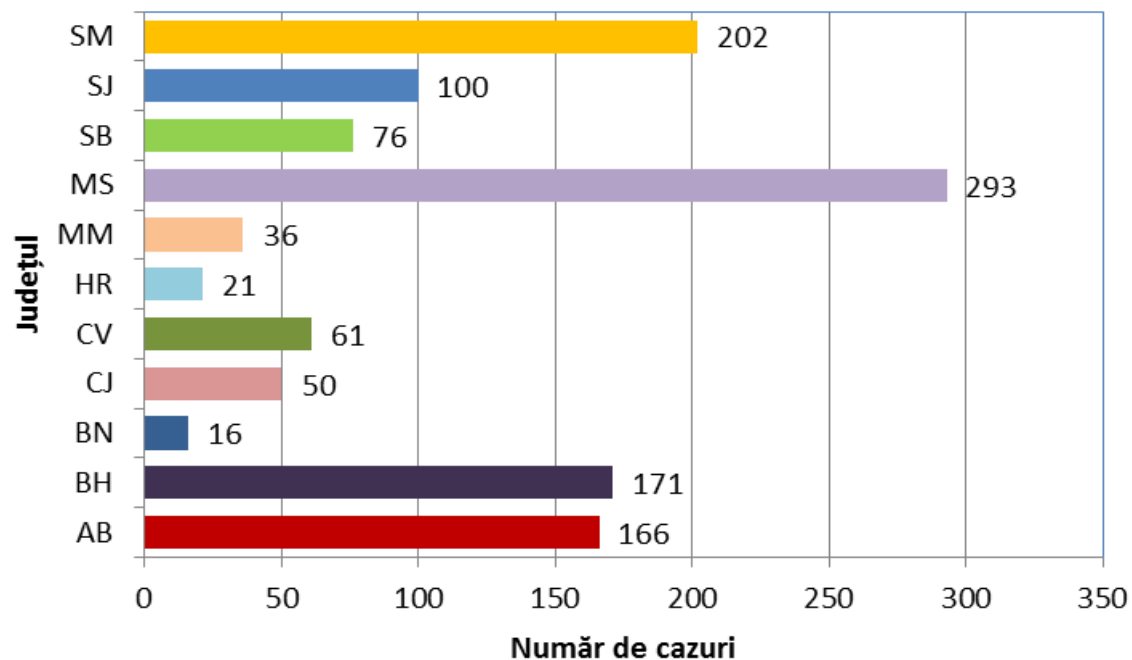
Scor durere în evoluție



Bare/Coloane

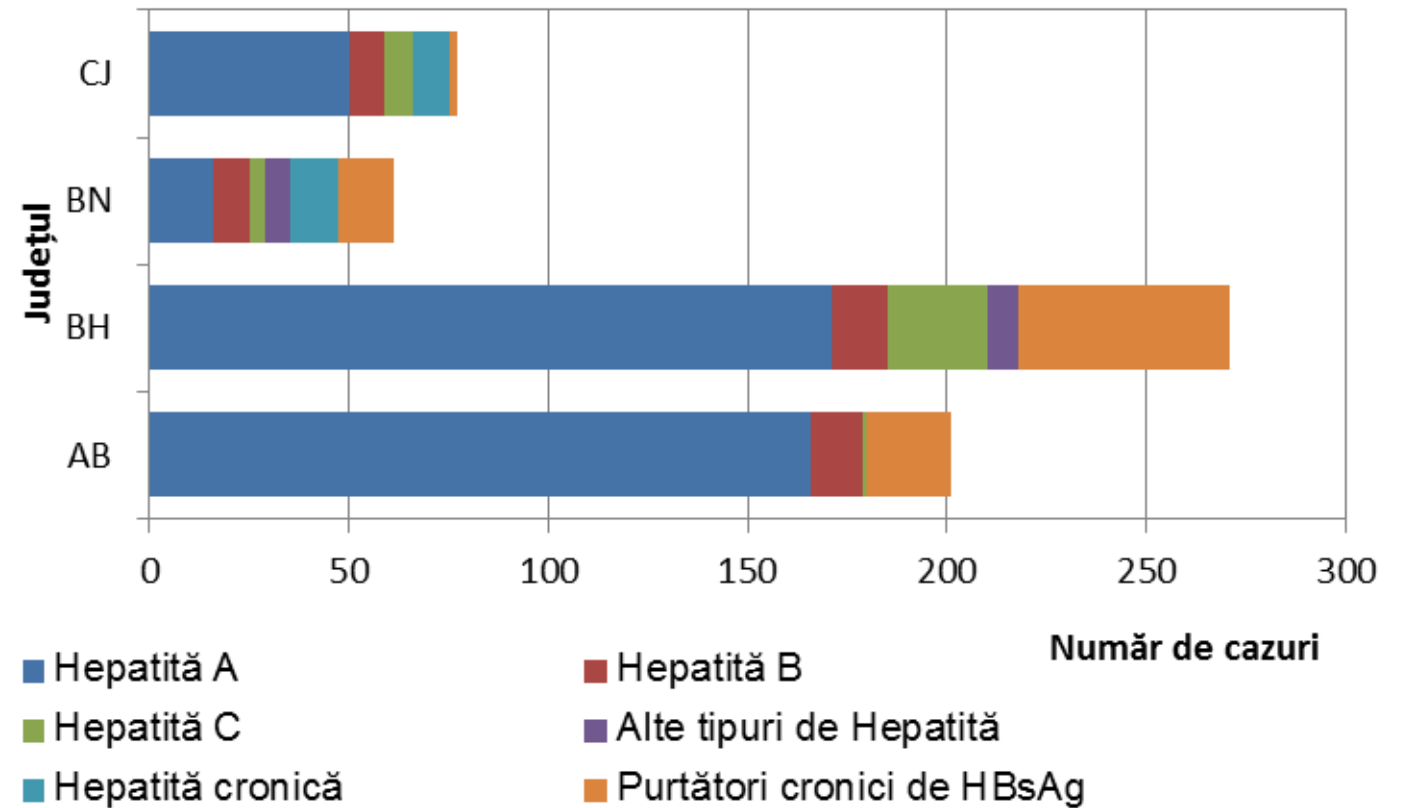
		BH	BN	CJ	CV	HR	MM	MS	SB	SJ	SM
Hepatita A		171	16	50	61	21	36	293	76	100	202

Cazuri de Hepatită A în funcție de domiciliu



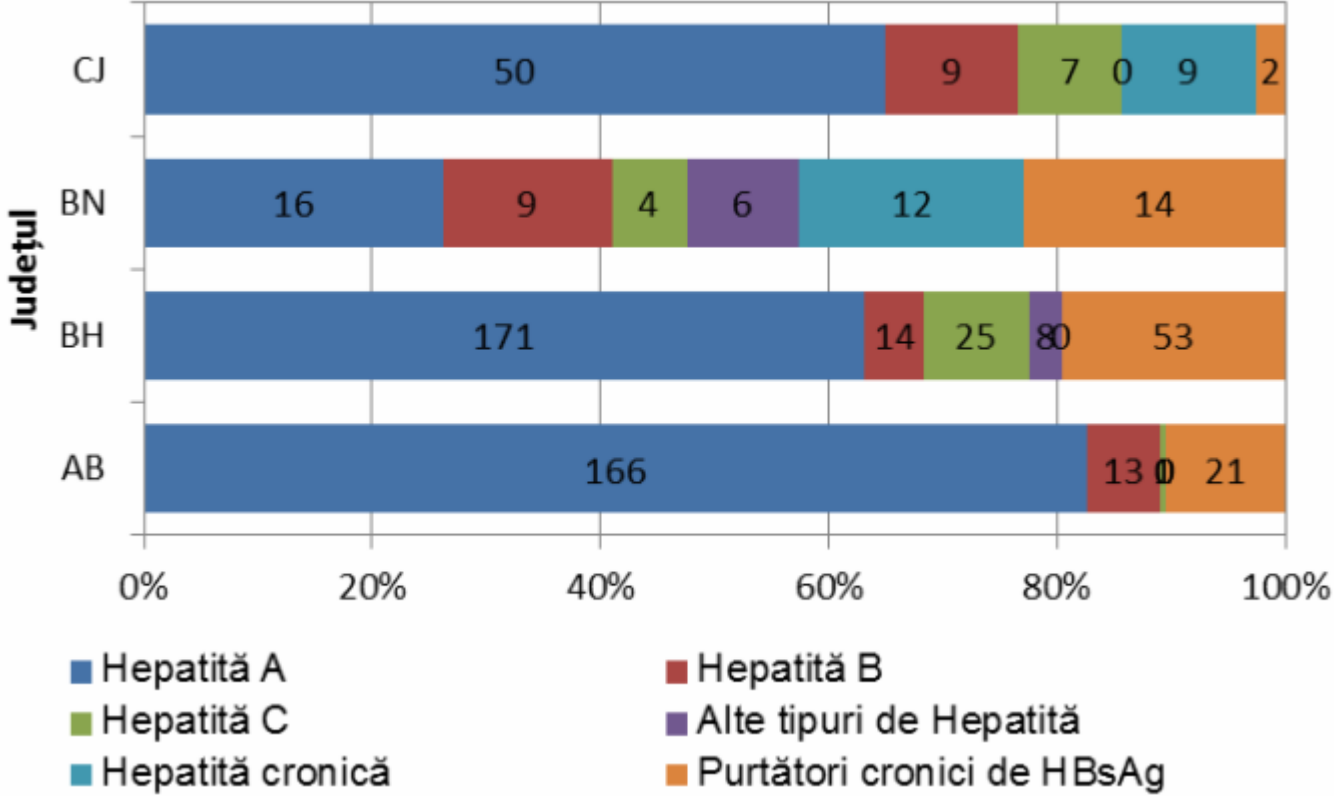
Stacked Bar

	AB	BH	BN	CJ
Hepatită A	166	171	16	50
Hepatită B	13	14	9	9
Hepatită C	1	25	4	7
Alte tipuri de Hepatită	0	8	6	0
Hepatită cronică	0	0	12	9
Purtători cronici de HBsAg	21	53	14	2



100% Stacked Bar

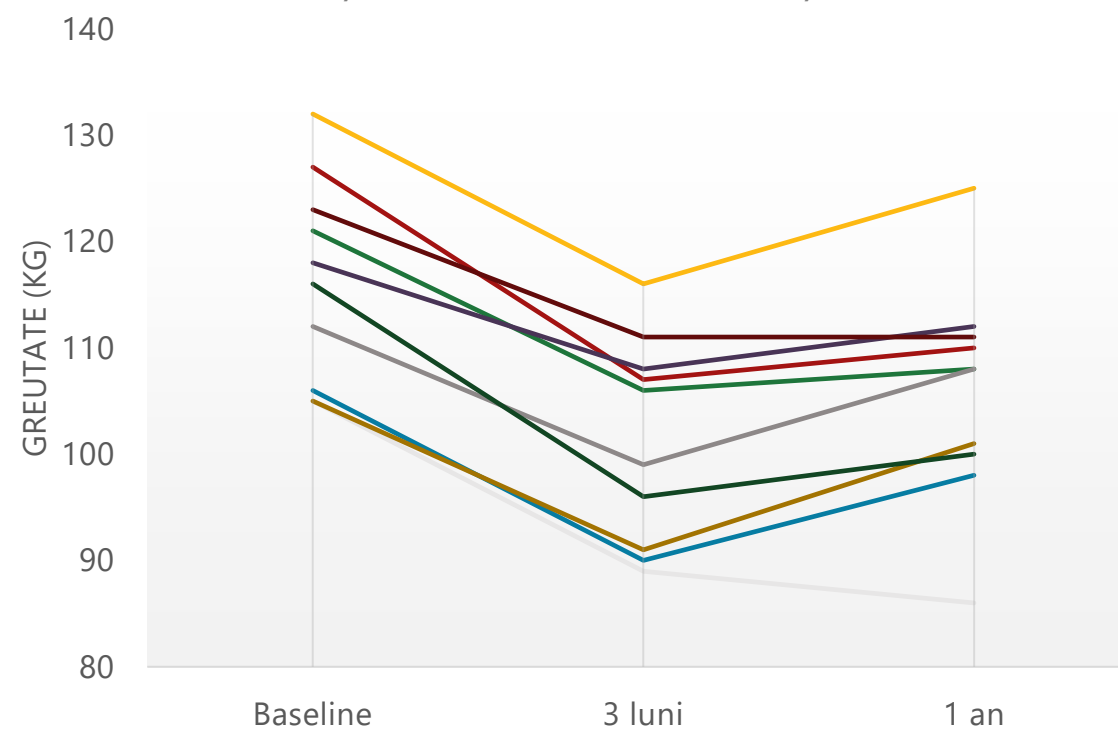
	BH	BN	CJ
Hepatită A	171	16	50
Hepatită B	14	9	9
Hepatită C	25	4	7
Alte tipuri de Hepatită	8	6	0
Hepatită cronică	0	12	9
Purtători cronici de HBsAg	53	14	2



Linie

ID	Baseline	3 luni	6 luni
1	132	116	125
2	127	108	111
3	123	111	111
4	121	106	108
5	118	108	112
6	116	96	100
7	112	99	108
8	106	91	98
9	105	91	101
10	105	89	86

Evoluția greutății post-dietă la un eșantion de 10 subiecți



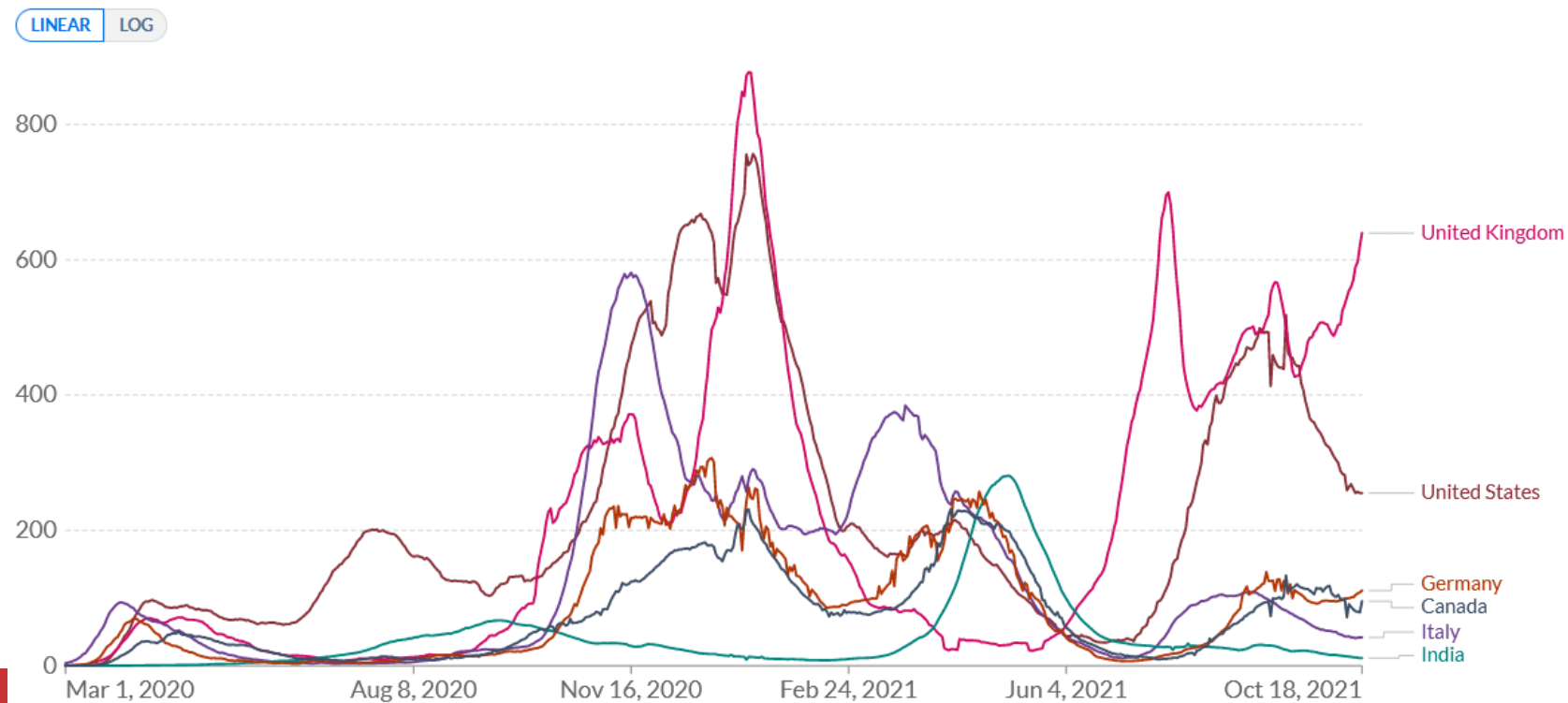
Linie

Date summarize

Daily new confirmed COVID-19 cases per million people

Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases; the main reason for that is limited testing.

Our World
in Data



Source: Johns Hopkins University CSSE COVID-19 Data

CC BY

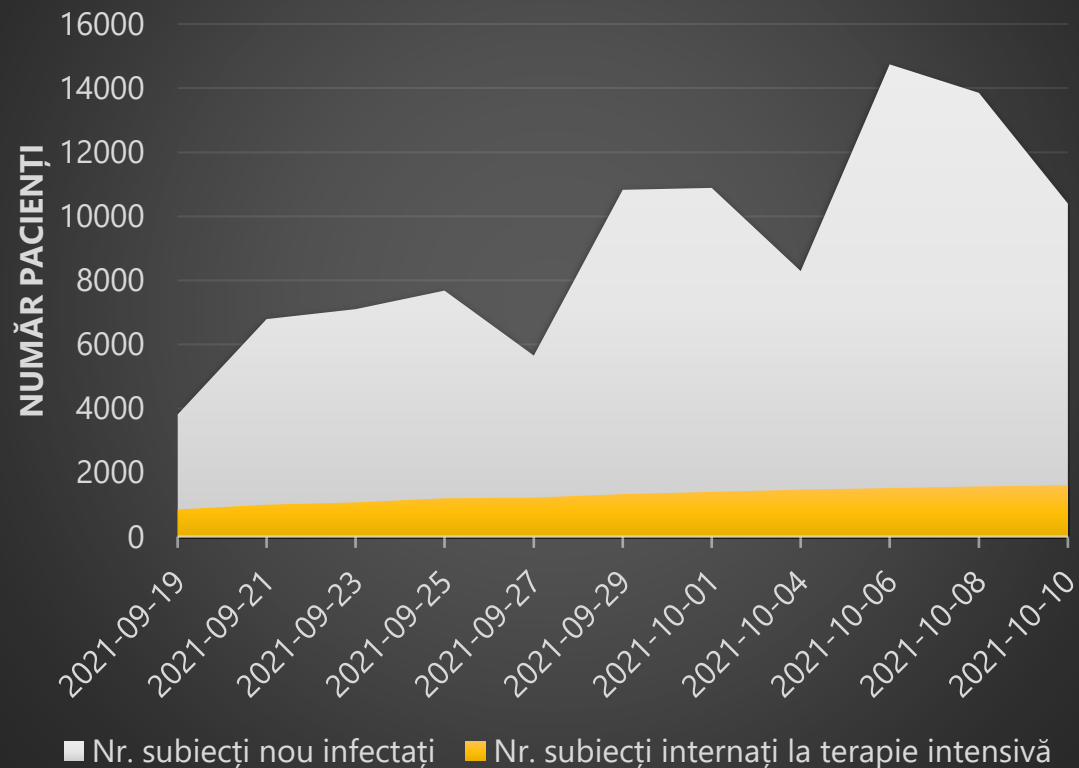
Jan 28, 2020

Oct 18, 2021

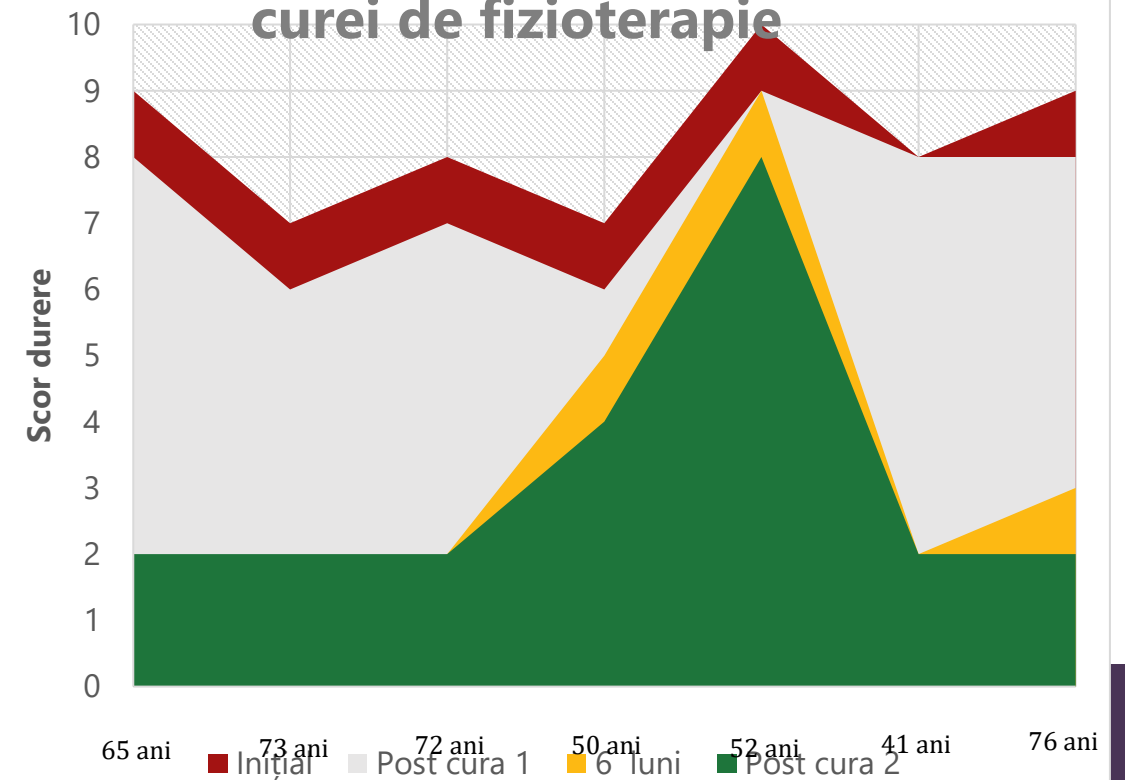
Data	Țara	Cazuri
------	------	--------

Mar 1, 2020		

Evoluția COVID-19 Romania 19/09-18/10/2021



Evoluția durerii în contextul curei de fizioterapie





REPREZENTARE GRAFICĂ

- Reprezentarea grafică trebuie să fie auto-explicativă
- Titlul și detaliile cu privire la reprezentarea grafică aparțin legendei – nu le includeți în figură.
- Figurile trebuie numerotate consecutiv în ordinea citării în text.
- Dacă figura a fost publicată anterior specificați sursa de unde a-ti luat-o și obțineți de la autori sau de la casa de editură acordul scris de acceptare al utilizării.
- Explicați clar în legendă semnificația simbolurilor utilizate precum și a altor semne (de exemplu săgeți, numere, litere, etc.)
- Evitați reprezentările 3D!

De reținut!

Variabilelor calitative se sumarizează ca frecvențe absolute și relative

Raportăm % doar dacă $n > 30$, în caz contrar folosim raportul (ex. 3/8)

Pentru variabilele dicotomiale este suficientă raportarea sumară a unei valori posibile.

Tabelul și graficul trebuie să se explice prin ele însele

Nu există rețeta universală pentru reprezentarea tabelară / grafică

Pe graficul sectorial punem o singură variabilă

Alegem cea mai bună reprezentare grafică care transmite cel mai corect informație

...