



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

TESTE NEPARAMETRICE TESTE HI-PĂTRAT

Sorana D. Bolboacă

Variabile de interes vs. test statistic

Calitative

Nominale/ordinale

- Frecvența: testul Hi-pătrat, testul Hi-pătrat corectat, testul exact al lui Fisher, Testul Z, Testul McNemar (grupuri perechi)

Calitative

Ordinale (> 5 clase; $p < 0,05$ – distribuție) + 1 variabilă calitativă de grupare

- 2 grupuri → independente - Testul Mann-Whitney; dependente - Testul Wilcoxon
- > 2 grupuri → independente - Testul Kruskal-Wallis; dependente - Testul Friedman

Cantitative

Discrete sau continue ($p < 0,05$ – distribuție)

- 2 grupuri → independente - Testul Mann-Whitney; dependente - Testul Wilcoxon
- > 2 grupuri → independente - Testul Kruskal-Wallis; dependente - Testul Friedman

Obiective educaționale

Teste pe frecvențe

- Testul Hi-pătrat
- Testul Hi-pătrat cu corecția Yates
- Testul exact al lui Fisher
- Testul McNemar
- Testul Z



Parametrul = frecvența (π)

Testul Hi-pătrat

PHANTOMS:

Testul Hi-pătrat (χ^2) – Chi-Squared

Testează cât de diferite sunt datele observate comparativ cu cele așteptate/teoretice
Evaluează asocierea dintre două variabile calitative

Abb	Ce?
P	Asocierea (două variabile calitative) / Frecvența
H	<div><div><div>Ipoteze (cu privire la parametrul populației) H_0: X și Y sunt independente. H_1: X și Y sunt dependente.</div><div>$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{(f_i^o - f_i^t)^2}{f_i^t}$</div></div><div>$\chi^2$ = statistica χ^2 f_i^o = frecvența observată f_i^t = frecvența teoretică r = numărul de rânduri c = numărul de coloane $df = (r-1) \cdot (c-1)$, df (grade de libertate)</div></div>
A	Condiții de aplicare: <ul style="list-style-type: none">○ <i>Independența</i>: observațiile din eșantion sunt independente (eșantion randomizat / asignare randomizată; eșantionare fără înlocuire $n < 10\% \cdot N$)○ Fiecare celulă din tabelul teoretic trebuie să conțină cel puțin 5 observații.
N	Denumirea testului: testul Hi-pătrat (Chi-squared test)
T	Statistica testului: χ^2 (regiunea de respingere = $[3.84; \infty)$ pentru tabelul de contingență de 2×2)
O	Valoarea p (nivel de semnificație α)
M	Decizia (respingem / nu avem suficiente dovezi să respingem H_0 – utilizăm statistica χ^2 vs. regiunea de respingere / valoarea p vs. α)
S	Concluzia testului în contextul problemei.

PHANTOMS:

Testul Hi-pătrat (χ^2) – Chi-Squared

H_0 : Obezitatea și sexul sunt **in**dependente (Obezitatea **nu** variază în funcție de sex).

H_1 : Obezitatea și sexul sunt dependente (Obezitatea variază în funcție de sex).

Table 2 Epidemiological, clinical and therapeutic differences between genders (mean \pm SD) n (%)

	Females ($n = 542$)	Males ($n = 486$)	P value
Obesity	224 (41.4)	160 (32.9)	0.005



Observat	Obezitate +	Obezitate -	Total
Feminin	224	318	542
Masculin	160	326	486
Total	384	644	1028

Teoretic	Obezitate +	Obezitate -	Total
Feminin	$=542 \cdot 384 / 1028$ 203	$=542 \cdot 644 / 1028$ 340	542
Masculin	$=486 \cdot 384 / 1028$ 182	$=486 \cdot 644 / 1028$ 305	486
Total	384	644	1028

PHANTOMS:

Testul Hi-pătrat (χ^2)-Chi-Squared

Observat	Obezitate +	Obezitate -	Total
Feminin	224	318	542
Masculin	160	326	486
Total	384	644	1028

Teoretic	Obezitate +	Obezitate -	Total
Feminin	$=542 \cdot 384 / 1028$ 203	$=542 \cdot 644 / 1028$ 340	542
Masculin	$=486 \cdot 384 / 1028$ 182	$=486 \cdot 644 / 1028$ 305	486
Total	384	644	1028

Abb	Asocierea dintre obezitate și gen.
P	Asocierea: 2 variabile calitative dicotomiale (sex: feminin/masculin & obezitate: da/nu)
H	H_0 : Obezitatea și sexul sunt in dependente (Obezitatea nu variază în funcție de gen). H_1 : Obezitatea și sexul sunt dependente (Obezitatea variază în funcție de gen).
A	Condiții de aplicare: Independența (observațiile din eșantion sunt independente) & Fiecare celulă din tabelul teoretic trebuie să conțină cel puțin 5 observații.
N	Denumirea testului: testul Hi-pătrat (Chi-squared test)
T	$\chi^2 = (224-203)^2/203 + (318-340)^2/340 + (160-182)^2/182 + (326-305)^2/305 = 7,7$
O	Valoarea P = 0,0055
M	$\chi^2 = 7,7$ aparține regiunii de respingere $[3,84; \infty) \rightarrow$ avem suficiente dovezi să respingem H_0 $P < 0,05 \rightarrow$ avem suficiente dovezi să respingem H_0
S	Obezitatea și sexul sunt dependente /Obezitatea variază în funcție de sex.

Putem însă să tragem concluzia că obezitatea este determinată de gen?

Cât de mare e riscul obezității?

Observat	Obezitate +	Obezitate -	Total
Feminin	224	318	542
Masculin	160	326	486
Total	384	644	1028

- Riscul obezității la femei: $224/(224+318) = 0,41$ IC95% [0,37 – 0,46]
 - Riscul obezității la bărbați: $160/(160+326) = 0,33$ IC95% [0,29 – 0,37]
 - Rata riscului (RR): $0.41/0.33 = 1.26$ (IC95% [1.07 – 1.48])
- riscul de obezitate la femei e de 1.26 mai mare față de riscul de obezitate la bărbați.
- în populație mă aștept ca acest risc să fie între 1.07 și 1.48
- Deoarece valoare 1 (RR=1 → am același risc de obezitate și la femei și la bărbați) nu aparține intervalului de încredere riscul e semnificativ statistic.



PHANTOMS:

Corecția Yates

Observat	Stângani	Dreptaci	Total
Feminin	1	44	45
Masculin	9	43	52
Total	10	87	97

Teoretic	Stângani	Dreptaci	Total
Feminin	$=45 \cdot 10 / 97$ 4,64	$=45 \cdot 87 / 97$ 40,36	45
Masculin	$=52 \cdot 10 / 97$ 5,36	$=52 \cdot 87 / 97$ 46,64	52
Total	10	87	97

Asocierea dintre mâna dominantă și gen

P	Asocierea: 2 variabile calitative dicotomiale (sex: feminin/masculin & dominanța: stângaci/dreptaci)
H	H_0 : Mâna dominantă nu variază cu sex / Mâna dominantă și sexul sunt independente. H_1 : Mâna dominantă variază cu sex / Mâna dominantă și sexul sunt dependente.
A	Condiții de aplicare: Independența (observațiile din eșantion sunt independente) & Toate frecvențele așteptate sunt mai mari de 1 și 80% din celulele așteptate sunt > 5 (3 din 4).
N	Denumirea testului: testul Hi-pătrat cu corecția Yate (<i>Chi-squared test with Yate's correction</i>)
T	$\chi^2 = (\text{abs}(1-4.64)-0.5)^2/4.64 + (\text{abs}(44-40.36)-0.5)^2/40.36 + (\text{abs}(9-5.36)-0.5)^2/5.36 + (\text{abs}(43-46.64)-0.5)^2/46.64 = 4.42$
O	Valoarea P (CHISQ.TEST) = 0.0148
M	$\chi^2 = 4.48$ aparține regiunii de respingere $[3.84; \infty) \rightarrow$ avem suficiente dovezi să respingem H_0 $P = 0.0148 < 0.05 \rightarrow$ avem suficiente dovezi să respingem H_0
S	Mâna dominantă și sexul sunt dependente.

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{(|f_i^o - f_i^t| - 0,5)^2}{f_i^t}$$

Testul exact al lui Fisher

PHANTOMS:

Testul Fisher

Observat	SDRA focal	SDRA non-focal	Total
Astm+	1	3	4
Astm-	31	84	115
Total	32	87	119

Teoretic	SDRA focal	SDRA non-focal	Total
Astm+	1,08	2,92	4
Astm-	30,92	84,04	115
Total	32	87	119

Asocierea dintre astm și morfologia pulmonară CT în sindromul de detresie respiratorie acută (SDRA)

P	Parametrii: Morfologie pulmonară CT (SDRA focal / non-focal) & Astm (prezent (+)/ absent (-))
H	π_1 = proporția subiecților cu astm și SDRA focal, π_2 = proporția subiecților cu astm și SDRA non-focal H_0 : $\pi_1 = \pi_2$ (Astmul nu este asociat cu morfologia pulmonară în SDRA) H_1 : $\pi_1 \neq \pi_2$ (Astmul este asociat cu morfologia pulmonară în SDRA)
A	Condiții de aplicare: Independența (observațiile din eșantion sunt independente) & Frecvența așteptată/expectată/teoretică < 2; > 20% din frecvențele așteptată/expectată/teoretică < 5
N	Denumirea testului: testul exact al lui Fisher
T	Nu există
O	Suma dintre valoarea p calculată pe tabelul observat și respectiv valoarea p calculată pe tabelul teoretic. (Suma valorilor p calculate pe toate tabelele posibile respectând suma de pe rânduri și suma de pe coloane).
M	$\underline{p} > 0.9999 \rightarrow$ Nu avem suficiente evidențe să respingem H_0
S	Astmul nu este asociat cu morfologia pulmonară în SDRA.

PHANTOMS:

Testul Fisher

$$p = \frac{(AP + FN)!(FP + AN)!(FN + AN)!(AP + FP)!}{n!AP!FP!FN!AN!}$$

$$p = \frac{(12 + 3)! * (3 + 14)! * (3 + 14)! * (12 + 3)!}{32! * 12! * 3! * 3! * 14!}$$

Observat	Magneziu	Placebo	Total
Răspuns+	12	3	15
Răspuns-	3	14	17
Total	15	17	32

(i) 0 15 15 2	(ii) 1 14 14 3	(iii) 2 13 13 4
(iv) 3 12 12 5	(v) 4 11 11 6	(vi) 5 10 10 7
(vii) 6 9 9 8	(viii) 7 8 8 9	(ix) 8 7 7 10
(x) 9 6 6 11	(xi) 10 5 5 12	(xii) 11 4 4 13
(xiii) 12 3 3 14	(xiv) 13 2 2 15	(xv) 14 1 1 16
(xvi) 15 0 0 17	Illustration of all the different ways of rearranging cell frequencies in table 1, but with the marginal totals remaining the same.	

TABLE 3					
Total	a	b	c	d	P-value
i	0	15	15	2	0.0000002
ii	1	14	14	3	0.0000180
iii	2	13	13	4	0.0004417
iv	3	12	12	5	0.0049769
v	4	11	11	6	0.0298613
vi	5	10	10	7	0.1032349
vii	6	9	9	8	0.2150728
viii	7	8	8	9	0.2765221
ix	8	7	7	10	0.2212177
x	9	6	6	11	0.1094916
xi	10	5	5	12	0.0328475
xii	11	4	4	13	0.0057426
xiii	12	3	3	14	0.0005469
xiv	13	2	2	15	0.0000252
xv	14	1	1	16	0.0000005
xvi	15	0	0	17	0.0000000

Total	a	b	c	d	p-value
i	0	15	15	2	0.0000002
ii	1	14	14	3	0.0000180
iii	2	13	13	4	0.0004417
iv	3	12	12	5	0.0049769
v	4	11	11	6	0.0298613
vi	5	10	10	7	0.1032349
vii	6	9	9	8	0.2150728
viii	7	8	8	9	0.2765221
ix	8	7	7	10	0.2212177
x	9	6	6	11	0.1094916
xi	10	5	5	12	0.0328475
xii	11	4	4	13	0.0057426
xiii	12	3	3	14	0.0005469
xiv	13	2	2	15	0.0000252
xv	14	1	1	16	0.0000005
xvi	15	0	0	17	0.0000000
sum					1.0000000

tabel observat

P (tabelul observat sau mai extrem decât cel observat) = P(xiii)+P(xiv)+P(xv)+P(xvi) = 0.0005726

PHANTOMS:

Testul Fisher

$$p = \frac{(AP + FN)!(FP + AN)!(FN + AN)!(AP + FP)!}{n!AP!FP!FN!AN!}$$

$$p = \frac{(12 + 3)! * (3 + 14)! * (3 + 14)! * (12 + 3)!}{32! * 12! * 3! * 3! * 14!}$$

Observat	Magneziu	Placebo	Total
Răspuns+	12	3	15
Răspuns-	3	14	17
Total	15	17	32

	X	Y	Z	AA	AB	AC
55	Total	a	b	c	d	p-value
56	i	0	15	15	2	0.0000002
57	ii	1	14	14	3	0.0000180
58	iii	2	13	13	4	0.0004417
59	iv	3	12	12	5	0.0049769
60	v	4	11	11	6	0.0298613
61	vi	5	10	10	7	0.1032349
62	vii	6	9	9	8	0.2150728
63	viii	7	8	8	9	0.2765221
64	ix	8	7	7	10	0.2212177
65	x	9	6	6	11	0.1094916
66	xi	10	5	5	12	0.0328475
67	xii	11	4	4	13	0.0057426
68	xiii	12	3	3	14	0.0005469
69	xiv	13	2	2	15	0.0000252
70	xv	14	1	1	16	0.0000005
71	xvi	15	0	0	17	0.0000000

$$= (\text{FACT}(Y56 + Z56) * \text{FACT}(AA56 + AB56) * \text{FACT}(Z56 + AB56)) * \text{FACT}(Y56 + AA56) / (\text{FACT}(Y56 + Z56 + AA56 + AB56) * \text{FACT}(Y56) * \text{FACT}(AA56) * \text{FACT}(AB56) * \text{FACT}(Z56))$$

- P (tabelul observat sau mai extrem decât cel observat) = P(xiii) + P(xiv) + P(xv) + P(xvi) = 0.0005726 (test unilateral)
- Valoarea P (test bilateral) = 2 * 0.0005726 = 0.001145211
- Valoarea P (test bilateral) = sum(P ≤ 0.0005726) = P(i) + P(ii) + P(iii) + P(xiii) + P(xiv) + P(xv) + P(xvi) = 0.0000002 + 0.0000180 + 0.0004417 + 0.0005469 + 0.0000252 + 0.0000005 + 0.0000000 = 0.0010325

Testul McNemar

PHANTOMS:

Testul McNemar

Observat	durere2+	durere2-	Total
durere1+	a	b	a+b
durere1-	c	d	c+d
Total	a+c	b+d	a+b+c+d

Abb Efectul unui analgezic asupra durerii articulare

P	Asocierea (două variabile calitative: durere înainte de analgezic (prezentă= durere1+/absentă=durere1-) și durere după analgezic (prezentă= durere2+/absentă=durere2-))		
H	$H_0: \pi_b = \pi_c$ vs. $H_1: \pi_b \neq \pi_c$		
A	Condiții de aplicare a testului: <ul style="list-style-type: none"> ○ Date perechi: înainte/după sau grupuri perechi ○ În același grup datele sunt mutual exclusive 	$\chi^2 = \frac{(b - c)^2}{b + c}$	$OR = \frac{b}{c}$
N	Testul McNemar		
T	$\chi^2 = (b-c)^2/(b+c)$ & $OR = b/c$		
O	Valoarea P (nivel de semnificație 5%)		
M	χ^2 aparține regiunii de respingere $[3.84; \infty) \rightarrow$ avem suficiente dovezi să respingem H_0 Valoarea P ($\alpha=5\%$) $< 0.05 \rightarrow$ avem suficiente dovezi să respingem H_0		
S	Analgezicul investigat are efect benefic asupra durerii?		

PHANTOMS:

Testul McNemar

S-a realizat un studiu pe femei după instalarea menopauzei dintr-o comunitate mică. Au fost identificate femeile cu cancer endometrial (cazuri). Pentru fiecare femeie din grupul caz s-a identificat o femeie de aceeași vârstă și cu aceeași durată a șederii în respectiva comunitate. Întrebarea medicală a fost dacă cancerul endometrial a fost legat de utilizarea medicamentelor cu estrogen.

Abb	Asocierea dintre utilizarea estrogenilor și cancerul endometrial.
P	Asocierea (2 variabile calitative dicotomiale: Estrogen (da/nu) și grupul (caz/martor))
H	$H_0: \pi_b = \pi_c$ $H_1: \pi_b \neq \pi_c$
A	Condiții de aplicare a testului: <ul style="list-style-type: none">○ Date perechi: înainte/după sau grupuri perechi<ul style="list-style-type: none">○ O variabilă dicotomială (utilizarea estrogenilor da/nu) și o variabilă independentă cu grupuri perechi (caz/martor)○ În același grup datele sunt mutual exclusive
N	Testul McNemar

Observat	Estrogen martor	Fără estrogen martor	Total
Estrogen caz	27	29	56
Fără estrogen caz	2	8	10
Total	29	37	66

PHANTOMS: McNemar

Abb	Asocierea dintre utilizarea estogenilor și cancerul endometrial.
T	χ^2 (formula with continuity correction) = $(b-c -1)^2/(b+c) = (29-2-1)^2/(29+2) = 21.81$
O	Valoarea P (nivel de semnificație = 5%) < 0.0001
M	χ^2 aparține regiunii de respingere $[3.84; \infty) \rightarrow$ avem suficiente evidențe să respingem H_0 Valoarea P < 0,05 \rightarrow avem suficiente evidențe să respingem H_0
S	Cancerul endo. etrial este asociat cu utilizarea estogenului.

$$OR = b/c = 29/2 = 14.5 - IC95\% = [4.46 - 60.76]$$

Riscul de a dezvolta cancer endometrial este de 14 ori mai mari pentru femeile care au fost pe terapie cu estrogen comparativ cu cele care nu au fost.

Care este riscul pe care mă aștept să-l văd în populație?



Testul Z de comparare a frecvențelor

Testul Z: 1. Compararea unei frecvențe observate cu o frecvență teoretică (π_0)

P	Frecvența
H	$H_0: \pi = \pi_0$ vs. $H_1: \pi \neq \pi_0$
A	Condiții de aplicare a testului: <ul style="list-style-type: none"> ○ $n \cdot f > 10$ (unde n = volumul eșantionului; f = frecvența observată în eșantion) ○ $n \cdot (1-f) > 10$
N	Testul Z
T	$Z = \frac{f - \pi_0}{\sqrt{\frac{\pi_0 \cdot (1 - \pi_0)}{n}}}$
O	Valoarea P (nivel de semnificație = 5%)
M	Z aparține regiunii de respingere $(-\infty; -1.96] \cup [1.96; \infty) \rightarrow$ avem suficiente dovezi să respingem H_0 Dacă valoarea P-value $< 0.05 \rightarrow$ avem suficiente dovezi să respingem H_0
S	

Testul Z

S-a investigat frecvența hepatitei B la personalul medical care lucrează în spitalele de boli infecțioase din Transilvania. Se știe din studiile anterioare că prevalența hepatitei B în populația generală a Transilvaniei este de 9%. S-a studiat un eșantion de 300 de persoane cu loc de muncă în spitalele de boli infecțioase și s-a obținut o frecvență de hepatită B egală cu 6%. Este frecvența hepatitei B la personalul medical care lucrează în spitalele de boli infecțioase semnificativ diferită față de cea din populație generală?

P	Frecvențe relative (cea din populație este cunoscută)
H	$H_0: \pi = \pi_0$ $H_1: \pi \neq \pi_0$
A	Condiții de aplicare: <ul style="list-style-type: none"> ○ $300 \cdot 0.06 > 10$ ○ $300 \cdot (1 - 0.06) > 10$
N	Testul Z
T	$Z = (f - \pi_0) / \sqrt{\pi_0 \cdot (1 - \pi_0) / n} = (0.06 - 0.09) / \sqrt{0.09 \cdot (1 - 0.09) / 300} = -1.8$
O	Valoarea P = 0.0790
M	Regiunea de respingere $(-\infty; -1.96] \cup [1.96; \infty)$: $-1.96 \leq -1.8 \leq 1.96 \rightarrow$ nu avem suficiente dovezi să respingem H_0
S	Hepatita B nu este mai frecventă la personalul medical care lucrează în spitalele de boli infecțioase în comparație cu populația generală.

Testul Z: 2.

Compararea a două frecvențe

P	Frecvențe relative
H	$H_0: \pi_1 = \pi_2$ $H_1: \pi_1 \neq \pi_2$
A	Condiții: <ul style="list-style-type: none">○ Numărul observațiilor este suficient de mare ($n_1 > 30$ și $n_2 > 30$)
N	Testul Z
T	$Z = (f_1 - f_2) / \sqrt{f(1-f)(1/n_1 + 1/n_2)}$, where $f = (f_1 n_1 + f_2 n_2) / (n_1 + n_2)$
O	Valoarea P ($\alpha=5\%$): =NORM.S.DIST(Z,FALSE)
M	Z aparține regiunii de respingere $(-\infty; -1.96] \cup [1.96; \infty) \rightarrow$ avem suficiente dovezi să respingem H_0 Valoarea P $< 0.05 \rightarrow$ avem suficiente dovezi să respingem H_0
S	

Testul Z: 2.

Compararea a două frecvențe

	Females (n = 542)	Males (n = 486)
Sedentariness	352 (70.5)	274 (56.4)

Sedentarismul și sexul	
P	Frecvența sedentarismului
H	$H_0: \pi_1 = \pi_2$ (Frecvența sedentarismului la femei nu este semnificativ diferită față de frecvența sedentarismului la bărbați) $H_1: \pi_1 \neq \pi_2$ (Frecvența sedentarismului la femei este semnificativ diferită față de frecvența sedentarismului la bărbați)
A	Condiții de aplicare: $n_1 = 542$ și $n_2 = 486$
N	Testul Z
T	$f_1 = 352/542 = 0.65$; $f_2 = 274/486 = 0.56$; $f = (0.65 \cdot 542 + 0.56 \cdot 486) / (542 + 486) = 0.6075$ $Z = (0.65 - 0.56) / \sqrt{0.6075 \cdot (1 - 0.6075) \cdot (1/542 + 1/486)} = 2.95$
O	Valoarea P ($\alpha = 5\%$): $= \text{NORM.S.DIST}(2.95, \text{FALSE}) = 0.0051$
M	Regiunea de respingere $(-\infty; -1.96] \cup [1.96; \infty)$: $Z = 2.95 > 1.98 \rightarrow$ Avem suficiente dovezi să respingem H_0 & Valoarea P = 0.0051 \rightarrow Avem suficiente dovezi să respingem H_0
S	Există o asociere semnificativă statistic între sex și sedentarism.

Testul Hi-pătrat

	Females (n = 542)	Males (n = 486)
Sedentariness	352 (70.5)	274 (56.4)

Asocierea dintre sedentarism și sex	
P	Frecvențe absolute
H	H ₀ : Sedentarisul și sexul sunt independente H ₁ : Sedentarisul și sexul sunt dependente
A	Condiții de aplicare: (vezi testul Hi-pătrat)
N	Testul Hi-pătrat
T	$\chi^2 = (352-330)^2/330 + (274-296)^2/296 + (190-212)^2/212 + (212-190)^2/190 = 7.93$
O	Valoarea P ($\alpha=5\%$): 0.0050 (=CHISQ.TEST(array-observat ,array-teoretic))
M	$\chi^2 = 7,93$ aparține regiunii de respingere $[3.84; \infty) \rightarrow$ avem suficiente dovezi să respingem H ₀ P < 0.05 \rightarrow avem suficiente dovezi să respingem H ₀
S	Sedentarismul și genul sunt dependente.

Observat	Feminin	Masculin	Total
sedentarism=da	352	274	626
sedentarism=nu	190	212	402
Total	542	486	1028

Teoretic	Feminin	Masculin	Total
sedentarism=da	330	296	626
sedentarism=nu	212	190	402
Total	542	486	1028

Lorraine Frazier,^{1,*} Erica Yu,¹ Jennifer Sanner,¹ Fang Liu,¹ Malini Udtha,¹ Stanley Cron,¹ Stephanie Coulter,² and Roberta C. Bogaev²

TABLE 1: Comparison of demographic and clinical health characteristics of acute coronary syndrome patients.

Variables	(<i>n</i> = 789)	Women (<i>n</i> = 248)	Men (<i>n</i> = 541)	<i>P</i> value
Race				0.0274*
White	485 (61.47)	141 (56.85)	344 (63.59)	
Black	152 (19.26)	63 (25.40)	89 (16.45)	
Spanish/Hispanic Latino Descent	123 (15.59)	37 (14.92)	86 (15.90)	
Some other race	29 (3.68)	7 (2.82)	22 (4.07)	
Marital Status				<0.0001*
Single, Divorced, Separated, Widowed	272 (34.56)	126 (51.01)	146 (27.04)	
Married	515 (65.44)	121 (48.99)	394 (72.96)	

Note: values in parenthesis are percentages.

*Statistical significant at level of $P < 0.05$.

[illegible]

Vaccination Behaviour Influences Self-Report of Influenza Vaccination Status: A Cross-Sectional Study among Health Care Workers

Anna Llupia, Alberto L. García-Basteiro, Guillermo Mena, José Ríos, Joaquim Puig, José M. Bayas, Antoni Trilla
Published: July 11, 2012 • <https://doi.org/10.1371/journal.pone.0039496>

Table 2. Validity of self-reported influenza vaccination in HCW compared to occupational health medical records (2007–2009).

Campaign		Record		McNemar P
		Yes	No	
2007	Declare Yes	28	43	$p < 0.0001$
	No	9	118	
2009s	Declare Yes	71	25	$p = 0.002$
	No	7	145	

		Controls	
		(+)	(-)
Cases	(+)	71	25
	(-)	7	145
		78	170

Measures of Association				
Test	Value	d.f.	P Values	
			1-tail	2-tail
McNemar:	16.13	1		0.001463
McNemar with continuity correction:	9.031	1		0.002654
Fisher exact			0.001051	0.002102
Mid-P exact			0.0006594	0.001319

There are 32 discordant pairs.
Because this number is ≥ 20 , the McNemar test can be used.

Odds-based Estimates				
Parameter	Point Estimate	95% Confidence Intervals Lower,Upper		Type
Pair-Matched Odds Ratio:	3.571	1.545, 8.257 ¹		Taylor series



- Testul Hi-pătrat este valid dacă cel puțin 80% dintre frecvențele probabile (teoretice/așteptate) depășesc 5 și toate frecvențele probabile depășesc 1
 - Regula se aplică frecvențelor teoretice/așteptate, nu celor observate!
 - Dacă criteriul nu este satisfăcut → se aplica corecția Yates (corecție de continuitate) sau testul exact al lui Fisher.
- Testul Hi-pătrat se poate aplica pe orice tabel de contingență ($r \times c$)
- Atenție la condițiile de aplicare!
- Testul McNemar se aplică dacă datele sunt perechi!