

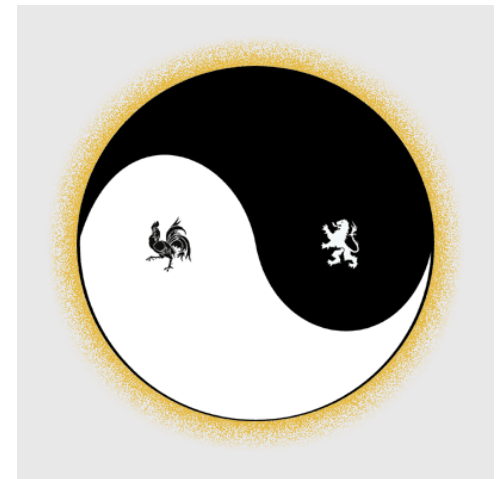
# Egy és (többváltozós) logisztikus regressziós vizsgálatok és alkalmazásaik a klinikumban



Dr. Prohászka Zoltán  
*Az MTA doktora*  
Semmelweis Egyetem  
III. Sz. Belgyógyászati Klinika  
2017-11-23

[prohoz@kut.sote.hu](mailto:prohoz@kut.sote.hu)

[www.kutlab.hu](http://www.kutlab.hu)



# Ismétlés: A reprodukálhatóság krízise a tudományos szakirodalomban

- **A jelenség**

a tudományos szakcikkeknek töredékét sikerült utánvizsgálatokban reprodukálni.

- **Az okok**

- Tervezési hibák, alacsony elemszám és az ismeretlen „statisztikai erő”
- Értelmezési hibák: a null hipotézis,  $p$  érték és a konfidencia intervallum fogalmában
- Eredmények szelekciója, a negatív kutatási eredmények nem nyilvánosak
- Rejtett (vagy eltitkolt) többszörös összehasonlítások
- Adatok manipulálása „szignifikancia” elérésére

- **A javítás módszerei**

- A biostatisztikai oktatása színvonalának javítása
- Hatásméret, elemszám, reprodukálhatóság szintjének tervezése
- Adatmanipulálási szokások tárgyalása
- Megbízható negatív eredmények számára közzéadási lehetőség
- Biostatisztikai módszertani leírás, biostatisztikai lektorálás
- Tervezés lektorálása, nyilvánosan regisztrált kutatások elterjesztése
- A reprodukálhatóság vizsgálatának tudományos értéként elismerése
- Open Data közzéadás

# Ismétlés

- Egyváltozós modellekkel megválaszolható kérdések
  - Egy megfigyelt különbség vagy arány a véletlen hatására alakult-e ki?
- Többváltozós modellekkel megválaszolható kérdések
  - Az adott független változó(k) szignifikáns kapcsolatban vannak-e a függő változóval?
  - Van-e kapcsolat a független változók között?
  - Mekkora a független változók egymáshoz képest mutatott relatív súlya a megfigyelt különbség (arány) kialakításában?
  - Más szavakkal: segít súlyozni, megítélni, típusba sorolni a független változóinkat.

Variable	Univariate Analysis	
	OR (95% CI)	P Value
Age	1.03 (1.02-1.05)	<0.001
Sex, male	0.91 (0.55-1.51)	0.72
Hypertension	1.82 (1.09-3.04)	0.02
Diabetes mellitus	2.43 (1.45-4.09)	0.001
pre-CKD	5.47 (3.18-9.39)	<0.001
Prior hospitalization for HF	1.28 (0.79-2.06)	0.31
LVEF <45%	1.16 (0.72-1.87)	0.53
NYHA, class IV	2.24 (1.39-3.63)	0.001
ln-NT-proBNP <sup>†</sup>	1.58 (1.23-2.04)	<0.001
SBP	1.01 (1.00-1.02)	0.06
DBP	1.01 (1.00-1.03)	0.12
Serum creatinine <sup>†</sup>	1.01 (1.01-1.02)	<0.001
Serum albumin	0.92 (0.88-0.96)	0.001
Hemoglobin	0.99 (0.98-0.99)	0.003
Use ACE/ARB preadmission	1.16 (0.71-1.91)	0.55
Use diuretic preadmission	1.22 (0.75-1.99)	0.42
ln-UACR <sup>†</sup>	1.63 (1.38-1.93)	<0.001
ln-uNGAL <sup>†</sup>	2.01 (1.65-2.44)	<0.001
ln-uAGT <sup>†</sup>	1.99 (1.66-2.37)	<0.001

- 321 beteg akut szívelégtelenséggel
- 30% 7 napon belül veseelégtelen is lett.
- Eset-kontroll design a rizikófaktorok megállapítására

# A független változó

Típusa	Jellegzetességei	Felismerési lehetőség	Példa
<b>Zavaró faktor (confounder)</b>	Kapcsolatban áll a rizikótényezővel és okozati összefüggésben a kimenetellel	Rétegzett analízis Többváltozós analízis A statisztikai analízis azonban önmagában nem elegendő, hogy a kettő között különbséget lehessen tenni.	Öngyújtó, dohányzás, tüdőrák
<b>Közbeeső változó (intervening variable)</b>	Kapcsolatban áll a rizikótényezővel és okozati összefüggésben a kimenetellel, a rizikótényező általa fejt ki hatását		Doh, cotinine, toxikus metabolitok, tüdőrák
<b>Gátló faktor (supresser)</b>	Kapcsolatban áll a rizikótényezővel és a kimenetellel, a rizikófaktor hatását elfedi	Rétegzett analízis a supresser változó szerint Többváltozós analízis, adjusztálás (illesztés) a supresser változóra	Aspirin, MI, halálozás Orvosi hivatás, válás, öngyilkosság
<b>Hatásmódosító vagy interakciós változó (interaction variable)</b>	Hatásmódosítás. A rizikótényező hatása a kimenetelre egy harmadik változó értékétől függ.	Rétegzett analízis az interakciós változó szerint. Többváltozós analízis „with interaction terms”	Expozíció, inokuláció mértéke, betegség súlyossága

# Többváltozós analízisek

<b>Függő változó</b>	<b>Példa a kimenetelre</b>	<b>A használandó többszörös analitikai módszer</b>
Folyamatos	Vérnyomás, testsúly, hőmérséklet	Többszörös lineáris regresszió
Dichotom (igen-nem)	Halál, betegség kialakulása, felvétel osztályra	Többszörös logisztikus regresszió
Az eseményig eltelt idő	Halálig, dg.-ig eltelt idő	Cox regresszió (proportional hazard analízis)

# A logisztikus regresszió

- Kategórikus függő (dependes) változó értékeinek becslésére, előrejelzésére szolgáló modell
  - Binary: két kategória
  - Multinomial: több, független kategória
  - Ordinális: trend a kategóriák között
- A modellt valós, mért, számolt valószínűségekre alapozzuk, ezeket összességükben magyarázó, független (independent) változóknak nevezzük (explanatory or predictor variables)
  - A független változók lehetnek nominális vagy kategória változók

# Problémabemutató, kérdésfeltevés:

- 195 szívelégtelenségben szenvedő beteg kereste fel az ambulanciát 1 év alatt.
- 92 betegnek súlyos szívelégtelensége volt szubjektív, ön-bevallásos skálán (NYHA)
  - Kik azok a betegek, akik várhatóan a súlyos csoportba tartoznak? *(milyen objektív mutatók alapján jósolható ez meg?)*
  - Ha több ilyen változó is van, akkor vajon ezek egymás hatásán keresztül vagy függetlenül jelzik a halálozási rizikót? *(akik anémiásak és/vagy veseelégtelenségük is van?)*
  - Ha van több független magyarázó változó, akkor ezeknek milyen az egymáshoz viszonyított fontossága, súlya?  
*(kétszer olyan fontos az anémia, mint a veseérintettség?)*



	Tüdőpangás nincs	Tüdőpangás van	<b>Össze n</b>
NYHA 1-2	82	21	<b>103</b>
NYHA 3-4	32	60	<b>92</b>
<b>Összesen</b>	<b>114</b>	<b>81</b>	<b>195</b>

Chi-square, df 38.39, 1  
P value  $P < 0.0001$

# A független (prediktor, magyarázó) változók

- A független változók lehetnek
  - Kategorikusak
    - Dichotomok – pl. fogyaszt-e alkoholt?
    - 0/1-nek kódoljuk – 0 - a referencia csoport, 1- a vizsgált csoport (Dummy változó)
    - Ha több kategória is előfordul, akár igen sok dummy változót is csinálhatunk: „milyen gyakran fogyaszt alkoholt?”
  - Folytonosak (vérnyomás, életkor)
    - Egységnyi változáshoz tartozó esélyhányados növekedést adja meg
    - Nehezen interpretálható
    - → folyamatos változó kategorikussá alakítása (pl.: medián alatt/felet, melyik tercilisbe tartozik)
    - → folyamatos változó normalizálása (1 SD [Z-score], IQ tartomány)
- Számuk korlátozott – ökölszabály: kb. 6-10x legyen több eset a kisebb csoportban, mint a független változók száma

	Tüdőpan- gás nincs	Tüdőpan- gás van	Össze- se- n
NYHA 1- 2	82	21	<b>103</b>
NYHA 3- 4	32	60	<b>92</b>
<b>Összesen</b>	<b>114</b>	<b>81</b>	<b>195</b>

- A példa kérdésfeltevése szerint („mivel jósolhatom meg, hogy ki a súlyos beteg”):
  - Melyik a független változó?
  - Melyik a függő változó?
  - Hány független változót tudnék elemezni ebben a vizsgálatban?
  - Ki tudná megsaccolni az odds-ot?

# Esély, odds

- Dichotom változónál az egyik esemény valószínűsége osztva a másik esemény valószínűségével
- Tartománya: 0 – végtelen
- „mennyivel (*milyen mértékben*) esélyesebb, hogy az adott egyén az eseményes csoportba kerül, mint a nem eseményesbe”
  - Dobókocka
    - 6-os legyen:  $1/6$
    - Ne legyen hatos:  $5/6$
    - Odds hogy ne legyen 6-os: 5 (hogy hatos legyen: 0,2)

# Esély arány, odds ratio (OR)

- OR – ezzel számszerűsíti a logisztikus regresszió a függő és független változó közötti összefüggést
- Az esélyek arányát hasonlítja össze a függő változó szerinti csoportokban, vagyis...
- ...megadja mennyivel *változik* az esélye annak, hogy az eseményes csoportba tartozik egy egyén, ha a *független változó értéke egy egységnyivel változik*

	Tüdőpan- gás nincs	Tüdőpan- gás van	Össze- se n
NYHA 1- 2	82	21	<b>103</b>
NYHA 3- 4	32	60	<b>92</b>
<b>Összesen</b>	<b>114</b>	<b>81</b>	<b>195</b>

- *A példa kérdésfeltevése szerint („mivel jósolhatom meg, hogy ki a súlyos beteg”):*
  - *Melyik a független változó?*
  - *Melyik a függő változó?*
  - *Hány független változót tudnék elemezni ebben a vizsgálatban?*
  - **Ki tudná megsaccolni az esély arányt (odds ratio)?**
  - **„Mennyi az esélye, hogy aki pang, súlyosabb csoportba tartozzon, ahhoz képest, mint aki nem pang?”**

	Tüdőpan- gás nincs	Tüdőpan- gás van	Össze- se n
NYHA 1- 2	82	21	<b>103</b>
NYHA 3- 4	32	60	<b>92</b>
<b>Összesen</b>	<b>114</b>	<b>81</b>	<b>195</b>

$$OR = \frac{[A/(A+C)] / [C/(A+C)]}{[B/(B+D)] / [D/(B+D)]} = \frac{A/C}{B/D} = \frac{AD}{BC}$$

$$OR = \frac{82/32}{21/60} = \frac{2,56}{0,35} = 7,31$$

- *A példa kérdésfeltevése szerint („mivel jósolhatom meg, hogy ki a súlyos beteg”):*
  - *Melyik a független változó?*
  - *Melyik a függő változó?*
  - *Hány független változót tudnék elemezni ebben a vizsgálatban?*
  - **Ki tudná megsaccolni az esély arányt (odds ratio)?**
  - **„Mennyi az esélye, hogy aki pang, súlyosabb csoportba tartozzon, ahhoz képest, mint aki nem pang?”**

	Tüdőpangás nincs	Tüdőpangás van	<b>Össze n</b>
NYHA 1-2	82	21	<b>103</b>
NYHA 3-4	32	60	<b>92</b>
<b>Összesen</b>	<b>114</b>	<b>81</b>	<b>195</b>

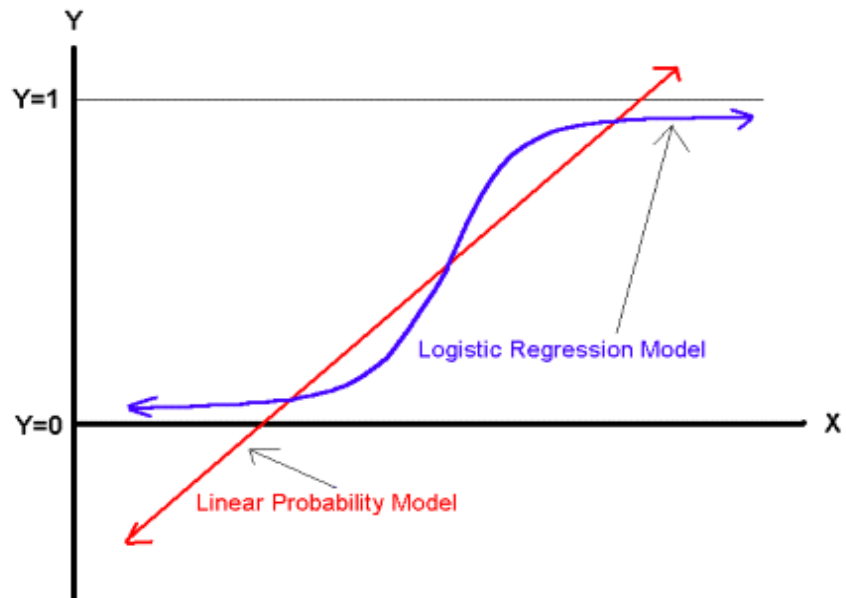
Chi-square, df 38.39, 1

P value  $P < 0.0001$

Odds ratio: 7,321, 95% confidence interval 3.847 to 13.93



## Comparing the LP and Logit Models



# Logisztikus regresszió egyenlete

$$\text{logit } [p(x)] = \log \left[ \frac{p(x)}{1-p(x)} \right] = a + b_1 x_1 + b_2 x_2 + b_3 x_3 \dots$$

- $p$  – a valószínűség, hogy az egyén valamelyik csoportba tartozik
- $a$  – a konstans
- $b$  – az adott prediktor regressziós koefficiense
  - Ez az  $\ln(\text{OR})!$
  - Sok helyen béta ( $\beta$ )
  - $\text{OR} = e^\beta$  – egymásból számíthatóak

# A regressziós együtthatókból kiolvashatjuk...

## **Béta (b):**

- ... hogy az adott változó hogyan befolyásolja az  $y$  (kimeneteli változó) bekövetkezésének esélyét:  
 $\beta > 0$  – növeli,  $\beta < 0$  – csökkenti,  $\beta = 0$  – nem befolyásolja

## **OR:**

- Folytonos magyarázó változóknál:
  - ... hogy egy egységnyi növekedés hányszorosára növeli a  $y$  bekövetkeztének esélyét (odd)
- ha az  $x$  magyarázó változó dichotom
  - ... hogy a vizsgált csoportban az  $y$  bekövetkezésének esélye (odd) hányszorososa a referenciacsoportbelinek
- ha a magyarázó változó több mint 2 kategóriás
  - ... hogy a szóban forgó csoportban az  $y$  bekövetkezésének esélye (odd) hányszorososa a referenciacsoportbelinek

Variable	Univariate Analysis	
	OR (95% CI)	P Value
Age	1.03 (1.02-1.05)	<0.001
Sex, male	0.91 (0.55-1.51)	0.72
Hypertension	1.82 (1.09-3.04)	0.02
Diabetes mellitus	2.43 (1.45-4.09)	0.001
pre-CKD	5.47 (3.18-9.39)	<0.001
Prior hospitalization for HF	1.28 (0.79-2.06)	0.31
LVEF <45%	1.16 (0.72-1.87)	0.53
NYHA, class IV	2.24 (1.39-3.63)	0.001
ln-NT-proBNP <sup>†</sup>	1.58 (1.23-2.04)	<0.001
SBP	1.01 (1.00-1.02)	0.06
DBP	1.01 (1.00-1.03)	0.12
Serum creatinine <sup>†</sup>	1.01 (1.01-1.02)	<0.001
Serum albumin	0.92 (0.88-0.96)	0.001
Hemoglobin	0.99 (0.98-0.99)	0.003
Use ACE/ARB preadmission	1.16 (0.71-1.91)	0.55
Use diuretic preadmission	1.22 (0.75-1.99)	0.42
ln-UACR <sup>†</sup>	1.63 (1.38-1.93)	<0.001
ln-uNGAL <sup>†</sup>	2.01 (1.65-2.44)	<0.001
ln-uAGT <sup>†</sup>	1.99 (1.66-2.37)	<0.001

- 321 beteg akut szívelégtelenséggel
- 30% 7 napon belül veseelégtelen is lett.
- Eset-kontroll design a rizikófaktorok megállapítására

# Wald chi négyzet teszt

- A regressziós egyenletben az egyes prediktorok (x) hatásának szignifikanciáját teszteli
- $H_0$  – a változónak nincs szerepe a függő változó kimenetében
- Ha a OR 95%-os konfidencia intervalluma tartalmazza az 1-es értéket, akkor az adott prediktor nem szignifikáns

**TABLE 7.1**

Frequency of primary PCP prophylaxis among patients whose AIDS-defining diagnosis was PCP.

Characteristic	Did not receive PCP prophylaxis n (%)	Received primary PCP prophylaxis n (%)	Adjusted P value	Adjusted odds ratio	95% confidence limits
Total	212 (65.0)	114 (35.0)			
Age group					
<35 years	61 (69.3)	27 (30.7)		1.0 (ref.)	
≥35 years	151 (63.5)	87 (36.6)	0.55	1.19	0.68, 2.07
Ethnicity					
Nonwhite	74 (77.9)	21 (22.1)		0.49	0.28, 0.87
White	138 (59.7)	93 (40.3)	0.01	1.0 (ref.)	
Sex					
Male	206 (64.6)	113 (35.4)		0.81	0.06, 10.12
Female	6 (85.7)	1 (14.3)	0.87	1.0 (ref.)	
Sexual orientation					
Gay/bisexual man	185 (62.7)	110 (37.3)		3.19	0.78, 13.03
Heterosexual	27 (87.1)	4 (12.9)	0.11	1.0 (ref.)	
Injection drug use					
Yes	35 (72.9)	13 (27.1)		1.11	0.49, 2.54
No	177 (63.7)	101 (36.3)	0.80	1.0 (ref.)	
Insurance					
None	52 (82.5)	11 (17.5)		0.35	0.17, 0.73
Public/private	151 (59.5)	103 (40.6)	0.005	1.0 (ref.)	

Adapted with permission from Schwarcz, S.K., et al. "Prevention of *Pneumocystis carinii* pneumonia: Who are we missing?" *AIDS* 1997; 11:1263-8. Copyright Rapid Science Publishers Ltd.

Wald Chi négyzet

OR a referencia csoportban mindig 1

Ha a OR CI-je nem tartalmazza az 1-et (50-50%), akkor szignifikáns.

Ha  $OR < 1$ , akkor az eseményre a referencia csoportnak van nagyobb esélye

# Összefoglalás

- Kontingencia tábla, esély-arány
- Binomiális (kategórikus) függő változó értékének jóslására a logisztikus regresszió pontosabb becslést ad, mint a lineáris
- A független változó kategóriái (értékei) miként jósolják a függő változót
  - Szignifikánsak?
  - (Függetlenek?)
  - Milyen az egymáshoz viszonyított súlyuk?

# Tavasszal: Modellépítés - Milyen független változók legyenek?

- A kérdésfeltevés, függő változó lényeges szempont
- Ha a betegség kimenetelével kapcsolatos:
  - Alap adatok – nem, életkor
  - Betegség súlyosságát jelző mutató
  - Az általunk tesztelni kívánt változó
  - Irodalmi adatokból ismert prediktorok
- Lehetséges irányvonal: egyváltozós modellben szoros összefüggést mutat



# „Automatikus” modellépítés

- „Stepwise regression”
  - Forward selection
  - Backward elimination
- A nem szignifikáns változókat kiválogatva megkeresi a legjobban illeszkedő modelleket
- Biológiaiailag releváns?

# Quick logit regression

		Model: Logistic regression (logit) N of 0's: 103 1's: 92 (Biostat_2_adatbázis Dep. var: NYHA_01 Loss: Max likelihood (MS-err. scaled to 1) Final loss: 114,02683255 Chi2( 1)=41,653 p=,00000 Modell illeszkedés	
		Const.B0	Pulmonal_congest ion_1-present
N=195			
Estimate	-0,9409834	1,990806	Béta
Standard Error	0,2084384	0,3282366	
t(193)	-4,514443	6,065155	A független változó hatása a függőre
p-level	0,00001102726	0,000000006822798	
-95%CL	-1,352093	1,343414	
+95%CL	-0,5298737	2,638197	
Wald's Chi-square	20,3802	36,7861	Az odds ratio
p-level	0,000006370944	0,000000001332023	
Odds ratio (unit ch)	0,3902439	7,321429	
-95%CL	0,2586983	3,832104	
+95%CL	0,5886793	13,98796	
Odds ratio (range)		7,321429	
-95%CL		3,832104	
+95%CL		13,98796	

Model: Logistic regression (logit) N of 0's: 103 1's: 92 (Biostat_2_adatbázis) Dep. var: NYHA_01 Loss: Max likelihood (MS-err. scaled to 1) Final loss: 111,45829993 Chi( 2)=46,790 p=,00000			
	Const.B0	age	Pulmonal_congestion_1-present
N=195			
Estimate	-3,196561	0,03286155	2,010672
Standard Error	1,059353	0,0149037	0,3346262
t(192)	-3,017467	2,204926	6,008711
p-level	0,002894578	0,02864696	0,0000000922290
-95%CL	-5,286025	0,003465547	1,350657
+95%CL	-1,107098	0,06225755	2,670688
Wald's Chi-square	9,105104	4,861698	36,10461
p-level	0,002550975	0,02746638	0,00000000189021
Odds ratio (unit ch)	0,04090263	1,033407	7,468338
-95%CL	0,005061844	1,003472	3,85996
+95%CL	0,3305168	1,064236	14,44991
Odds ratio (range)		8,874458	7,468338
-95%CL		1,258897	3,85996
+95%CL		62,55954	14,44991

# Hogy interpretáljuk az előző modellt?

- A súlyos szívelégtelenség fennállása kapcsolatot mutat a tüdőpangás jelenlétével. Kb. 7-szeres az esélye, ha valakinek pang a tüdeje, hogy a „súlyos” csoportba tartozzon.
- Ha súlyos és nem súlyos csoportba tartozó betegek egyenlő idők lennének, a tüdőpangás akkor is kb. 7-szeresére növelné az esélyét a „súlyos” csoportba tartozásnak, vagyis a pangás hatás a független az életkortól.
- „Adjusztálás”
- Fordítva is igaz: age OR: 1,03...nehéz interpretálni, elképzelni, mit is jelent ez a rizikó a gyakorlatban

Model: Logistic regression (logit) N of 0's: 103 1's: 92 (Biostat_2_adatbázis) Dep. var: NYHA_01 Loss: Max likelihood (MS-err. scaled to 1) Final loss: 111,79871325 Chi2( 2)=46,109 p=,00000			
	Const.B0	Pulmonal_congestion_1-present	age_65_hilo
N=195			
Estimate	-1,421601	2,038329	0,7181076
Standard Error	0,3227381	0,3357446	0,3452933
t(192)	-4,404813	6,07107	2,079703
p-level	0,0000175728	0,00000000666594	0,03888085
-95%CL	-2,058168	1,376108	0,0370523
+95%CL	-0,7850335	2,700551	1,399163
Wald's Chi-square	19,40238	36,85789	4,325166
p-level	0,0000106218	0,00000000128488	0,03756069
Odds ratio (unit ch)	0,2413274	7,677771	2,050549
-95%CL	0,1276876	3,95946	1,037747
+95%CL	0,4561044	14,88793	4,051807
Odds ratio (range)		7,677771	2,050549
-95%CL		3,95946	1,037747
+95%CL		14,88793	4,051807

Dummy var

„idős vs. fiatal”

# Az életkor és a közbeeső változó (vagy confounder?) kapcsolata a kimenettel

Model: Logistic regression (logit) N of 0's: 103 1's: 92 (Biostat_2_adatbázis) Dep. var: NYHA_01 Loss: Max likelihood (MS-err. scaled to 1) Final loss: 106,47776169 Chi2( 3)=56,751 p=,00000				
	Const.B0	Pulmonal_congesti on_1-present	HGB	age_65_hilo
N=195				
Estimate	3,138476	2,177778	-0,03215856	0,5294558
Standard Error	1,462641	0,3559044	0,01021556	0,359107
t(191)	2,14576	6,118996	-3,147996	1,474367
p-level	0,03315431	0,0000000522774	0,001907781	0,142029
-95%CL	0,253472	1,47577	-0,05230834	-0,1788692
+95%CL	6,023479	2,879786	-0,01200875	1,237781
Wald's Chi-square	4,604284	37,44211	9,909889	2,173759
p-level	0,03189967	0,000000000952580	0,00164544	0,1403925
Odds ratio (unit ch)	23,06868	8,82667	0,968353	1,698008
-95%CL	1,288491	4,374402	0,9490362	0,8362152
+95%CL	413,0132	17,81046	0,988063	3,447953
Odds ratio (range)		8,82667	0,05901568	1,698008
-95%CL		4,374402	0,01002038	0,8362152
+95%CL		17,81046	0,3475766	3,447953

# A reprodukálhatóság krízise a tudományos szakirodalomban: ma tárgyalt vagy érintett aspektusok

- **A jelenség**  
a tudományos szócikkeknek töredékét sikerült utánvizsgálatokban reprodukálni.
- **Az okok**
  - Értelmezési hibák: a null hipotézis, p érték és a konfidencia intervallum fogalmában
  - 
  - Példa
- **A javítás módszerei**
  - A biostatistikai oktatása színvonalának javítása
  - Hatásméret, elemszám, reprodukálhatóság szintjének tervezése
  - **A logisztikus regresszió során legtöbbször „értelmes” csoportokra bontjuk a függő változót: ha nem egyforma nagyok a csoportok, nagy a veszélyes a statisztikai erő elvesztésének. A vizsgálatot a legkisebb csoport elemszámának megfelelően kell tervezni.**
  - Adatmanipulálási szokások tárgyalása
  - Megbízható negatív eredmények számára közlési lehetőség
  - Biostatistikai módszertani leírás, biostatistikai lektorálás
  - **A többváltozós elemzés hiánya, „remaining confounding”, hibás következtetés levonása**

- A kutatók eseteket és egészséges kontrollokat gyűjtöttek, és megállapították, hogy X biomarker szintje szignifikánsan ( $p < 0.05$ ) magasabb a betegekben, mint a kontrollokban
  - ROC analízissel megállapították az optimális vágópontot, amit alkalmazva 0,75 sen és 0,82 spec értékeket kaptak
  - Azt a következtetést vonták le, hogy a biomarker alkalmas diagnosztikai markere lehet a betegségnek
  - A biomarker összefüggését a betegség diagnosztikáját megalapozó mutatókkal nem elemezték, többváltozós (adjusztált) analízist nem végeztek.
- 
- Helyes a következtetés?
  - Ki tudja megfogalmazni ennek a vizsgálatnak az **orvosi** „nullhipotézisét”?
  - Vagyis: egy eset kontrollos vizsgálatban mit kereshetünk, ha a kontrollok egészséges személyek?