VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
MATEMATINĖS STATISTIKOS KATEDRA

Baigiamasis magistro darbas

Sergančiųjų lėtiniu virusiniu C hepatitu demografinių rizikos faktorių bei biocheminių ir histologinių ligos charakteristikų statistinė analizė

Parengė: PAULIUS BAKAITIS
Darbo vadovas: DOC. V. KAZAKEVIČIUS

Vilnius
2006

TURINYS:

1. Įvadas 4
1.1. Trumpa apžvalga 4
1.2. Darbo tikslas ir uždaviniai 5
2. Duomenų aprašymas 6
3. Teorija 8
3.1. Kepenų būklės įvertinimo rodikliai 8
3.2. Statistiniai metodai 9
4. Duomenų analizė 16
4.1. Pagrindinės duomenų charakteristikos 16
4.2. Alkoholio vartojimo įtakos nustatymas 17
4.3. Kepenų cirozės įtakoti rodiklių pakitimai 20
4.4. Palyginimas tarp skirtingų HAI kategorijų 24
4.5. Kitų rodiklių ryšio nustatymas 28
5. Išvados 30
6. Literatūra 31
7. Priedai 32

SUTRUMPINIMAI:

ALT – alaninė aminotransferazė.

AST – asparagininė aminotransferazė.

CIK – cirkuliuojantys imuniniai kompleksai.

DTL – didelio tankio lipoproteinų cholesterolis.

ENG – eritrocitų nusėdimo greitis.

GGT – gama-gliutamiltransferazė.

HAI – histologinis aktyvumo indeksas.

HB – hemoglobinas.

HCV – hepatito C virusas.

IgA – imunoglobulinas-a.

IgG – imunoglobulinas-g.

IgM – imunoglobulinas-m.

KMI – kūno masės indeksas.

LEU – leukocitai.

LHC – lėtinis hepatitas C.

MTL – mažo tankio lipoproteinų cholesterolis.

NBT – nitro-blue-tetrazoliumas.

RNR – ribonukleino rūgštis.

ŠF – šarminė fosfatazė.

TR – trombocitai.

1. ĮVADAS

1.1. Trumpa apžvalga

Hepatitas – tai kepenų uždegimas, kai dėl įvairių priežasčių pažeistos kepenų ląstelės žūva, į kepenų audinį įsiskverbia uždegiminės ląstelės, o žuvusių ląstelių vietą pamažu užpildo jungiamasis audinys. Kepenų uždegimą gali sukelti įvairios priežastys. Viena dažniausių – hepatitų virusai A, B ir C. Bene pavojingiausias iš jų yra hepatito C virusas (HCV).

HCV - tarsi uždelsto veikimo bomba, apie kurį visuomenė šiandien turi ne itin daug informacijos. Šios ligos sukėlėjas buvo išskirtas gana vėlai – tik 1989 m. Iki tol šis hepatitas buvo vadinamas “nei A nei B hepatitu”. HCV yra labai atsparus ir daugelį dienų gali būti gyvybingas netgi išdžiuvusiame kraujo lašelyje. Vakcinos bei imunoglobulino, apsaugančio nuo HCV, nėra, kadangi žmogaus organizmas nepajėgus gaminti antikūnų visiems ligos genotipams. Tai labai apsunkina vakcinos sukūrimą bei gydymo procesą. Pasaulio sveikatos organizacijos duomenimis hepatitu C yra užsikrėtę nuo 150 iki 200 milijonų pasaulio gyventojų. Nors oficialiai Lietuvoje yra diagnozuota apie 1,5 tūkstančio pacientų, sergančių hepatitu C, manoma, kad užsikrėtusių gali būti apie 50 - 70 tūkstančių.

HCV į žmogaus organizmą patenka per kraują, todėl didžiausią riziką užsikrėsti yra:

- ⇒ Persodinant donoro organus, audinius arba perpilant kraują.
- ⇒ Medicinos darbuotojams, susižeidusiems darbo metu.
- ⇒ Narkomanams, naudojantiems nesterilius švirštus.
- ⇒ Darantis tatuiruotes arba veriantis auskarus.
- ⇒ Lytiniu keliu, dažnai keičiantiems sekso partnerius.

Yra galimybė užsikrėsti HCV ir buityje, kai viruso nešiotojas ir jo artimieji ar kiti asmenys naudoja bendrus asmens higienos daiktus, pvz.: dantų šepetukus, skustuvus, manikiūro-pedikiūro reikmenis ar pan. Taip pat HCV motina gali

perduoti savo vaisiui, todėl maždaug 5-6 proc. hepatitu C infekuotų motinų kūdikių yra užsikrėtę.

Hepatitas C gali būti ūminis arba lėtinis. Susirgus ūminiu hepatitu C, tik apie 10–15 proc. ligonių pasveiksta, kitiems, jei liga trunka ilgiau kaip 6 mėnesius, vystosi lėtinis virusinis hepatitas C (LHC). LHC yra slaptas procesas, kuris lėtai progresuoja, keletą dešimtmečių nepasireikšdamas jokiais simptomais ir fiziniais požymiais. Maždaug 20 proc. ligonių LHC progresuoja į kepenų cirozę, o nuo 1 iki 4 proc. ligonių kasmet diagnozuojamas pirminis kepenų vėžys. LHC nėra pilnai pagydomas, tačiau, laiku diagnozavus ligą bei pradėjus gydymą, įmanoma sustabdyti ligos progresavimą, sumažinti kepenų pažeidimo laipsnį ir tokiu būdu atitolinti kepenų cirozės ar pirminio kepenų vėžio išsivystymą, o tuo pačiu ir poreikį transplantuoti kepenis.

1.2. Darbo tikslas ir uždaviniai

Darbo tikslas:

Ištirti pacientų, kuriems diagnozuotas hepatitas C, ligos raiškos ir eigos ypatumus.

Darbo uždaviniai:

- ⇒ Apibendrinti ir apžvelgti turimus duomenis, išvesti pagrindines skaitines charakteristikas.
- ⇒ Nustatyti alkoholio vartojimo įtaką ligonių histologiniams bei biocheminiams rodikliams.
- ⇒ Nustatyti kepenų cirozės įtakotus rodiklių pakitimus.
- ⇒ Nustatyti ryšį tarp ligos histologinio aktyvumo indekso ir kraujo biocheminių rodiklių.
- ⇒ Priklausomybės tarp kitų rodiklių nustatymas.

Duomenų analizei naudosiu SPSS v11.5 ir SAS statistinius paketus.

2. DUOMENŲ APRAŠYMAS

Iš Vilniaus universiteto ligoninės Santariškių klinikų buvo gauta 1005 pacientų (amžius nuo 16 iki 87 metų), sergančių LHC, morfologiniai, histologiniai ir biocheminių tyrimų duomenys. Buvo surinkti tokie pagrindiniai duomenys apie pacientus: amžius, lytis, kūno masės indeksas (KMI). Pacientai taip pat buvo apklausti dėl alkoholio vartojimo ir pagrindinių rizikos faktorių užsikrėsti HCV (donorystė, kraujo perpylimas, ilgalaikis gydymas, traumų bei operacijų skaičius, profesinė rizika, hemodializės procedūros, narkotikų vartojimas, tatuiruočių darymasis, kontaktas su kitais ligoniais). Kepenų biopsijos pagalba buvo gauti morfologiniai duomenys: nustatytas histologinis aktyvumo indeksas (HAI) bei kepenų fibrozės laipsnis. Echoskopijos pagalba buvo išmatuota kepenų ir blužnies dydis. Atlikus kraujo tyrimus buvo nustatyti tokie rodikliai:

⇒ Viruso genotipas.

⇒ Kepenų fermentai (ALT, AST, ŠF, GGT).

⇒ Pagrindiniai kraujo parametrai (hemoglobinas, leukocitai, eritrocitų nusėdimo greitis, trombocitai).

⇒ Lipidai (bendras cholesterolis, MTL ir DTL cholesteroliai, trigliceridai).

⇒ Tulžies pigmentai - bilirubinai (bendras, tiesioginis ir netiesioginis).

⇒ Imuniniai kompleksai (CIK, NBT).

⇒ Antikūniai – imunoglobulinai (IgA, IgG ir IgM).

Taigi turime tokius kiekybinius kintamuosius:

Eil. Nr.	Kintamasis	Matavimo vnt.	Trumpinys
1	Paciento amžius	metai	AMZIUS
2	Kūno masės indeksas	balai	KMI
3	Histologinis aktyvumo indeksas	balai	HAI
4	Kepenų dydis	mm	KEPENYS
5	Blužnies dydis	mm	BLUZNIS
6	Alaninė aminotransferazė	U / L	ALT
7	Asparagininė aminotransferazė	U / L	AST
8	Šarminė fosfatazė	U / L	SF
9	Gama-gliutamiltransferazė	U / L	GGT

Eil. Nr.	Kintamasis	Matavimo vnt.	Trumpinys
10	Hemoglobinas	g / l	HB
11	Leukocitai	$*10^9$ / l	LEU
12	Eritrocitų nusėdimo greitis	mm / h	ENG
13	Trombocitai	$*10^9$ / l	TR
14	Bendras cholesterolis	mmol / l	B_CHOLES
15	DTL-cholesterolis	mmol / l	DTL_CHOL
16	MTL-cholesterolis	mmol / l	MTL_CHOL
17	Trigliceridai	mmol / l	TRIGLIC
18	Bendras bilirubinas	μmol / l	BILIRUB
19	Tiesioginis bilirubinas	μmol / l	T_BILIR
20	Netiesioginis bilirubinas	μmol / l	N_BILIR
21	Imuniniai kompleksai	μg / ml	CIK
22	Nitro-blue-tetrazoliumas	U / L	NBT
23	Imunoglobulinas-A	IU / ml	IGA
24	Imunoglobulinas-G	IU / ml	IGG
25	Imunoglobulinas-M	IU / ml	IGM

Kokybiniai nominalūs kintamieji:

Pilnas pavadinimas	Trumpinys	Variantai	Kodas
Alkoholio vartojimas	ALKOHOL	Nevartoja	0
		Vartoja	1
Lytis	LYTIS	Vyras	1
		Moteris	2
Cirozė	CIROZE	Yra	1
		Nėra	0
Genotipas	GENOTIP	1 tipas	1
		2 tipas	2
		3 tipas	3
Ribonukleino rūgšties testas	RNR	Neigiamas	0
		Teigiamas	1

Kokybiniai ranginiai kintamieji:

Pilnas pavadinimas	Trumpinys	Variantai	Kodas
Histologinis aktyvumo indeksas (ranguotas)	HAI_R	HAI nuo 1 iki 3	1
		HAI nuo 4 iki 8	2
		HAI nuo 9 iki 12	3
		HAI nuo 13 iki 18	4
Kepenų fibrozės laipsnis	FIBROZE	Be fibrozės	0
		Neryški fibrozė	1
		Vidutinė fibrozė	2
		Ryški fibrozė	3
		Labai ryški fibrozė (cirozė)	4

3. TEORIJA

3.1. Kepenų būklės įvertinimo rodikliai

Histologinis aktyvumo indeksas (HAI):

HAI parodo kepenų uždegimo lygį. Jis nustatomas biopsijos pagalba paėmus kepenų bioptatą (mėginį) ir ištyrus jį per mikroskopą. HAI yra suma kepenų pažeidimo požymių įvertintų balais:

- ⇒ Periportinės ir periseptinės laiptinės nekrozės: 0 - nėra. 1 - pavienės. 2 - pavienės / vidutinės. 3 - vidutinės. 4 - didelės.
- ⇒ Susiliejančios nekrozės: 0 – nėra. 1 - židininės susiliejančios nekrozės. 2 - trečiosios zonos nekrozė keliose srityse. 3 - trečiosios zonos nekrozė daugelyje sričių. 4 - trečiosios zonos nekrozė ir pavieniai portocentriniai tiltai. 5 - trečiosios zonos nekrozė ir dauginiai portocentriniai tiltai. 6 - visuotinė skiltelių ar daugiaskiltelinė nekrozė.
- ⇒ Židininės nekrozės, apoptozė ir židininis uždegimas: 0 - nėra. 1 - vienas židinys ar mažiau viename regėjimo lauke. 2 - 2–4 židiniai regėjimo lauke. 3 - 5–10 židinių regėjimo lauke. 4 - daugiau negu 10 židinių regėjimo lauke.
- ⇒ Vartų uždegimas: 0 – nėra. 1 - nežymus, keliuose ar visuose vartų laukuose. 2 - vidutinis, keliuose ar visuose vartų laukuose. 3 - vidutinis / didelis, visuose vartų laukuose. 4 - didelis, visuose vartų laukuose.

Maksimalus HAI laipsnis – 18.

Kepenų fibrozė:

Kepenų fibrozė (skaidulinio audinio vešėjimas kepenyse) taip pat nustatoma ištyrus kepenų bioptatą ir įvertinama laipsniais pagal tokius požymius:

0 – nėra fibrozės. 1 – fibrozuoti išplėsti vartų laukai, nėra užtvarų. 2 – pavienės vartų užtvaros. 3 – gausios užtvaros, tačiau nėra cirozės. 4 – cirozė.

Kepenų fermentai:

GGT yra vienas jautriausių kepenų pažeidimo rodiklių. Vienkartinis alkoholio ar hepatotoksinių medžiagų (medikamentų) panaudojimas gali sukelti GGT aktyvumo padidėjimą serume. Kartais nežinant šio parametro galimybių kepenų pažeidimui įvertinti naudojami mažiau jautrūs rodikliai – ALT ir AST. Nors GGT - tai fermentas, kurio randama daugelyje organų, jis praktiškai išryškėja tik kepenų, tulžies takų ir kasos patologijų atvejais. GGT nepakeičiamas įvertinant hepatitų, kurie nepasireiškia gelta, eigą.

Padidėjęs AST fermento aktyvumas dažniausiai vertinamas, kaip kepenų ir širdies patologijos rodiklis. AST padidėja daugelio kepenų patologijų metu, tačiau labai dideli AST aktyvumo padidėjimai vistik dažniausiai būna susiję su kepenų patologija, nors gali būti randami ir raumeninio audinio griūvimo, navikų nekrozės, gangrenos, šoko, hipoksijos metu. Ūmaus kepenų funkcijos nepakankamumo metu padidėja AST ir ALT, sumažėja gliukozės koncentracija ir atsiranda respiratorinė alkalozė. Kepenų ligų atvejais AST ir ALT rodikliai didėja nebūtinai kartu ir nebūtinai vienodu mastu.

3.2. Statistiniai metodai

Stjudento t-testas (*populiacijų dispersijos lygios*):

Dvi intervalinių duomenų imtys $(x_1, x_2, ..., x_n)$ ir $(y_1, y_{2,}..., y_n)$ gautos matuojant du nepriklausomus normaliuosius atsitiktinius dydžius $X \sim N(\mu_x, \sigma^2)$ ir $Y \sim N(\mu_x, \sigma^2)$. Vidurkiai μ_x, μ_y ir dispersija σ^2 nežinomi. Formuluojame statistinę hipotezę: $\begin{cases} H_0 : \mu_x = \mu_y, \\ H_1 : \mu_x \neq \mu_y. \end{cases}$

Paprasčiausias populiacijos vidurkių skirtumo $\mu_x - \mu_y$ įvertis yra $\overline{X} - \overline{Y}$. Tačiau vertinant vidurkių skirtumas reikia atsižvelgti ir į duomenų sklaidą, t. y. stebimų dydžių dispersijas. Todėl apskaičiuojama kriterijaus statistika bus lygi:

$$t = \frac{\bar{x} - \bar{y}}{s_p \sqrt{1/n + 1/m}} = \frac{\bar{x} - \bar{y}}{\sqrt{(n-1)s_x^2 + (m-1)s_y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}};$$

čia $\bar{x}, \bar{y}$ yra imčių vidurkiai, s_x^2, s_y^2 - imčių dispersijos, o n, m – imčių dydžiai.

Tegul reikšmingumo lygmuo α. Tada hipotezę H_0 atmetame, jeigu $|t| > t_{\alpha/2}(n+m-2)$, kur $t_{\alpha/2}(n+m-2)$ - yra Stjudento skirstinio su (n-1) laisvės laipsnių $\alpha/2$ lygmens kritinė reikšmė. Hipotezės H_0 neatmetame, jei $|t| \le t_{\alpha/2}(n+m-2)$.

Stjudento t-testas (*populiacijų dispersijos nelygios*):

Dvi intervalinių duomenų imtys $(x_1, x_2, ..., x_n)$ ir $(y_1, y_{2,}, ..., y_n)$ gautos matuojant du nepriklausomus normaliuosius atsitiktinius dydžius $X \sim N(\mu_x, \sigma_x^2)$ ir $Y \sim N(\mu_x, \sigma_y^2)$. Vidurkiai μ_x, μ_y ir dispersijos σ_x^2, σ_y^2 nežinomi. Formuluojame statistinę hipotezę: $\begin{cases} H_0 : \mu_x = \mu_y, \\ H_1 : \mu_x \neq \mu_y. \end{cases}$

Apskaičiuojama kriterijaus statistika: $t = \frac{\bar{x} - \bar{y}}{\sqrt{s_x^2/n + s_y^2/m}}$;

čia $\bar{x}, \bar{y}$ yra imčių vidurkiai, s_x^2, s_y^2 - imčių dispersijos, o n, m – imčių dydžiai.

Tegul reikšmingumo lygmuo α. Tada hipotezę H_0 atmetame, jeigu $|t| > t_{\alpha/2}(k)$, kur $t_{\alpha/2}(k)$ - yra Stjudento skirstinio su k laisvės laipsnių $\alpha/2$ lygmens kritinė reikšmė. Laisvės lapsnių skaičius k yra mažiausias sveikasis skaičius, tenkinantis sąlygą: $k \le \frac{(s_x^2/n + s_y^2/m)^2}{s_x^4/n^3 + s_y^4/m^3}$.

Hipotezės H_0 neatmetame, jei $|t| \le t_{\alpha/2}(k)$.

Mano–Vitnio–Vilkoksono kriterijus:

Stjudento t kriterijų galima taikyti tik tada, kai kintamieji normalieji, tačiau normalumo sąlyga tenkinama ne visada arba šios sąlygos patikrinti neįmanoma. Tada yra taikomas Mano–Vitnio–Vilkoksono neparametrinis kriterijus. Tarkim,

turime dviejų nepriklausomų kintamųjų X ir Y stebėjimus $(x_1, x_2, ..., x_{n1})$ ir $(y_1, y_{2,}, ..., y_{n1})$. Duomenys gauti matavimams naudojant santykių, intervalų arba rangų skalę. Formuluojame statistinę hipotezę:

H_0: kintamųjų skirstiniai vienodi,

H_1: kintamųjų skirstiniai skirtingi.

Apskaičiuojame kriterijaus statistiką:

$Z = \frac{U_1 - \mu}{\sigma}$, kur $\mu = \frac{n_1 n_2}{2}$, $\sigma = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$.

Tarkime, reikšmingumo lygmuo yra α. Jei $|Z| > z_{\alpha/2}$, tai hipotezė H_0 atmetama, t. y. skirstiniai skiriasi. Čia $z_{\alpha/2}$ yra standartinio normaliojo skirstinio $\alpha/2$ lygmens kritinė reikšmė. Priešingu atveju H_0 neatmetame.

χ^2 nepriklausomumo kriterijus:

χ^2 nepriklausomumo kriterijus yra naudojamas tada, kai kintamieji yra kokybiniai. Tarkime, turime porinių stebėjimų imtį $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$, gautą stebint kategorinių kintamųjų porą (X,Y). Reikia patikrinti hipotezę, kad kintamieji X ir Y nepriklausomi. Iš duomenų aibės sudarome porinę dažnių lentelę:

	y_1	y_2	…	y_c	Σ
x_1	o_{11}	o_{12}	…	o_{1c}	$n_{1.}$
x_2	o_{21}	o_{22}	…	o_{2c}	$n_{2.}$
…	…	…	…	…	…
x_r	o_{r1}	o_{r2}	…	o_{rc}	$n_{r.}$
Σ	$n_{.1}$	$n_{.2}$	…	$n_{.c}$	n

kur $n_{i.} = \sum_{j=1}^{c} o_{ij}$ - imties narių, kurių požymio X reikšmė yra x_i, skaičius,

$n_{.j} = \sum_{i=1}^{r} o_{ij}$ - imties narių, kurių požymio Y reikšmė yra y_j, skaičius,

$n = \sum_{i=1}^{r} \sum_{j=1}^{c} o_{ij}$ - imties didumas.

Tarkime, $p_{ij}=P(X=x_i,Y=y_j)$ – populiacijos dalis, kuriai matuojamų požymių pora (X,Y) įgyja reikšmę (x_i,y_j), $p_i=P(X=x_i)$ - populiacijos dalis, kuriai požymis X

įgyja reikšmę x_i, o $q_j=P(Y=y_j)$ - populiacijos dalis, kuriai požymis Y įgyja reikšmę y_j: $\sum_{i=1}^{r} p_i = \sum_{j=1}^{c} q_j = 1$

Kintamieji yra nepriklausomi, jei $P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$ su visais i=1,…,r ir j=1,…,c. Todėl tikriname hipotezę:

$$\begin{cases} H_0 : p_{ij} = p_i q_j, \text{ su visais } i = 1,\ldots,r;\ j = 1,\ldots,c \\ H_1 : p_{ij} \neq p_i q_j, \text{ bent vienai porai } (i, j) \end{cases}$$

Jei hipotezė teisinga, tai p_{ij} yra nežinomų parametrų p_i, q_i, i=1,…,r, j=1,…,c, funkcijos. Stebint kintamųjų porą (X,Y), gali įvykti r·c nesutaikomų įvykių, todėl kiekvieną įvykį apibūdina du indeksai i ir j. Kriterijaus statistika lygi:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - n\hat{p}_{ij})^2}{\hat{p}_{ij}}.$$

Nežinomų parametrų įverčiai, gauti χ^2 minimumo metodu, yra:

$\hat{p}_i = \frac{n_{i\cdot}}{n}$, i=1,…,r, $\hat{q}_j = \frac{n_{\cdot j}}{n}$, j=1,…,c.

Tikėtinieji dažniai apskaičiuojami pagal formulę:

$$E_{ij} = n\hat{p}_{ij} = \hat{p}_i \cdot \hat{q}_j = \frac{n_{i\cdot}}{n} \cdot \frac{n_{\cdot j}}{n} = \frac{n_{i\cdot} n_{\cdot j}}{n}.$$

Tarkime, pasirinktas reikšmingumo lygmuo α. Hipotezė H_0 apie kintamųjų nepriklausomumą atmetama, jei statistikos χ^2 reikšmė yra didesnė už χ^2 skirstinio su (r-1)(c-1) laisvės laipsnių α lygmens kritinę reikšmę. χ^2 kriterijaus statistika aproksimuojama χ^2 skirstiniu. Kriterijaus rezultatai laikomi patikimais, jei n ne mažesnis kaip 30 ir bent 75% dažnių lentelės gardelių tikėtini dažniai ne mažesni kaip 5. Visi stebėjimai turi būti nepriklausomi ir kiekvienas stebėjimas turi patekti tik į vieną gardelę.

Logistinė regresija:

Kai turime dvireikšmį kategorinį kintamąjį, jo reikšmei prognozuoti galima taikyti logistinę regresiją. Tarkime, kad kategorinis kintamasis įgyja dvi reikšmes: 0 ir 1. Logistinės regresijos modelis atrodys taip:

$$p_i = \frac{\exp\{z(\vec{x}_i)\}}{1+\exp\{z(\vec{x}_i)\}}, \quad z(\vec{x}_i) = a + b_1 x_{1i} + ... + b_k x_{ki} = \ln\frac{p_i}{1-p_i}.$$

Tikimybių santykis $\frac{p_i}{1-p_i}$ vadinamas galimybe (galimybės įvertinimu) įvykti įvykiui $Y_i = 1$. Įvykio Y=1 galimybė įvykti yra didesnė už 1 tada ir tik tada, kai P(Y=1)>P(Y=0). Logistinė regresija tinka galiojant gana bendroms prielaidoms: kintamieji $X_1,...,X_k$ nebūtinai turi būti normalieji; nereikalaujama normaliai pasiskirsčiusių paklaidų, nekalbama apie priklausomo kintamojo homoskedastiškumą.

Tarkime, duomenis sudaro stebėjimai $(y_i,x_{1i},x_{2i},...,x_{ki})$, i=1,2,...,n, čia y_i yra arba 0, arba 1, o $x_{1i},x_{2i},...$ – intervalinių kintamųjų arba pseudokintamųjų reikšmės. Kaip ir tiesinės regresijos atveju, parametrus reikia parinkti taip, kad logistinės regresijos modelis su turimais duomenimis būtų kuo geriau suderintas. Tam taikomas didžiausiojo tikėtinumo metodas. Kiekvienam stebėjimui skaičiuojama tikimybė p_i , o parametrų įverčiai parenkami taip, kad tikėtinumo funkcija $L = \prod_{i:y_i=1} p_i \prod_{i:y_i=0}(1-p_i)$ būtų maksimali. Parametrų įverčių skaičiavimas – iteracinis procesas, t. y. pasirenkamos pradinės jų reikšmės ir, taikant tam tikrą skaičiavimo metodą, keičiamos tol, kol stabilizuojasi.

Gauti koeficientai naudojami $\hat{z}(\vec{x})$ skaičiuoti. Be to, visada galima rasti tikimybės, kad Y=1, įvertį. Tikimybės P(Y=1) įvertis, kai nepriklausomų kintamųjų reikšmių $x_1,x_2,.,.,x_k$ vektorius yra $\vec{x} = (x_1, x_2, ..., x_k),$, skaičiuojamas pagal formulę

$$\hat{P}(Y=1\,|\,\vec{x}) = \exp\{\hat{z}(\vec{x})\}/(1+\exp\{\hat{z}(\vec{x})\}, \text{ čia } \hat{z}(\vec{x}) = \hat{a} + \hat{b}_1 x_1 + ... + \hat{b}_k x_k.$$

Žinoma, tikimybės įvertis skaičiuojamas tik tokioms x reikšmėms, kurios patenka į duomenų aibės intervalą. Tiriant intervalinius kintamuosius, b_j parodo, kiek padidėja (sumažėja) $\hat{z}(\vec{x})$ reikšmė, kai x_j padidėja vienu vienetu, o likusieji $x_1,...,x_{j-1},x_{j+1},\ldots,x_k$ yra fiksuoti. Pseudokintamųjų atveju skirtingoms b_j reikšmėms gaunamos skirtingos galimybės logaritmo formulės. Tačiau galimybės logaritmo pokyčius sunku interpretuoti. Todėl paprastai skaičiuojamas vadinamasis galimybių santykis $\exp\{\hat{b}_j\}$, kuris parodo, kaip kinta Y galimybė įgyti reikšmę 1.

Žinodami $P(Y=1|\vec{x})$ įvertį, galime nesunkiai nusakyti taisyklę Y reikšmei prognozuoti: jeigu $\hat{P}(Y=1|\vec{x})>0.5,$, tai prognozuojame, kad Y reikšmė yra 1. Priešingu atveju Y reikšmė yra 0. Jeigu mus domina tik Y reikšmės prognozė, visai nebūtina skaičiuoti $P(Y=1|\vec{x})$ įverčio. Iš tikrųjų $\hat{P}(Y=1|\vec{x})>0.5,$ tada ir tik tada, kai $\hat{z}(\vec{x})$ >0. Todėl Y reikšmėms prognozuoti patogiau taikyti tokią taisyklę: į formulę $\hat{z}(\vec{x})=\hat{a}+\hat{b}_1x_1+...+\hat{b}_kx_k$ įstatomos $x_1,x_2,...,x_k$ reikšmės. Jeigu $\hat{z}(\vec{x})$ >0, tai prognozuojama, kad Y=1, priešingu atveju Y=0. Jei $\hat{z}(\vec{x})$ >0, tai rekomenduojama sprendimą priimti metant monetą.

Logistinėje regresijoje prognozuojame tikimybes ir pagal jas sprendžiame, kad Y=1 arba Y=0 (suklystame ar nesuklystame). Todėl logistinėje regresijoje prasmingiau kalbėti ne apie liekamąsias paklaidas, o apie teisingų prognozių procentą. Duomenis sudaro stebėjimai $(y_i,x_{1i},x_{2i},...,x_{ki})$, i=l,2,...,n, ir žinomos y_i reikšmės, todėl galima patikrinti, ar dažnai jos sutampa su reikšmėmis, kurias remdamiesi logistine regresija prognozuotume. Tikrinant, ar logistine regresija prognozėms tinka, šis žingsnis būtinas. Jeigu duomenų, iš kurių gauti parametrų įverčiai, prognozės netikslios, tai logistinės regresijos taikyti negalima, net jei kiti rodikliai bei statistinės išvados rodytų, kad ji tinkama. Nėra griežtų taisyklių, nusakančių, koks teisingumo klasifikavimo procentas laikytinas patenkinamu. Natūralu reikalauti, kad kiekvienos kategorijos jis būtų ne mažesnis kaip 50%, t. y. dažniau teisingai prognozuotume, nei klystume. Bendrasis teisingai atpažintų objektų procentas dar nebūtinai rodo, kad logistine regresija tinkama.

ANOVA (vienfaktorinė dispersinė analizė):

Vienfaktorinė dispersinė analizė (ANOVA) yra Stjudento kriterijaus apibendrinimas keletui nepriklausomų imčių. Dispersinės analizės tikslas – nuspręsti, ar priklausomo kintamojo, išmatuoto skirtingose populiacijose, vidurkiai skiriasi. Tarkime, turime k imčių: (x11,…x1n1),…,(xk1,…,xknk), gautų matuojant neprikl. a. d. $X_1\sim \mathrm{N}(\eta_1,\sigma^2),...,X_k\sim \mathrm{N}(\eta_k,\sigma^2)$. Vidurkių ir dispersijos nežinome. Reikia patikrinti hipotezę:

$H_0 : \mu_1 = \mu_2 = ... = \mu_g$

H1: bent du vidurkiai skiriasi.

Kriterijaus statistika skaičiuojama pagal formulę: $F = \frac{MSB}{MSW}$,

kur $MSW = \frac{(n_1 - 1)S_1^2 + ... + (n_k - 1)S_k^2}{n_1 + ... + n_k - k}$, $MSB = \frac{n_1(\overline{X}_1 - \overline{X}\ )^2 + ... + n_k(\overline{X}_k - \overline{X})^2}{k-1}$.

H_0 atmetama su reikšmingumo lygmeniu α, jeigu gauta empirinė statistika $F > F_\alpha(k-1, n-k)$, čia $F_\alpha(k-1, n-k)$ - Fišerio skirstinys su k-1 ir n-k laisvės laipsnių α lygmens kritinė reikšmė.

Kruskalo-Voliso kriterijus:

Kai netenkinamos ANOVA prielaidos, hipotezė apie dviejų ar daugiau populiacijų skirstinių lygybę tikrinama naudojantis Kruskalo-Voliso kriterijumi – tai yra neparametrinis One-Way ANOVA analogas. Nepriklausomų tolydžiųjų kintamųjų X, Y, Z,... stebėjimai yra $(x_1, x_2, x_3, ..., x_{n1}), (y_1, y_2, ..., y_{n2}), (z_1, z_2, ..., z_{n3}), ...$ Duomenys gauti matavimams naudojant santykių, intervalų arba rangų skalę. Tikriname statistinę hipotezę:

H_0 : kintamųjų skirstiniai yra vienodi,

H_1 : kintamųjų skirstiniai nėra vienodi.

Kriterijaus statistika: $H = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(n+1);$

čia *k* – kintamųjų skaičius, n_j - j-ojo kintamojo stebėjimų skaičius $(n = n_1 + n_2 + ... + n_k)$, R_j^2 - j-osios imties rangų sumos kvadratas. Jei hipotezė H_0 teisinga, tai statistika *H* apitiksliai pasiskirsčiusi pagal χ^2 dėsnį su (k-1) laisvės laipsnių, t.y. $H \approx \chi^2(k-1)$. Tarkime, reikšmingumo lygmuo yra α=0.05. Jei $H > \chi^2_{0.05}(k-1)$, hipotezę apie skirstinių lygybę atmetame. Priešingu atveju H_0 neatmetame.

4. DUOMENŲ ANALIZĖ

4.1. Pagrindinės duomenų charakteristikos

Turime duomenis apie 1005 pacientus, sergančius LHC. Išveskime kiekybinių duomenų pagrindines skaitines charakteristikas:

Pavadinimas	Galiojančios reikšmės	Trūkstamos reikšmės	Vidurkis	Standart. nuokrypis	Min	Max
Amžius	1005	0	49.03	15.42	16	87
KMI	974	31	26.06	4.55	17	47
HAI	719	286	6.20	2.86	1	14
Kepenų dydis	932	73	149.41	18.23	93	230
Blužnies dydis	792	213	117.96	21.82	70	240
ALT	997	8	132.31	145.75	5	2500
AST	998	7	99.71	99.49	8	1460
ŠF	968	37	92.58	56.31	16	719
GGT	969	36	99.57	143.87	6	1908
Hemoglobinas	981	24	137.43	19.03	61	183
Leukocitai	982	23	5.94	5.01	1.3	135
ENG	954	51	13.28	13.73	1	120
TR	957	48	183.68	86.02	4	1314
Cholesterolis	416	589	4.62	1.38	1.57	13.90
DTL	184	821	1.28	0.64	0.2	6.7
MTL	182	823	2.84	1.14	0.4	7.67
Trigliceridai	221	784	1.34	0.85	0.39	7.88
Bilirubinas	877	128	22.31	32.62	3.2	600
T bilirubinas	367	638	16.60	33.07	0	350
N bilirubinas	367	638	20.64	17.93	1.9	250
IK	305	700	0.15	0.11	0.025	1
NBT	156	849	26.13	13.10	2	66
IgA	363	642	2.93	1.86	0.3	11.8
IgG	364	641	16.73	5.92	3.28	46.50
IgM	363	642	1.96	2.08	0.28	33.40

Kokybinių kintamųjų dažnių lentelės:

Kintamasis	Variantai	Skaičius	Procentai
ALKOHOL	0	696	69.3
	1	309	30.7
LYTIS	1	551	54.8
	2	454	45.2
HAI_R	1	177	20.9
	2	460	54.3
	3	189	22.3
	4	21	2.5
FIBROZE	0	3	0.3
	1	143	14.2
	2	313	31.1
	3	179	17.8

Kintamasis	Variantai	Skaičius	Procentai
	4	237	23.6
GENOTIP	1	160	15.9
	2	23	2.3
	3	61	6.1
RNR	0	36	3.6
	1	721	71.7

Pastaba: dėl didelio išsibarstymo ir mažo reikšmių skaičiaus į analizę neįtrauktas kintamasis HCV genotipo potipis. Apžvelgiant duomenis, galima pasakyt, kad 1a potipis nustatytas 4 pacientams (0.4%), 1b potipis – 150 pacientams (14.9%), kombinuotas 1a/1b – 6 pacientams (0.6%), 2a – 2 pacientams (0.2%), 2a/2b – 1 pacientui (0.1%), 2a/2c – 18 pacientų, 3a – 47 pacientams (4.7%).

Pacientai taip pat buvo apklausiami dėl galimo užsikrėtimo HCV būdo. Buvo gauti tokie rezultatai:

Galimas užsikrėtimo būdas	Teigiamai atsakiusių pacientų skaičius	Procentas
Operacijos	723	71.9
Kraujo perpylimas	225	22.4
Donorystė	192	19.1
Sunkios traumos	177	17.6
Ilgalaikis gydymas	139	13.8
Darbo metu	59	5.9
Tatuiruočių darymasis	50	5.0
Narkomanija	48	4.8
Kontaktas su kitais ligoniais	18	1.8
Hemodializės procedūros	17	1.7
Nežino	74	7.4

4.2. Alkoholio vartojimo įtakos nustatymas

Alkoholio vartojimas, kaip žinia, neigiamai veikia kepenų būklę net ir sveikiems žmoniems, o ligoniams, sergantiems hepatitu C, jis dar labiau pažeidžia viruso jau ir taip paveiktas kepenis ir tuo labai pablogina ligos eigą. Pamėginsime rasti priklausomybę tarp alkoholio vartojimo ir kitų rodiklių. Suskirstom pacientus į dvi grupes: į vartojančius alkoholinius gėrimus (309 pacientai, 30.7%) ir nevartojančius (696 pacientai, 69.3%).

Kiekybinių duomenų normališkumui tikrinti naudosim Kolmogorovo-Smirnovo testą vienai imčiai imant atskirai vartojančius ir nevartojančius alkoholį

pacientus. Šis testas tikrina nulinę hipotezę, kad duomenų skirstinys nesiskiria nuo normalaus skirstinio. Pasirenkame reikšmingumo lygmenį alpha=0.01. Jei gauta p-reikšmė mažesnė už alpha, tai nulinę hipotezę atmetame, t. y. duomenys nepasiskirstę pagal normalų skirstinį. Kintamiesiems, pasiskirsčiusiems pagal normalų skirstinį, pritaikysime t-testą, o likusiems – neparametrinį Mano-Vitnio-Vilkoksono testą. T-testas tikrina nulinę hipotezę apie dviejų grupių vidurkių lygybę, o Mano-Vitnio-Vilkoksono testas – apie skirstinių lygybę. T-testo ir Mano-Vitnio-Vilkoksono testo atveju pasirenkame reikšmingumo lygmenį alpha=0.1. Jei p-reikšmė mažesnė už alpha, tai kintamasis statistiškai reikšmingai skiriasi tarp vartojančių ir nevartojančių alkoholį pacientų. SPSS pagalba gauname tokias p-reikšmės:

Kintamojo pavadinimas	**Kolmogorovo-Smirnovo testas vienai imčiai**		**T-testas**	**Mano-Vitnio-Vilkoksono testas**
	Nevartoja alkoholio	**Vartoja alkoholį**		
Amžius	0.007	**0.517**	-	**0.000**
KMI	0.000	**0.039**	-	0.185
HAI	0.000	0.001	-	**0.000**
Kepenų dydis	**0.027**	**0.187**	**0.000**	-
Blužnies dydis	0.000	**0.120**	-	0.112
ALT	0.000	0.000	-	**0.000**
AST	0.000	0.000	-	**0.000**
ŠF	0.000	0.000	-	**0.000**
GGT	0.000	0.000	-	**0.000**
Hemoglobinas	**0.014**	0.000	-	**0.000**
Leukocitai	0.000	0.000	-	**0.000**
ENG	0.000	0.000	-	0.300
TR	**0.010**	**0.328**	0.608	-
Cholesterolis	**0.426**	**0.131**	0.154	-
DTL	**0.038**	**0.437**	**0.072**	-
MTL	**0.212**	**0.785**	0.384	-
Trigliceridai	0.001	0.006	-	**0.023**
Bilirubinas	0.000	0.000	-	**0.002**
T bilirubinas	0.000	0.000	-	**0.000**
N bilirubinas	0.000	0.000	-	0.918
IK	0.000	0.000	-	0.110
NBT	**0.014**	**0.339**	0.268	-
IgA	0.000	0.004	-	**0.003**
IgG	0.001	**0.101**	-	0.329
IgM	0.000	**0.047**	-	0.781

Norėdami sužinoti, koks būtent tas skirtumas, t-testo atveju palyginame abiejų grupių vidurkius, o neparametrinio Mano-Vitnio-Vilkoksono testo atveju – grupių vidutinius rangus:

Vidurkių lentelė:

Kintamojo pavadinimas	Alkoholio vartojimas	N	Mean	Std. Deviation	Std. Error Mean
Kepenų dydis	Nevartoja	639	146.23	17.018	.673
	Vartoja	293	156.33	18.871	1.102
DTL cholesterolis	Nevartoja	127	1.335	.6862	.0609
	Vartoja	57	1.151	.5131	.0680

Vidutinių rangų lentelė:

Kintamojo pavadinimas	Alkoholio vartojimas	N	Mean Rank	Sum of Ranks
Amžius	Nevartoja	696	528.57	367883.50
	Vartoja	309	445.41	137631.50
Histologinis aktyvumo indeksas	Nevartoja	496	337.81	167554.00
	Vartoja	223	409.35	91286.00
Kepenų fermentas ALT	Nevartoja	690	473.64	326815.00
	Vartoja	307	555.99	170688.00
Kepenų fermentas AST	Nevartoja	691	465.59	321724.50
	Vartoja	307	575.82	176776.50
Kepenų fermentas ŠF	Nevartoja	667	461.67	307934.00
	Vartoja	301	535.09	161062.00
Kepenų fermentas GGT	Nevartoja	667	408.30	272338.50
	Vartoja	302	654.39	197626.50
Hemoglobinas	Nevartoja	681	459.77	313104.00
	Vartoja	300	561.89	168567.00
Leukocitai	Nevartoja	682	459.71	313525.50
	Vartoja	300	563.76	169127.50
Trigliceridai	Nevartoja	150	104.29	15643.00
	Vartoja	71	125.18	8888.00
Bendras bilirubinas	Nevartoja	603	421.17	253963.50
	Vartoja	274	478.25	131039.50
Tiesioginis bilirubinas	Nevartoja	237	168.95	40041.00
	Vartoja	130	211.44	27487.00
Imunoglobulinas-a	Nevartoja	241	170.39	41064.50
	Vartoja	122	204.93	25001.50

Taigi iš lentelių matome, kad pas vartojančius alkoholį padidėję šie rodikliai: kepenų dydis, histologinis aktyvumo indeksas, kepenų fermentai (ALT, AST, ŠF, GGT), hemoglobinas, leukocitai, trigliceridai, bilirubinai (bendras, tiesioginis) ir imunoglobulinas-a. Pas vartojančius sumažėjęs tik DTL-cholesterolis. Taip pat matom, kad alkoholį vartoja jaunesni pacientai.

Kokybiniams duomenims tikrinti panaudosime χ^2 nepriklausomumo testą, kuris tikrina nulinę hipotezę, kad požymiai yra nepriklausomi. Pasirenkame reikšmingumo lygmenį alpha=0.01.

SPSS pagalba gauname tokias p-reikšmes:

Kintamojo pavadinimas	p-reikšmė
Lytis	0.000
Fibrozės laipsnis	0.004
Genotipas	0.922
Ribonukleo rūgštis	0.563

Matome, kad hipotezę apie požymių nepriklausomumą (t. y. požymiai susiję) galima atmesti dviems kintamiesiems:

⇒ Lytis (tarp vartojančių alkoholį 87.7% vyrai ir tik 12.3% moterys).

⇒ Fibrozės laipsnis (tarp vartojančių daugiau turi kepenų fibrozės laipsnį 3 ir 4, nei tarp nevartojančių).

4.3. Kepenų cirozės įtakoti rodiklių pakitimai

Pradžioje reikia suskirstyti pacientus į tuos, kuriems jau išsivysčiusi kepenų cirozė ir kuriems jos dar nėra. Tam panaudosime kintamąjį FIBROZĖ, kuris parodo kepenų fibrozės laipsnį pagal METAVIR sistemą. Remiantis ja, didžiausias, t. y. 4 kepenų fibrozės laipsnis reiškia kepenų cirozę. Taigi sukuriame naują kintamąjį CIROZE, kuris įgyja dvi reikšmes: pacientams su ketvirtu fibrozės laipsniu jis lygus “1” (237 pacientai, 23.6%), o visiems kitiems jis lygus “0” (768 pacientus, 76.4%). Kaip ir 4.2. skyrelyje, Kolmogorovo-Smirnovo testo vienai imčiai pagalba patikrinsime duomenų normališkumą ir kintamiesiems, kuriems ši sąlyga tenkinama, pritaikysime t-testą, o likusiems – Mano-Vitnio-Vilkoksono kriterijų. Reikšmingumo lygmuo Kolmogorovo-Smirnovo testo atveju alpha=0.01, o t-testo ir Mano-Vitnio Vilkoksono testo atveju alpha=0.1.

Gauname tokias p-reikšmes:

Kintamojo pavadinimas	Kolmogorovo-Smirnovo testas vienai imčiai		T-testas	Mano-Vitnio-Vilkoksono testas
	Su ciroze	Be cirozės		
Amžius	0.000	**0.035**	-	**0.000**
KMI	**0.053**	0.000	-	**0.000**
HAI	**0.061**	0.000	-	**0.033**
Kepenų dydis	**0.673**	0.008	-	0.190
Blužnies dydis	**0.179**	0.000	-	**0.000**
ALT	**0.042**	0.000	-	**0.017**
AST	0.000	0.000	-	**0.000**
ŠF	**0.012**	0.000	-	**0.000**

Kintamojo pavadinimas	Kolmogorovo-Smirnovo testas vienai imčiai		T-testas	Mano-Vitnio-Vilkoksono testas
	Su ciroze	Be cirozės		
GGT	0.000	0.000	-	**0.000**
Hemoglobinas	**0.195**	**0.094**	**0.000**	-
Leukocitai	0.000	0.000	-	**0.000**
ENG	0.000	0.000	-	**0.000**
TR	0.002	0.000	-	**0.000**
Cholesterolis	**0.066**	**0.119**	**0.000**	-
DTL	**0.464**	**0.028**	**0.016**	-
MTL	**0.498**	**0.334**	**0.065**	-
Trigliceridai	**0.091**	0.000	-	0.444
Bilirubinas	0.000	0.000	-	**0.000**
T bilirubinas	0.000	0.000	-	**0.000**
N bilirubinas	0.000	0.000	-	**0.071**
IK	0.009	0.000	-	0.262
NBT	**0.572**	0.007	-	0.746
IgA	**0.020**	0.000	-	**0.000**
IgG	**0.052**	**0.016**	**0.000**	-
IgM	**0.012**	0.000	-	**0.002**

Kintamiesiems, kurių skirtumas tarp grupių statistiškai reikšmingas, t-testo atveju palyginame abiejų grupių vidurkius, o neparametrinio Mano-Vitnio-Vilkoksono testo atveju – grupių vidutinius rangus:

Vidurkių lentelė (t-testo atveju):

Kintamojo pavadinimas	Kepenų cirozė	N	Mean	Std. Deviation	Std. Error Mean
Hemoglobinas	Nėra	746	141.24	16.582	.607
	Yra	235	125.34	21.165	1.381
Bendras cholesterolis	Nėra	307	4.7866	1.39780	.07978
	Yra	109	4.1437	1.23551	.11834
DTL cholesterolis	Nėra	136	1.345	.6849	.0587
	Yra	48	1.087	.4547	.0656
MTL cholesterolis	Nėra	134	2.9306	1.16991	.10106
	Yra	48	2.5775	1.00905	.14564
Imunoglobulinas g	Nėra	276	15.7472	5.07833	.30568
	Yra	88	19.8320	7.21722	.76936

Vidutinių rangų lentelė:

Kintamojo pavadinimas	Kepenų cirozė	N	Mean Rank	Sum of Ranks
Amžius	Nėra	768	438.43	336715.00
	Yra	237	712.24	168800.00
Kūno masės indeksas	Nėra	751	464.70	348989.00
	Yra	223	564.29	125836.00
Histologinis aktyvumo indeksas	Nėra	637	354.12	225571.50
	Yra	82	405.71	33268.50
Blužnies dydis	Nėra	588	334.91	196924.50
	Yra	221	591.50	130720.50
Kepenų fermentas ALT	Nėra	763	511.01	389904.00

Kintamojo pavadinimas	Kepenų cirozė	N	Mean Rank	Sum of Ranks
	Yra	234	459.82	107599.00
Kepenų fermentas AST	Nėra	763	455.30	347395.50
	Yra	235	643.00	151105.50
Kepenų fermentas ŠF	Nėra	734	422.86	310376.50
	Yra	234	677.86	158619.50
Kepenų fermentas GGT	Nėra	741	450.30	333669.50
	Yra	228	597.79	136295.50
Leukocitai	Nėra	746	521.54	389068.50
	Yra	236	396.54	93584.50
Eritrocitų nusėdimo greitis	Nėra	722	422.22	304841.50
	Yra	232	649.54	150693.50
TR	Nėra	727	559.58	406812.50
	Yra	230	224.31	51590.50
Bendras bilirubinas	Nėra	656	383.27	251427.50
	Yra	221	604.41	133575.50
Tiesioginis bilirubinas	Nėra	213	146.46	31196.50
	Yra	154	235.92	36331.50
Netiesioginis bilirubinas	Nėra	213	175.49	37380.00
	Yra	154	195.77	30148.00
Imunoglobulinas-a	Nėra	275	165.12	45409.00
	Yra	88	234.74	20657.00
Imunoglobulinas-m	Nėra	275	172.28	47376.00
	Yra	88	212.39	18690.00

Gavosi, kad daugelis rodiklių skiriasi, tarp pacientų su kepenų ciroze ir be. Pas turinčius cirozę didesni šie rodikliai: amžius, KMI, HAI, blužnies dydis, kepenų fermentai (AST, ŠF, GGT), eritrocitų nusėdimo greitis, bilirubinai (bendras, tiesioginis, netiesioginis) ir imunoglobulinai (G ir M). O sumažėję šie rodikliai: kepenų fermentas AST, hemoglobinas, leukocitai, trombocitai, cholesterolis (bendras, DTL ir MTL) bei imunoglobulinas-a.

Kepenų cirozę galima nustatyti tik atlikus kepenų biopsiją, t. y. paėmus nedidelį kepenų mėginį ir išnagrinėjus jį per mikroskopą. Iš t-testo ir Mano-Vitnio-Vilkoksono kriterijaus rezultatų matome, kad daugelis kraujo biocheminių rodiklių statistiškai reikšmingai skiriasi tarp pacientų, kuriems jau yra kepenų cirozė ir kuriems ji dar neišsivysčiusi. Pamėginkime nustatyt, ar pagal kraujo tyrimus galima prognozuoti, kad žmogui jau yra išsivysčiusi kepenų cirozė. Tam panaudosime logistinę regresiją. Priklausomas kintamasis bus CIROZE, o nepriklausomais imsim visus, kuriems nustatėme statistiškai reikšmingą skirtumą. Dėl didelio tuščių reikšmių skaičiaus teko atmesti kintamuosius DTL_CHOL, MTL_CHOL, T_BILIR, N_BILIR.

SAS paketo pagalba, atlikus logistinę regresiją, gavosi tokie rezultatai:

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.0693	2.4465	1.5739	0.2096
ALT	1	**-0.0130**	0.00492	7.0176	**0.0081**
AST	1	**0.00914**	0.00522	3.0654	**0.0800**
ŠF	1	**0.0102**	0.00491	4.2861	**0.0384**
GGT	1	0.00192	0.00170	1.2774	0.2584
HB	1	-0.0110	0.0174	0.4033	0.5254
LEU	1	0.0445	0.0477	0.8694	0.3511
ENG	1	**0.0526**	0.0246	4.5690	**0.0326**
TR	1	**-0.0225**	0.00495	20.7281	**<.0001**
B_CHOLES	1	**-0.4075**	0.2323	3.0760	**0.0795**
BILIRUB	1	**0.0163**	0.00981	2.7476	**0.0974**
IGA	1	0.2111	0.1380	2.3388	0.1262
IGG	1	**-0.00113**	0.0462	0.0006	**0.9806**
IGM	1	**-0.0743**	0.1995	0.1388	**0.7095**

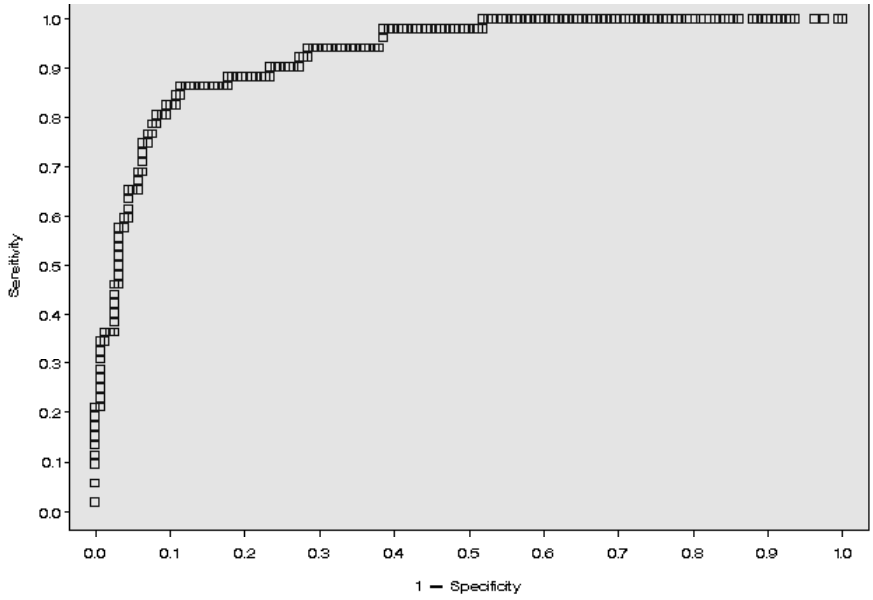
Taigi cirozę įtakoja šie rodikliai: kepenų fermentus ALT, AST ir ŠF, eritrocitų nusėdimo greitį, trombocitus, bendrą cholesterolį, imunoglobulinus G ir M. Šiems rodikliams paskaičiuoti šansų santykiai yra tokie:

Effect	Point Estimate	95% Wald Confidence Limits	
ALT	0.987	0.978	0.997
AST	1.009	0.999	1.020
_F	1.010	1.001	1.020
ENG	1.054	1.004	1.106
TR	0.978	0.968	0.987
B_CHOLES	0.665	0.422	1.049
BILIRUB	1.016	0.997	1.036
IGG	0.999	0.912	1.093
IGM	0.928	0.628	1.373

Atlikus klasifikavimą gavome tokią lentelę:

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.100	45	109	49	7	73.3	86.5	69.0	52.1	6.0
0.200	43	126	32	9	80.5	82.7	79.7	42.7	6.7
0.300	41	136	22	11	84.3	78.8	86.1	34.9	7.5
0.400	39	140	18	13	85.2	75.0	88.6	31.6	8.5
0.500	34	144	14	18	84.8	65.4	91.1	29.2	11.1
0.600	28	146	12	24	82.9	53.8	92.4	30.0	14.1
0.700	24	150	8	28	82.9	46.2	94.9	25.0	15.7
0.800	20	150	8	32	81.0	38.5	94.9	28.6	17.6
0.900	14	152	6	38	79.0	26.9	96.2	30.0	20.0

Matome, kad didėjant slenksčiui, jautrumas mažėja, o specifiškumas didėja. Kai slenksčio reikšmės kinta intervale 0.2 – 0.4, gauname labai gerus klasifikavimo rezultatus. Tai iliustruoja žemiau pateikiama ROC kreivė:



Nors klasifikavimas buvo atliekamas tai pačiai apmokymo imčiai ir nebuvo bandoma klasifikuoti testinės imties, yra pagrindo teigti, kad modelis tinkamas cirozės prognozavimui pagal kraujo tyrimų rodiklius.

4.4. Palyginimas tarp skirtingų HAI kategorijų

Duomenis analizuojame suskirstę juos į 4 kategorijas pagal histologinį aktyvumo indeksą (HAI):

⇒ 1 kategorija – labai nedidelio aktyvumo uždegimas, kai HAI nuo 1 iki 3.

⇒ 2 kategorija – nedidelio aktyvumo uždegimas, kai HAI nuo 4 iki 8.

⇒ 3 kategorija – vidutinio aktyvumo uždegimas, kai HAI nuo 9 iki 12.

⇒ 4 kategorija – didelio aktyvumo uždegimas, kai HAI nuo 13 iki 18.

Kiekvienai kategorijai atskirai Kolmogorovo-Smirnovo testo pagalba patikriname duomenų normališkumo sąlygą. Pasirenkame reikšmingumo lygmenį alpha=0.01. Jei gauta p-reikšmė didesnė už alpha, tai duomenys pasiskirstę pagal normalų skirstinį. Kintamiesiems, kurių duomenys visose 4 kategorijose pasiskirstę pagal normalų skirstinį, pritaikome ANOVA

(vienfaktorinę dispersinę analizę). Ji tikrina nulinę hipotezę, kad visų 4 imčių (kategorijų) vidurkiai statistiškai reikšmingai nesiskiria. Pasirenkame reikšmingumo lygmenį alpha=0.1. Jei gauta p-reikšmė mažesnė už alpha, tai nulinę hipotezę atmetam, t. y. bent 2 imčių (kategorijų) vidurkiai statistiškai reikšmingai skiriasi. Kintamiesiems, kurie nepasiskirstę pagal normalų skirstinį, pritaikome Kruskalo-Voliso testą. Jis tikrina nulinę hipotezę, kad kintamųjų skirstiniai imtyse (kategorijose) vienodi. Pasirenkame reikšmingumo lygmenį alpha=0.1. Jei gauta p-reikšmė mažesnė už alpha, tai nulinę hipotezę atmetam, t. y. kintamųjų skirstiniai skirtingose imtyse (kategorijose) statistiškai reikšmingai skiriasi.

Gaunam tokias p-reikšmes:

Kintamojo Pavadinimas	**Kolmogorovo-Smirnovo testas vienai imčiai**				**ANOVA**	**Kruskalo-Voliso testas**
	1 kat.	**2 kat.**	**3 kat.**	**4 kat.**		
Amžius	**0.524**	**0.156**	**0.139**	**0.712**	**0.005**	-
KMI	**0.022**	0.000	**0.027**	**0.337**	-	**0.054**
Kepenų dydis	**0.176**	**0.110**	**0.866**	**0.833**	**0.000**	-
Blužnies dydis	**0.200**	0.000	0.007	**0.269**	-	0.567
ALT	0.000	0.000	**0.039**	**0.301**	-	**0.000**
AST	0.000	0.000	0.000	**0.190**	-	**0.000**
ŠF	0.000	0.000	0.003	**0.460**	-	**0.000**
GGT	0.000	0.000	0.000	**0.020**	-	**0.000**
Hemoglobinas	**0.651**	**0.298**	**0.395**	**0.825**	0.106	-
Leukocitai	0.000	0.005	**0.030**	**0.497**	-	0.981
ENG	0.000	0.000	0.000	**0.484**	-	0.308
TR	**0.612**	0.000	**0.262**	**0.958**	-	**0.000**
Cholesterolis	**0.835**	**0.406**	**0.852**	**0.189**	**0.000**	-
DTL	**0.871**	**0.056**	**0.411**	**0.934**	0.583	-
MTL	**0.825**	**0.461**	**0.849**	**0.412**	**0.067**	-
Trigliceridai	**0.048**	0.000	**0.255**	**0.858**	-	**0.093**
Bilirubinas	0.000	0.000	0.000	**0.061**	-	**0.000**
T bilirubinas	**0.030**	0.000	0.000	**0.630**	-	**0.000**
N bilirubinas	**0.124**	0.000	**0.229**	**0.800**	-	0.689
IK	0.001	0.001	0.002	**0.687**	-	0.122
NBT	**1.000**	**0.054**	**0.512**	**0.793**	0.257	-
IgA	**0.051**	**0.016**	**0.023**	**0.408**	**0.005**	-
IgG	**0.718**	**0.084**	**0.965**	**0.174**	**0.003**	-
IgM	**0.050**	0.009	0.000	**0.156**	-	0.600

Norint nustatyti, kurių būtent kategorijų vidurkiai statistiškai reikšmingai skiriasi, naudojami poriniai Post-Hock testai. Fišerio mažiausiai reikšmingo skirtumo LSD post-hoc testas tikrina vidurkių skirtumą naudodamas t-testą, imant visas imčių poras atskirai. Kiekvienai porai yra paskaičiuojama p-reikšmė,

kuri parodo, ar vidurkių skirtumas statistiškai reikšmingas tarp dviejų kategorijų. Taip yra, kai p-reikšmė mažesnė už alpha=0,1.

Gaunam tokius rezultatus:

Kintamasis	HAI kategorijos	Vidurkio skirtumas	p-reikšmė
Amžius	1-2	-3.13	**0.017**
	1-3	-5.27	**0.001**
	1-4	-6.64	**0.052**
	2-3	-2.14	**0.095**
	2-4	-3.51	0.288
	3-4	-1.37	0.688
Kepenų dydis	1-2	-1.23	0.440
	1-3	-7.46	**0.000**
	1-4	-5.59	0.190
	2-3	-6.23	**0.000**
	2-4	-4.36	0.291
	3-4	1.87	0.659
Cholesterolis	1-2	0.5658	**0.003**
	1-3	0.5217	**0.017**
	1-4	-1.0005	**0.024**
	2-3	-0.0441	0.806
	2-4	-1.5664	**0.000**
	3-4	-1.5222	**0.001**
MTL cholesterolis	1-2	0.4657	**0.053**
	1-3	0.5106	**0.054**
	1-4	-0.2510	0.553
	2-3	0.0449	0.840
	2-4	-0.7167	**0.074**
	3-4	-0.7616	**0.068**
IgA	1-2	-0.196	0.430
	1-3	-0.850	**0.002**
	1-4	-0.267	0.585
	2-3	-0.653	**0.002**
	2-4	-0.071	0.878
	3-4	0.582	0.219
IgG	1-2	-2.4432	**0.002**
	1-3	-1.5613	**0.059**
	1-4	-4.4253	**0.004**
	2-3	0.8819	0.180
	2-4	-1.9821	0.165
	3-4	-2.8640	**0.050**

Norėdami nustatyti skirtumą tarp grupių Kruskalo-Voliso testo atveju, palyginame skirtingų kategorijų vidutinius rangus. Priminsime, kad šiuo atveju hipotezė formuluojama pasiskirstymo funkcijoms, o ne vidurkiams. T. y. tikriname, ar kintamųjų skirstiniai skiriasi postūmio parametru grupėse. Testas yra galingiausias būtent tokiu atveju.

Vidutinių rangų lentelė:

Kintamojo pavadinimas	Histologinis aktyvumo indeksas (rang.)	N	Mean Rank
Kūno masės indeksas	1	174	389.70
	2	449	406.97
	3	183	455.09
	4	21	407.55
Kepenų fermentas ALT	1	174	199.27
	2	457	427.27
	3	187	579.53
	4	21	670.17
Kepenų fermentas AST	1	174	199.47
	2	458	419.61
	3	187	595.58
	4	21	712.17
Kepenų fermentas ŠF	1	166	334.72
	2	440	386.43
	3	183	492.70
	4	21	604.55
Kepenų fermentas GGT	1	167	277.03
	2	443	401.43
	3	184	523.61
	4	21	575.24
TR	1	173	462.18
	2	432	394.62
	3	180	356.15
	4	19	477.39
Trigliceridai	1	42	96.17
	2	87	92.75
	3	48	84.94
	4	9	133.94
Bendras bilirubinas	1	152	306.55
	2	386	365.03
	3	169	395.00
	4	20	518.75
Tiesioginis bilirubinas	1	42	90.65
	2	140	125.59
	3	64	152.74
	4	13	192.65

Atlikus χ^2 nepriklausomumo testą kokybiniams kintamiesiems, gautos tokios p-reikšmės:

Kintamojo pavadinimas	p-reikšmė
Fibroze	0.000
Lytis	0.208
RNR	0.150

Dėl per mažo reikšmių skaičiaus į analizę neįtrauktas kintamasis GENOTIPAS. Taigi, iš kokybinių kintamųjų tik tarp HAI ir fibrozes egzistuoja priklausomybė.

4.5. Kitų rodiklių ryšio nustatymas

Kintamųjų ryšio nustatymui naudojamas Spirmeno ranginės koreliacijos koeficientas. Jis parodo, ar du kintamieji koreliuoja ir taip pat įvertina tos koreliacijos stiprumą. Pasirenkame reikšmingumo lygmenį alpha=0.01 ir tikriname nulinę hipotezę, kad požymiai nekoreliuoja. Jei p-reikšmė mažesnė už alpha, tai nulinę hipotezę atmetame, t. y. tarp dviejų dydžių egzistuoja statistiškai reikšminga priklausomybė. Kuo koreliacijos koeficientas didesnis, tuo priklausomybė yra stipresnė. Teigiamas ženklas prie koeficiento parodo tiesioginę priklausomybę, neigiamas – netiesioginę. Koreliacijos stiprumą vertinsime tokiu principu:

- ⇒ Kai koreliacijos koeficientas < 0.2, tai koreliacija labai silpna.
- ⇒ Kai 0.2 <= koreliacijos koeficientas < 0.4, tai koreliacija silpna.
- ⇒ Kai 0.4 <= koreliacijos koeficientas < 0.6, tai koreliacija vidutinė.
- ⇒ Kai 0.6 <= koreliacijos koeficientas < 0.8, tai koreliacija stipri.
- ⇒ Kai koreliacijos koeficientas >= 0.8, tai koreliacija labai stipri.

SPSS paketo pagalba buvo gauti koreliacijos koeficientai ir juos atitinkančios p-reikšmės visoms nenominalių kintamųjų poroms (pilną lentelę galima pažiūrėti priede nr. 2). Labai stipri koreliacija gavosi tik tarp bendro cholesterolio ir MTL cholesterolio (koeficientas: 0.864, p-reikšmė: 0.000). Stipri koreliacija gavosi tarp šių kintamųjų:

Kintamasis-1	Kintamasis-2	Koreliacijos koeficientas	p-reikšmė
ALT	AST	0.790	0.000
Bilirubinas	Tiesioginis bilirubinas	0.727	0.000
Bilirubinas	Netiesioginis bilirubinas	0.729	0.000
ŠF	Tiesioginis bilirubinas	0.602	0.000
Hemoglobinas	Eritrocitų nusėdimo greitis	-0.615	0.000

O vidutinio stiprumo koreliacija gavosi tarp šių kintamųjų:

Kintamasis-1	Kintamasis-2	Koreliacijos koeficientas	p-reikšmė
Amžius	Fibrozė	0.485	0.000
Amžius	Eritrocitų nusėdimo greitis	0.412	0.000
Fibrozė	Blužnies dydis	0.486	0.000
Fibrozė	AST	0.445	0.000

Kintamasis-1	Kintamasis-2	Koreliacijos koeficientas	p-reikšmė
Fibrozė	ŠF	0.463	0.000
Fibrozė	Trombocitai	-0.564	0.000
Fibrozė	Tiesioginis bilirubinas	0.475	0.000
Fibrozė	Histologinis aktyvumo indeksas	0.428	0.000
Blužnies dydis	Trombocitai	-0.468	0.000
ALT	GGT	0.422	0.000
ALT	Histologinis aktyvumo indeksas	0.538	0.000
AST	ŠF	0.415	0.000
AST	GGT	0.546	0.000
AST	Tiesioginis bilirubinas	0.422	0.000
AST	Histologinis aktyvumo indeksas	0.567	0.000
ŠF	GGT	0.481	0.000
GGT	Tiesioginis bilirubinas	0.411	0.000
Leukocitai	Trombocitai	0.430	0.000
Eritrocitų nusėdimo greitis	Imunoglobulinas-g	0.414	0.000
Bendras cholesterolis	DTL cholesterolis	0.410	0.000
Imunoglobulinas-a	Imunoglobulinas-g	0.416	0.000

Kitoms kintamųjų poroms koreliacija gavosi silpna, labai silpna arba dydžiai nepriklausomi.

5. IŠVADOS

Ištyrus 1005 pacientų, sergančių hepatitu C, histologinius, morfologinius bei biocheminių kraujo tyrimų duomenis, buvo nustatyta statistiškai reikšminga priklausomybė tarp alkoholio vartojimo ir kepenų būklės, taip pat pas vartojančius alkoholį pastebimi pakitę daugelis kraujo rodiklių.

Atrinkus ligonius, kuriems jau yra išsivysčiusi kepenų cirozė, nustatyti cirozę įtakojantys kraujo biocheminiai rodikliai ir jiems pritaikyta logistinė regresija, siekiant patikrinti galimybę prognozuoti kepenų cirozės egzistavimą pagal kraujo tyrimų duomenis. Klasifikavimo procedūra davė labai gerus rezultatus, su aukštais jautrumo ir specifiškumo lygiais.

Palyginus skirtingas histologinio aktyvumo indekso kategorijas, nustatyti rodikliai, kurie kinta ligos progresavimo metu, t. y. esant skirtingiems kepenų uždegimo lygiams.

Taip pat buvo paskaičiuoti koreliacijų koeficientai tarp visų nenominalių kintamųjų ir rasti nauji koreliuojantys požymiai iki tol nepaminėti šiame darbe.

6. LITERATŪRA

1. V. Čekanavičius, G. Murauskas “Statistika ir jos taikymai, I dalis”, 2000.
2. V. Čekanavičius, G. Murauskas “Statistika ir jos taikymai, II dalis”, 2002.
3. А. Бююль, П. Цефель “SPSS: искусство обработки информации”, 2001.
4. “Sveikatos enciklopedija”, 2003.
5. Vytautas Kazakevičius “Kokybinių duomenų analizė”, 2005.
6. Filomena Barokienė “Virusinis hepatitas C”, 2004.

7. PRIEDAI

(PRIEDAS NR. 1)

The SAS System 16:03 Sunday, May 26, 2002 1

The LOGISTIC Procedure

Model Information

Data Set	WORK.DUOMENYS	
Response Variable	CIROZE	CIROZE
Number of Response Levels	2	
Number of Observations	210	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	CIROZE	Total Frequency
1	0	158
2	1	52

Probability modeled is CIROZE=1.

NOTE: 795 observations were deleted due to missing values for the response or explanatory variables.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	237.076	146.849
SC	240.423	193.709
-2 Log L	235.076	118.849

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	116.2268	13	<.0001
Score	87.1190	13	<.0001
Wald	46.2780	13	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.0693	2.4465	1.5739	0.2096
ALT	1	-0.0130	0.00492	7.0176	0.0081
AST	1	0.00914	0.00522	3.0654	0.0800
_F	1	0.0102	0.00491	4.2861	0.0384
GGT	1	0.00192	0.00170	1.2774	0.2584
HB	1	-0.0110	0.0174	0.4033	0.5254
LEU	1	0.0445	0.0477	0.8694	0.3511
ENG	1	0.0526	0.0246	4.5690	0.0326
TR	1	-0.0225	0.00495	20.7281	<.0001
B_CHOLES	1	-0.4075	0.2323	3.0760	0.0795
BILIRUB	1	0.0163	0.00981	2.7476	0.0974
IGA	1	0.2111	0.1380	2.3388	0.1262
IGG	1	-0.00113	0.0462	0.0006	0.9806
IGM	1	-0.0743	0.1995	0.1388	0.7095

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
ALT	0.987	0.978	0.997
AST	1.009	0.999	1.020
_F	1.010	1.001	1.020
GGT	1.002	0.999	1.005
HB	0.989	0.956	1.023
LEU	1.045	0.952	1.148
ENG	1.054	1.004	1.106
TR	0.978	0.968	0.987
B_CHOLES	0.665	0.422	1.049
BILIRUB	1.016	0.997	1.036
IGA	1.235	0.942	1.619
IGG	0.999	0.912	1.093
IGM	0.928	0.628	1.373

Association of Predicted Probabilities and Observed Responses

Percent Concordant	93.0	Somers' D	0.860
Percent Discordant	7.0	Gamma	0.860
Percent Tied	0.1	Tau-a	0.322
Pairs	8216	c	0.930

The LOGISTIC Procedure

Classification Table

Prob Level	Correct Event	Correct Non-Event	Incorrect Event	Incorrect Non-Event	Percentages Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.100	45	109	49	7	73.3	86.5	69.0	52.1	6.0
0.200	43	126	32	9	80.5	82.7	79.7	42.7	6.7
0.300	41	136	22	11	84.3	78.8	86.1	34.9	7.5
0.400	39	140	18	13	85.2	75.0	88.6	31.6	8.5
0.500	34	144	14	18	84.8	65.4	91.1	29.2	11.1
0.600	28	146	12	24	82.9	53.8	92.4	30.0	14.1
0.700	24	150	8	28	82.9	46.2	94.9	25.0	15.7
0.800	20	150	8	32	81.0	38.5	94.9	28.6	17.6
0.900	14	152	6	38	79.0	26.9	96.2	30.0	20.0

(PRIEDAS Nr. 2)

Spirmeno koreliacijos koeficientų lentelė

		amzius	kmi	fibroze	kepenys	bluznis	alt	ast	šf	ggt	hb	leu	eng	tr	B_choles	Dtl_chol	Mtl_chol	triglic	bilirub	T_bilir	N_bilir	CIK	NBT	iga	igg	igm	Hai_r
amzius	Corr. Coeffic.	1.000	.346 (**)	.485 (**)	-.041	.189 (**)	-.083 (**)	.172 (**)	.286 (**)	.137 (**)	-.253 (**)	-.207 (**)	.412 (**)	-.380 (**)	-.073	.077	-.132	.044	.183 (**)	.211 (**)	-.043	.036	-.079	.236 (**)	.270 (**)	.086	.122 (**)
	Sig.	.	.000	.000	.210	.000	.009	.000	.000	.000	.000	.000	.000	.000	.139	.298	.075	.513	.000	.000	.412	.526	.327	.000	.000	.101	.000
kmi	Corr. Coeffic.	.346 (**)	1.000	.261 (**)	.153 (**)	.260 (**)	.060	.129 (**)	.135 (**)	.153 (**)	.090 (**)	-.097 (**)	.139 (**)	-.196 (**)	.017	-.204 (**)	-.070	.159 (*)	.106 (**)	.003	.034	.126 (*)	-.025	.145 (**)	.176 (**)	-.076	.085 (*)
	Sig.	.000	.	.000	.000	.000	.063	.000	.000	.000	.006	.003	.000	.000	.736	.007	.358	.020	.002	.953	.528	.033	.766	.007	.001	.155	.014
fibroze	Corr. Coeffic.	.485 (**)	.261 (**)	1.000	.150 (**)	.486 (**)	.062	.445 (**)	.463 (**)	.364 (**)	-.297 (**)	-.191 (**)	.397 (**)	-.564 (**)	-.251 (**)	-.179 (*)	-.189 (*)	-.070	.370 (**)	.475 (**)	.041	.099	-.059	.338 (**)	.339 (**)	.159 (**)	.428 (**)
	Sig.	.000	.000	.	.000	.000	.066	.000	.000	.000	.000	.000	.000	.000	.000	.028	.020	.349	.000	.000	.467	.105	.493	.000	.000	.005	.000
kepenys	Corr. Coeffic.	-.041	.153 (**)	.150 (**)	1.000	.201 (**)	.145 (**)	.186 (**)	.141 (**)	.306 (**)	.031	.100 (**)	.093 (**)	-.044	-.045	-.128	-.028	.182 (**)	.026	.140 (**)	-.025	.091	-.083	.116 (*)	-.035	-.041	.146 (**)
	Sig.	.210	.000	.000	.	.000	.000	.000	.000	.000	.349	.003	.006	.194	.371	.098	.722	.009	.449	.009	.644	.125	.320	.033	.527	.452	.000
bluznis	Corr. Coeffic.	.189 (**)	.260 (**)	.486 (**)	.201 (**)	1.000	-.008	.237 (**)	.280 (**)	.213 (**)	-.194 (**)	-.231 (**)	.177 (**)	-.468 (**)	-.271 (**)	-.377 (**)	-.205 (*)	-.043	.321 (**)	.347 (**)	.110	.020	.110	.187 (**)	.200 (**)	.067	.045
	Sig.	.000	.000	.000	.000	.	.812	.000	.000	.000	.000	.000	.000	.000	.000	.000	.020	.588	.000	.000	.051	.775	.250	.002	.001	.283	.254
alt	Corr. Coeffic.	-.083 (**)	.060	.062	.145 (**)	-.008	1.000	.790 (**)	.159 (**)	.422 (**)	.292 (**)	.012	-.168 (**)	-.054	-.026	-.087	-.002	.071	.078 (*)	.018	-.074	.024	.120	.092	.013	.035	.538 (**)
	Sig.	.009	.063	.066	.000	.812	.	.000	.000	.000	.000	.717	.000	.093	.602	.244	.980	.294	.022	.735	.156	.677	.138	.079	.801	.501	.000
ast	Corr. Coeffic.	.172 (**)	.129 (**)	.445 (**)	.186 (**)	.237 (**)	.790 (**)	1.000	.415 (**)	.546 (**)	.008	-.072 (*)	.146 (**)	-.284 (**)	-.111 (*)	-.173 (*)	-.093	.025	.334 (**)	.422 (**)	.037	.085	.061	.279 (**)	.229 (**)	.184 (**)	.567 (**)
	Sig.	.000	.000	.000	.000	.000	.000	.	.000	.000	.803	.024	.000	.000	.023	.019	.213	.717	.000	.000	.476	.141	.447	.000	.000	.000	.000
šf	Corr. Coeffic.	.286 (**)	.135 (**)	.463 (**)	.141 (**)	.280 (**)	.159 (**)	.415 (**)	1.000	.481 (**)	-.156 (**)	.009	.291 (**)	-.271 (**)	-.057	-.227 (**)	-.137	.141 (*)	.325 (**)	.602 (**)	.041	.151 (**)	.004	.276 (**)	.218 (**)	.095	.256 (**)
	Sig.	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.779	.000	.000	.253	.002	.067	.038	.000	.000	.440	.009	.964	.000	.000	.071	.000
ggt	Corr. Coeffic.	.137 (**)	.153 (**)	.364 (**)	.306 (**)	.213 (**)	.422 (**)	.546 (**)	.481 (**)	1.000	.097 (**)	.082 (*)	.120 (**)	-.179 (**)	.054	-.136	.122	.249 (**)	.303 (**)	.411 (**)	-.051	.105	.055	.341 (**)	.110 (*)	.150 (**)	.362 (**)
	Sig.	.000	.000	.000	.000	.000	.000	.000	.000	.	.003	.012	.000	.000	.276	.068	.103	.000	.000	.000	.339	.070	.501	.000	.038	.004	.000
hb	Corr. Coeffic.	-.253 (**)	.090 (**)	-.297 (**)	.031	-.194 (**)	.292 (**)	.008	-.156 (**)	.097 (**)	1.000	.259 (**)	-.615 (**)	.196 (**)	.127 (*)	.122	.011	.105	-.040	-.360 (**)	-.035	-.090	.004	-.045	-.152 (**)	-.143 (**)	.078 (*)
	Sig.	.000	.006	.000	.349	.000	.000	.803	.000	.003	.	.000	.000	.000	.010	.100	.881	.123	.238	.000	.501	.120	.958	.400	.004	.007	.026
leu	Corr. Coeffic.	-.207 (**)	-.097 (**)	-.191 (**)	.100 (**)	-.231 (**)	.012	-.072 (*)	.009	.082 (*)	.259 (**)	1.000	-.119 (**)	.430 (**)	.141 (**)	.015	.100	.157 (*)	-.079 (*)	-.023	.075	-.009	.027	.088	-.024	-.010	.008
	Sig.	.000	.003	.000	.003	.000	.717	.024	.779	.012	.000	.	.000	.000	.004	.845	.182	.020	.020	.667	.156	.870	.737	.095	.656	.855	.815
eng	Corr. Coeffic.	.412 (**)	.139 (**)	.397 (**)	.093 (**)	.177 (**)	-.168 (**)	.146 (**)	.291 (**)	.120 (**)	-.615 (**)	-.119 (**)	1.000	-.226 (**)	-.016	-.112	.030	.033	.098 (**)	.347 (**)	-.001	.247 (**)	.014	.294 (**)	.414 (**)	.254 (**)	.055
	Sig.	.000	.000	.000	.006	.000	.000	.000	.000	.000	.000	.000	.	.000	.755	.136	.689	.635	.004	.000	.986	.000	.860	.000	.000	.000	.120
tr	Corr. Coeffic.	-.380 (**)	-.196 (**)	-.564 (**)	-.044	-.468 (**)	-.054	-.284 (**)	-.271 (**)	-.179 (**)	.196 (**)	.430 (**)	-.226 (**)	1.000	.328 (**)	.246 (**)	.239 (**)	.154 (*)	-.327 (**)	-.263 (**)	-.034	-.115 (*)	-.024	-.192 (**)	-.291 (**)	-.132 (*)	-.130 (**)
	Sig.	.000	.000	.000	.194	.000	.093	.000	.000	.000	.000	.000	.000	.	.000	.001	.001	.024	.000	.000	.519	.049	.769	.000	.000	.014	.000
B_choles	Corr. Coeffic.	-.073	.017	-.251 (**)	-.045	-.271 (**)	-.026	-.111 (*)	-.057	.054	.127 (*)	.141 (**)	-.016	.328 (**)	1.000	.410 (**)	.864 (**)	.382 (**)	-.165 (**)	-.182 (*)	-.016	-.028	.067	-.073	-.178 (**)	-.154 (*)	-.071
	Sig.	.139	.736	.000	.371	.000	.602	.023	.253	.276	.010	.004	.755	.000	.	.000	.000	.000	.001	.022	.845	.695	.495	.264	.006	.018	.187
Dtl_chol	Corr. Coeffic.	.077	-.204 (**)	-.179 (*)	-.128	-.377 (**)	-.087	-.173 (*)	-.227 (**)	-.136	.122	.015	-.112	.246 (**)	.410 (**)	1.000	.072	-.209 (**)	-.188 (*)	-.334 (**)	-.064	-.257 (**)	.013	-.052	.064	-.006	-.072
	Sig.	.298	.007	.028	.098	.000	.244	.019	.002	.068	.100	.845	.136	.001	.000	.	.331	.005	.014	.004	.593	.007	.924	.577	.491	.950	.375
Mtl_chol	Corr. Coeffic.	-.132	-.070	-.189 (*)	-.028	-.205 (*)	-.002	-.093	-.137	.122	.011	.100	.030	.239 (**)	.864 (**)	.072	1.000	.356 (**)	-.159 (*)	-.099	.019	.086	.179	-.278 (**)	-.146	-.121	-.087
	Sig.	.075	.358	.020	.722	.020	.980	.213	.067	.103	.881	.182	.689	.001	.000	.331	.	.000	.039	.408	.876	.378	.187	.003	.118	.194	.288

		amziu s	kmi	fibroz e	kepen ys	bluzni s	alt	ast	šf	ggt	hb	leu	eng	tr	B_chole s	Dtl_chol	Mtl_chol	triglic	bilirub	T_bilir	N_bilir	CIK	NBT	iga	igg	igm	Hai_r
triglic	Corr. Coeffic.	.044	.159 (*)	-.070	.182 (**)	-.043	.071	.025	.141 (*)	.249 (**)	.105	.157 (*)	.033	.154 (*)	.382 (**)	-.209 (**)	.356 (**)	1.000	-.105	.071	-.220 (*)	.085	-.143	-.153	-.239 (**)	-.168 (*)	.000
	Sig.	.513	.020	.349	.009	.588	.294	.717	.038	.000	.123	.020	.635	.024	.000	.005	.000	.	.133	.508	.038	.357	.281	.073	.004	.049	1.000
B_bilirub	Corr. Coeffic.	.183 (**)	.106 (**)	.370 (**)	.026	.321 (**)	.078 (*)	.334 (**)	.325 (**)	.303 (**)	-.040	-.079 (*)	.098 (**)	-.327 (**)	-.165 (**)	-.188 (*)	-.159 (*)	-.105	1.000	.727 (**)	.729 (**)	.036	-.063	.177 (**)	.233 (**)	.101	.170 (**)
	Sig.	.000	.002	.000	.449	.000	.022	.000	.000	.000	.238	.020	.004	.000	.001	.014	.039	.133	.	.000	.000	.542	.452	.001	.000	.065	.000
T_bilir	Corr. Coeffic.	.211 (**)	.003	.475 (**)	.140 (**)	.347 (**)	.018	.422 (**)	.602 (**)	.411 (**)	-.360 (**)	-.023	.347 (**)	-.263 (**)	-.182 (*)	-.334 (**)	-.099	.071	.727 (**)	1.000	.191 (**)	-.077	.118	.300 (**)	.236 (**)	.080	.315 (**)
	Sig.	.000	.953	.000	.009	.000	.735	.000	.000	.000	.000	.667	.000	.000	.022	.004	.408	.508	.000	.	.000	.431	.404	.000	.006	.355	.000
N_bilir	Corr. Coeffic.	-.043	.034	.041	-.025	.110	-.074	.037	.041	-.051	-.035	.075	-.001	-.034	-.016	-.064	.019	-.220 (*)	.729 (**)	.191 (**)	1.000	-.134	-.101	-.071	.053	-.101	-.063
	Sig.	.412	.528	.467	.644	.051	.156	.476	.440	.339	.501	.156	.986	.519	.845	.593	.876	.038	.000	.000	.	.171	.476	.413	.542	.243	.311
CIK	Corr. Coeffic.	.036	.126 (*)	.099	.091	.020	.024	.085	.151 (**)	.105	-.090	-.009	.247 (**)	-.115 (*)	-.028	-.257 (**)	.086	.085	.036	-.077	-.134	1.000	.067	.119	.262 (**)	.360 (**)	.117
	Sig.	.526	.033	.105	.125	.775	.677	.141	.009	.070	.120	.870	.000	.049	.695	.007	.378	.357	.542	.431	.171	.	.432	.060	.000	.000	.061
NBT	Corr. Coeffic.	-.079	-.025	-.059	-.083	.110	.120	.061	.004	.055	.004	.027	.014	-.024	.067	.013	.179	-.143	-.063	.118	-.101	.067	1.000	.062	.035	-.070	.108
	Sig.	.327	.766	.493	.320	.250	.138	.447	.964	.501	.958	.737	.860	.769	.495	.924	.187	.281	.452	.404	.476	.432	.	.485	.697	.434	.204
iga	Corr. Coeffic.	.236 (**)	.145 (**)	.338 (**)	.116 (*)	.187 (**)	.092	.279 (**)	.276 (**)	.341 (**)	-.045	.088	.294 (**)	-.192 (**)	-.073	-.052	-.278 (**)	-.153	.177 (**)	.300 (**)	-.071	.119	.062	1.000	.416 (**)	.202 (**)	.186 (**)
	Sig.	.000	.007	.000	.033	.002	.079	.000	.000	.000	.400	.095	.000	.000	.264	.577	.003	.073	.001	.000	.413	.060	.485	.	.000	.000	.001
igg	Corr. Coeffic.	.270 (**)	.176 (**)	.339 (**)	-.035	.200 (**)	.013	.229 (**)	.218 (**)	.110 (*)	-.152 (**)	-.024	.414 (**)	-.291 (**)	-.178 (**)	.064	-.146	-.239 (**)	.233 (**)	.236 (**)	.053	.262 (**)	.035	.416 (**)	1.000	.260 (**)	.119 (*)
	Sig.	.000	.001	.000	.527	.001	.801	.000	.000	.038	.004	.656	.000	.000	.006	.491	.118	.004	.000	.006	.542	.000	.697	.000	.	.000	.037
igm	Corr. Coeffic.	.086	-.076	.159 (**)	-.041	.067	.035	.184 (**)	.095	.150 (**)	-.143 (**)	-.010	.254 (**)	-.132 (*)	-.154 (*)	-.006	-.121	-.168 (*)	.101	.080	-.101	.360 (**)	-.070	.202 (**)	.260 (**)	1.000	.073
	Sig.	.101	.155	.005	.452	.283	.501	.000	.071	.004	.007	.855	.000	.014	.018	.950	.194	.049	.065	.355	.243	.000	.434	.000	.000	.	.204
Hai_r	Corr. Coeffic.	.122 (**)	.085 (*)	.428 (**)	.146 (**)	.045	.538 (**)	.567 (**)	.256 (**)	.362 (**)	.078 (*)	.008	.055	-.130 (**)	-.071	-.072	-.087	.000	.170 (**)	.315 (**)	-.063	.117	.108	.186 (**)	.119 (*)	.073	1.000
	Sig.	.000	.014	.000	.000	.254	.000	.000	.000	.000	.026	.815	.120	.000	.187	.375	.288	1.000	.000	.000	.311	.061	.204	.001	.037	.204	.

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).