

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
MATEMATINĖS STATISTIKOS KATEDRA

Roma Puronaitė

(parašas)

**T BANGOS KAITOS ANALIZĖ, NAUDOJANT MODIFIKUOTĄ
SLENKANČIO VIDURKIO METODĄ: ĮVAIRIŲ LAIKO EILUČIŲ
IŠLYGINIMO IR PANAŠUMO NUSTATYMO BŪDŲ Palyginimas**

Magistro baigiamasis darbas

Vilnius, 2012

Darbo vadovas:
Docentas, matematikos mokslų daktaras, Pranas Vaitkus (parašas)

Recenzentas:
Docentė, matematikos mokslų daktarė, Rūta Levulienė (parašas)

Registracijos Nr.:
Darbo gynimo data:

TURINYS

ĮVADAS.....	4
1. Širdies veikla ir elektrokardiograma.....	6
1.1. Širdies fiziologija.....	6
1.1.1. Širdies sandara ir kardiociklas.....	6
1.1.2. Širdies elektrinis laidumas.....	7
1.2. Elektrokardiograma.....	8
1.2.1. Istorija.....	8
1.2.2. Derivacijų sistemos.....	8
1.2.3. EKG interpretacija.....	9
1.2.4. T bangos kaita.....	10
2. Analizės metodai.....	10
2.1. Modifikuotas slenkančio vidurkio metodas.....	10
2.2. Laiko eilučių išlyginimo metodai.....	12
2.2.1. Dinaminis laiko skalės kraipymas.....	12
2.2.2. Atviros pradžios ir pabaigos dinaminis laiko skalės kraipymas.....	15
2.2.3. Išlyginimas pikų atžvilgiu.....	15
2.2.4. Išlyginimas naudojant kryžminę koreliaciją.....	16
2.3. Laiko eilučių panašumo nustatymo būdai.....	17
2.3.1. Čebyševio atstumas.....	17
2.3.2. Absoliutinis skirtumas tarp maksimumo taškų.....	17
2.3.3.-2.3.4. Kosinusų panašumas ir koreliacija.....	17
2.3.5. Tikėtimumo santykio atstumas.....	18
2.3.6. Lp normos apibendrinimas laiko eilutėms.....	20
2.4. Diskrečioji vilnelių transformacija (DBT).....	22
2.4.1. Diskrečioji vilnelių transformacija.....	22
2.4.2. Vilnelių panaudojimas EKG triukšmo ir izoelektrinės linijos šalinimui.....	23
2.4.3. Vilnelių panaudojimas T bangų išskyrimui.....	24
2.5. Statistinės analizės metodai.....	26
2.5.1. Išskirčių paieška.....	26
2.5.2. Neparametriniai kriterijai.....	27
2.5.3. Logistinė regresija.....	30
2.5.4. Kelių biomarkerių apjungimas min-max kombinacija.....	31
3. Duomenų analizė.....	33
3.1. Duomenys.....	33
3.1.1. TWA duomenų bazė.....	33
3.1.2. PTB duomenų bazė.....	34
3.1.3. Generuoti duomenys.....	34
3.2. Duomenų paruošimas analizei.....	36
3.2.1. Triukšmo ir izoelektrinės linijos dreifo šalinimas.....	36
3.2.2. T bangų išskyrimas.....	38
3.3. TWA duomenų bazės ir generuotų duomenų analizė.....	38
3.4. PTB duomenų analizė.....	40
3.4.1. Pradinė analizė.....	40
3.4.2. Homogeniškumo tikrinimas.....	45

3.4.3. Logistinė regresija.....	50
3.4.4. Kelių biomarkerių apjungimas.....	51
Išvados ir rekomendacijos.....	54
Santrauka.....	56
Summary.....	57
Literatūros sąrašas.....	58
Priedai.....	60
1. TBK analizei naudoti R paketai.....	60
2. Sukurtų funkcijų komentarai	61
2. Laiko eilučių išlyginimo pavyzdžiai.....	64
4. Ritmo pavyzdžiai	71
5. Kendalo koreliacijos generuotiems duomenims ir TWA duomenų bazės duomenims.....	72

ĮVADAS

T bangos kaita (TBK) – vienas iš galimų kardiologinių biomarkerių nagrinėjant įvairias aritmijas ir sutrikimus susijusius su širdies veiklos pasikeitimais. Didelis TBK laikomas pavojingų gyvybei skilvelinių aritmijų rizikos veiksniu ir papildoma priežastimi taikyti implantuojamų kardioverterių defibriliatorių (IKD) gydymą.[15]

T bangos aptikimas yra gana komplikauta TBK analizės dalis. Algoritmai, naudojami šiam tikslui neužtikrina šimtaprocentinio tikslumo. Dėl triukšmo įtakos ir elektrokardiogramos nepastovumo atsiradę netikslumai lemia T bangų pradžios ir pabaigos taškų pasislinkimą. Tai tampa ypač aktualu lyginant T bangas vieną su kita. Kitą vertus, T bangų trukmė ir forma priklauso nuo širdies ritmo pasikeitimo, t. y. RR intervalo trukmės pasikeitimo. T bangų lyginimas kiekviename taške tokiu atveju duos klaidinančius rezultatus dėl morfologinių pasikeitimų. Stengiantis išvengti tokių paklaidų, taikomi įvairūs apribojimai EKG įrašymo momentu, pvz. pastovus kvėpavimas, tačiau toks analizės būdas apriboja galimybes tirti pacientus prieš įvykstant aritmijoms, kai idealių sąlygų nėra įmanoma sukurti ar įprastomis gyvenimo sąlygomis, kas yra aktualu tuo atveju, jei TBK pasikeitimą būtų bandoma pritaikyti kaip staigios pavojingos aritmijos indikatorių IKD aparatuose, stimuliuojančiuose širdį siekiant sušvelninti aritmijų poveikį ar net laiku jų išvengti. [21] Vadinasi svarbiausi uždaviniai yra: kuo tikslesnis bangų išlyginimas ir kuo kokybiškesnis jų palyginimas. Šiame darbe siekiama palyginti keletą galimų T bangų išlyginimo metodų ir panašumo nustatymo būdų ieškant optimalaus sprendimo.

Palyginimo momentu bangos gali būti pasislinkę viena kitos atžvilgiu dėl kelių priežasčių:

- RR intervalų trukmės kaitos (t. y. T bangos ilgio pasikeitimo dėl širdies ritmo);
- T bangos pradžios ir pabaigos taškų aptikimo paklaidos;
- suvidurkintos ir naujos bangos skirtumo dėl suvidurkintosios ir naujos T bangos trukmės nesutapimo.

Modifikuoto slenkančio vidurkio metode TBK įvertinamas palyginant dvi suvidurkintas bangas, kurių viena yra suvidurkinta lyginė, o kita – nelyginė. Tradiciniame metode palyginimas atliekamas surandant didžiausią skirtumą tarp atitinkamų bangų taškų, tačiau aktualu gali būti ne

tik didžiausias skirtumas, bet ir bangų panašumas apskritai, t. y. T bangos gali smarkiai skirtis tik viename taške, o gali ir visoje T bangoje.

Alternatyvūs panašumo nustatymo būdai galėtų būti tokie:

- absoliutinis skirtumas tarp maksimumo taškų;
- kosinusų panašumas;
- koreliacija;
- tikėtimumo santykio atstumas;
- apibendrinta L_p norma laiko eilutėms (Lee – Verleysen atstumas).

Laiko eilučių išlyginimo metodai:

- dinaminis laiko skalės kraipymas (ang. DTW – dynamical time warping);
- dinaminis laiko skalės kraipymas, esant neapribotiems pradžios ir pabaigos taškams su simetriniu judėjimo šablonu;
- dinaminis laiko skalės kraipymas, esant neapribotiems pradžios ir pabaigos taškams su asimetriniu judėjimo šablonu;
- išlyginimas naudojant kryžminę koreliaciją;
- išlyginimas pagal maksimumo taškus (pikus).

1. Širdies veikla ir elektrokardiograma

1.1. Širdies fiziologija

1.1.1. Širdies sandara ir kardiociklas

Širdis – organas atsakingas už kraujo paskirstymą po visą organizmą. [18] Ją sudaro 4 pagrindinės dalys, kairysis ir dešinysis prieširdžiai bei kairysis ir dešinysis skilveliai. Kairioji širdies pusė kontroliuoja kraujo tekėjimą po visą kūną, o dešinioji – kraujo tekėjimą mažuoju kraujo apytakos ratu, t. y. į plaučius ir iš jų.

Kairysis prieširdis gauna deguonies prisotintą kraują iš plaučių, tuomet kraujas teką į kairįjį skilvelį, kuris išstumia kraują į aortą ir arterinę sistemą, kuri išnešioja deguonį po visą organizmą. Dešinysis prieširdis priima grįžtantį kraują ir šis teka į dešinįjį skilvelį, kuris išstumia kraują į plaučius, kur kraujas vėl pasiima deguonį.

Didžioji dalis raumenų masės sukaupta kairiojoje širdies pusėje, nes ši dalis turi aptarnauti didesnę organizmo dalį, t. y. reikalinga didesnė išstūmimo jėga. Širdies sienelės turi specialias raumenų ląsteles, kurios sukuria elektrinį sudirginimą ir kontroliuoja vožtuvus bei prieširdžių ir skilvelių susitraukimus. Širdies raumenyną sudaro tokie sluoksniai: endokardas – plonas sluoksnis, dengiantis vidines prieširdžių ir skilvelių sienelės, miokardas – storiasis sluoksnis, sudaro prieširdžių bei skilvelių sienelės ir epikardas – suaugęs su miokardu ir jį dengiantis sluoksnis.

Miokardas atsako į elektrinius signalus susitraukimais. Jo pažeidimas yra itin pavojingas, nes gali sutrikdyti širdies susitraukimų dinamiką, dalinai ar visiškai nutraukti kraujo apytaką ir deguonies pernešimą organizme.

Skilvelius nuo prieširdžių skiria vožtuvai, neleidžiantys kraujui tekėti atgaline kryptimi, vožtuvai skiria ir skilvelius nuo arterijų.

Prieširdžius ir skilvelius jungia Hiso kojųčių pluoštas, kuris svarbus elektrinio signalo perdavimui tarp šių širdies dalių. Šio pluošto sutrikimai lemia neefektyvų, netolygų, vėluojantį ar ankstyvą susitraukimą.

Visas kardiociklas apima širdies ląstelių repoliarizaciją ir depoliarizaciją, t. y. širdies raumenų susitraukimą ir atsipalaidavimą. Laikas, kai raumenys susitraukia yra vadinamas sistole, kai atsipalaiduoja – diastole.

1.1.2. Širdies elektrinis laidumas

Širdies raumens susitraukimai sąlygojami elektrinio impulso. Miokardo ląstelės, kurios sudaro didžiausią širdies dalį, ypatingos tuo, kad turi ilgesnį depoliarizacijos periodą ir leidžia lėtesnį impulso sklidimą ląstelėmis. Ritmo vedlio peismeikerinės (stimuliuojančios), trumpiau vadinamos P ląstelėmis, nesusitraukia, tačiau kaupia potencialą ir sukuria impulsą, kai jis peržengia slenkstį. Impulsas perduodamas artimiausioms miokardo ląstelėms. Teigiami jonai prasiskverbia pro ląstelių membraną. Miokardo ląstelių depoliarizacija priverčia jas sumažėti ir susitraukti. Sujaudinimas plinta iš vieno ląstelės į kitas ląsteles. Kai teigiami jonai išstumiami iš miokardo ląstelių, jos atsipalaiduoja ir atgauna savo ankstesnį dydį.

Širdies ląstelės turi keletą cheminių kanalų, kurie lemia susitraukimo procesą ir dviejų tipų potencialus. Didžioji ląstelių masė sukaupia prieširdžiuose, skilveliuose ir jungiamuosiuose pluoštuose yra greito atsako į impulsą ląstelės. Jos turi lėtus kalcio ir greitus natrio kanalus. Greiti kanalai yra tokie patys ir širdies ląstelėms, ir skeleto raumenims, o lėto atsako kalcio kanalai prailgina depoliarizacijos laiką miokardo ląstelėms ir užtikrina širdies ląstelių susitraukimus.

P ląstelės, sinusiniame, atrioventrikuliniame mazguose ir Hiso kojų pluošte neturi greitų natrio kanalų.

Kiekvieną bioelektrinį potencialą sudaro 4 fazės [9].

0 fazė: pradinė būklė, teigiamos krypties greita depoliarizacija dėl natrio jonų.

1 fazė: nežymi repoliarizacija dėl kalio jonų netekimo;

2 fazė: stabilizavimasis kalio jonams kompensuojant kalio jonų praradimą;

3 fazė: atpalaidavimo pradžia, kai kalio tekėjimas liaujasi, o natriis pasišalina iš ląstelių;

4 fazė: poilsio periodas, esantis tarp repoliarizacijos pabaigos ir depoliarizacijos pradžios, kai kalis vėl pradeda skverbtis į ląsteles.

Lėto atsako potencialai vyksta sinusiniame ir atrioventrikuliniame mazguose. P ląstelės susižadina pačios dėl pastovaus jonų tekėjimo, t. y. poilsio fazės nebuvimo. Vis dėlto pastebimas smarkus potencialo didėjimas 4 fazėje, vadinamasis peismeikerinis (stimuliuojantis) potencialas.

Ilga depoliarizacija 0 fazėje vyksta dėl lėto kalcio jonų tekėjimo, kuris prasideda tada, kai stimuliuojantis potencialas peržengia slenkstį (~ -40 mV). 3 repoliarizacijos fazė prasideda, kai kalio jonai pasišalina iš ląstelės. Repoliarizacija baigiasi, kai potencialas pasiekia ~ -60 mV.

1.2. Elektrokardiograma

1.2.1. Istorija

Elektrokardiograma (EKG) – tai elektrinio širdies aktyvumo, matuojamo specialiu prietaisu, elektrokardiografu, užrašimas laike. [19]

Pirmoji žmogaus elektrokardiograma buvo užrašyta Desire Waller 1887 metais. Vėliau Eithoven`as patobulino EKG užrašymo mechanizmą, pasiūlė galūnių derivacijas (atvadás).

Šiuo metu EKG gali būti atliekama arba įprastu elektrokardiografu, atspausdinat ant specialaus popieriaus, arba įrašant į kompiuterį. Tokiu atveju galima išsamesnė EKG analizė.

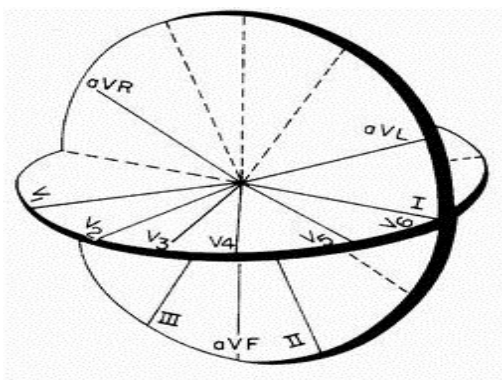
Šiuolaikiniais aparatais galimas imties dažnis yra nuo 250 Hz iki 1000 Hz.

1.2.2. Derivacijų sistemos

Elektriniai impulsai veikiantys širdyje gali būti fiksuojami nuo odos paviršiaus ir užrašomi derivacijomis (atvadais), elektrodams kontaktuojant su oda per lipnius kontaktus.

Plačiausiai naudojama 12 derivacijų sistema, bet įmanomos ir kitos kombinacijos.

12 derivacijų sistema. Ši sistema fiksuoja širdies elektrinį aktyvumą trijose dimensijose. Galūnių derivacijos registruoja širdies bioelektrinių potencialų projekciją į kūno frontaliąją plokštumą, krūtininės derivacijos V1 – V6 – horizontaliąją plokštumą bei fiksuoja skilvelių elektrinį aktyvumą, aVF ir V2 – statmeną frontaliajai ir horizontaliajai plokštumoms plokštumą.



1 pav.

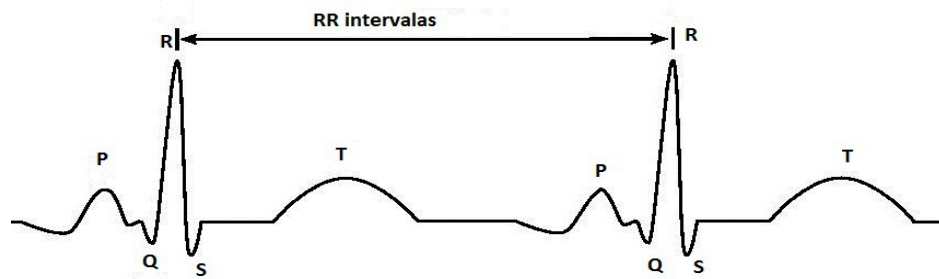
Visa sistema apima 6 galūnių ir 6 krūtinines derivacijas. Galima kiekvieną derivaciją įrašyti atskirai, tačiau kai kurias derivacijas galima išskaičiuoti naudojant formules.

$$\begin{aligned}
III &= I - II \\
aVR &= -\frac{1}{2}(I + II) \\
aVL &= I - \frac{1}{2}II \\
aVF &= II - \frac{1}{2}I \\
aVR + aVL + aVF &= 0
\end{aligned}
\tag{1}$$

Holter`io EKG. Holter`io EKG įrašoma naudojant mobilų aparatą, dviejuose vietose kontaktuojant su paciento kūnu. Viena derivacija gaunama prijungus elektrodą prie krūtinės, o kita – tvirtinant elektrodą paciento nugaroje. Naujaisi aparatai suteikia galimybę gauti derivacijas atitinkančias V5 ir aVF derivacijas. Holter`io EKG dažniausiai naudojama esant reikalui stebėti pacientą visą parą, kuo mažiau keičiant jo gyvenimo ritmą, tikintis užfiksuoti aritmijas, kurios yra trumpalaikės, gali išstikti nenustatytu laiku. Tokiu būdu įrašyta EKG gali suteikti papildomų žinių atsižvelgiant į jos kaitą, pvz. dienos ir nakties metu, pacientui atliekant tam tikrus veiksmus.

1.2.3. EKG interpretacija

Kiekviena EKG turi bendrus, joms būdingas bangas ir pikus, kurie naudojami kardiologų nustatant ligas ar sutrikimus, sprendžiant apie širdies darbo kokybę.



2 pav. EKG būdingos bangos

EKG laiko eilutėje atsispindi kiekvienas kardiociklas. Priimta vieno kardiociklo trukmę matuoti nuo vieno R piko iki kito R piko, nes šie taškai labiausiai pastebimi, tačiau fiziologiškai ciklas prasideda ties viena P banga ir baigiasi ties kita. P banga rodo pirminį sinusinio mazgo impulsą ir dešiniojo bei kairiojo skilvelių susitraukimą, atrioventrikulinio mazgo sužadimą. QRS kompleksas žymi skilvelių depolarizaciją, o T banga – skilvelių repolarizaciją.

T bangos paprastai yra asimetriškos, dešinioji pusė – statesnė, o kairioji nuožulnesnė. Simetriškumas ir statusas gali rodyti ligą. Pastebėta, kad vyresniems žmonėms ir moterims T bangos amplitudė būna mažesnė.

1.2.4. T bangos kaita

T bangos kaita (TBK) – tai T bangos amplitudės ar formos pasikeitimas kiekviename sekančiame širdies cikle. Šie pasikeitimai siejami su skilvelių repoliarizacijos sutrikimais. Tam tikrais atvejais TBK gali būti matomi ir plika akimi, pvz. esant dideliame amplitudės pasikeitimui. Tačiau dažniausiai T bangos pasikeitimai nėra lengvai nustatomi. Atsiradus galimybei EKG rezultatus gauti skaitmeniniu formatu bei atlikti matavimus μV tikslumu, galima išsamesnė TBK analizė.

TBK nustatymui ir įvertinimui medicinos praktikoje naudojami įvairūs metodai, vienas iš jų yra Nearing ir Verrier pasiūlytas modifikuotas slenkančio vidurkio metodas.

2. Analizės metodai

2.1. Modifikuotas slenkančio vidurkio metodas

Nearing ir Verrier [14] 2002 metais pasiūlytas TBK analizės metodas, kurio pagrindinis privalumas yra galimybė fiksuoti T bangos kaitą laike. Metodo esmė yra lyginių ir nelyginių bangų rekursinis vidurkinimas įvedant netiesinius apribojimus, skirtus artefaktų (neaiškios kilmės išsiskiriančių reikšmių) įtakai sumažinti.

Pasirinktam T bangų skaičiui N , T bangos paskirstomos į du blokus, A – nelygines T bangas ir B – lygines T bangas. Kiekviename bloke atskirai rekursiškai atnaujinami vidurkiai $\bar{T}_{A,k,i}$ ir $\bar{T}_{B,k,i}$. Apribojimai (3) sumažina galimą kaitą kiekviename bloke. Didžiausiais leidžiamas pasikeitimas tarp vidurkio bangos ir sekančios T bangos atitinkamų taškų yra $256 \mu V$, o vidurkis bus atnaujinamas daugiausiai $32 \mu V$.

$$\begin{aligned}
 &T_{A,k,i} - A \text{ bloko } k\text{-tosios } T \text{ bangos amplitudė } (\mu V), \\
 &k = 1, \dots, \frac{N}{2}, \quad i = 1, \dots, N_{A,k}, \\
 &N - T \text{ bangų skaičius,} \\
 &N_{A,k} - A \text{ bloko } k\text{-tosios } T \text{ bangos elementų skaičius}
 \end{aligned} \tag{2}$$

$$\bar{T}_{A,1,i} = T_{A,1,i}$$

$$\eta_{A,k,i} = \frac{1}{8}(T_{A,k,i} - \bar{T}_{A,k-1,i})$$

$$\begin{aligned} \Delta_{A,k,i} &= -32, & \text{jei } \eta_{A,k,i} &\leq -32 \\ \Delta_{A,k,i} &= \eta_{A,k,i}, & \text{jei } -32 < \eta_{A,k,i} &\leq -1 \\ \Delta_{A,k,i} &= -1, & \text{jei } -1 < \eta_{A,k,i} &< 0 \\ \Delta_{A,k,i} &= 0, & \text{jei } \eta_{A,k,i} &= 0 \\ \Delta_{A,k,i} &= 1, & \text{jei } 0 < \eta_{A,k,i} &\leq 1 \\ \Delta_{A,k,i} &= \eta_{A,k,i}, & \text{jei } 1 < \eta_{A,k,i} &\leq 32 \\ \Delta_{A,k,i} &= 32, & \text{jei } \eta_{A,k,i} &> 32 \end{aligned} \quad (3)$$

$$\bar{T}_{A,k,i} = \bar{T}_{A,k-1,i} + \Delta_{A,k,i}$$

$$2 \leq k \leq \frac{N}{2}$$

$$\Lambda_k = \max_i |\bar{T}_{A,k,i} - \bar{T}_{B,k,i}| \quad (4)$$

Analogiški veiksmai atliekami ir B blokui.

Algoritme yra du momentai kada atliekamas palyginimas: tuo metu, kada skaičiuojami skirtumai $T_{A,k,i} - \bar{T}_{A,k-1,i}$ bei $\bar{T}_{A,k,i} - \bar{T}_{B,k,i}$, tad bangų ilgiai palyginimo momentu turėtų sutapti, tačiau realiai bangos nebūtinai yra vienodo ilgio, todėl prieš atliekant palyginimą jos turi būti išlyginamos viena kitos atžvilgiu.

Tarp apskaičiuotųjų vidurkių panašumui vertinti Nearing ir Verrier algoritme naudojamas Λ_k , tačiau galbūt pakeitus jį kitu, galima gauti geresnius rezultatus.

2.2. Laiko eilučių išlyginimo metodai

2.2.1. Dinaminis laiko skalės kraipymas

Dinaminis laiko skalės kraipymas (dynamic time warping) – metodas skirtas laiko eilučių išlyginimui. Šis metodas taikomas kalbos atpažinimo, klasifikavimo uždaviniams spęsti.

B.Huang, W.Kinsner [16] naudoja DLSK metodą EKG klasifikavimui, o Cuesta-Frau ir kt. [1][2] naudoja TBK analizėje T bangų išlyginimui.

DLSK apibėžimas pagal Rabiner ir Juang. [13]

Tarkime turime dvi diskretaus laiko eilutes X ir Y, kur

$$\begin{aligned} X &= (x_1, x_2, \dots, x_{t_x}), \\ Y &= (y_1, y_2, \dots, y_{t_y}), \\ t_x, t_y &- X \text{ ir } Y \text{ trukmės.} \end{aligned} \quad (5)$$

Tuomet apibrėžiame X ir Y laiko indeksus atitinkamai, $i=1,2, \dots, t_x$ ir $j=1,2, \dots, t_y$.

Sukonstruojama dydžio $t_x \times t_y$ matrica D.

$$\begin{aligned} D &= (D_{i,j})_{t_x \times t_y}, \\ D_{i,j} &= d(x_i, y_j) = (x_i - y_j)^2 \\ i &= 1, 2, \dots, t_x, \quad j = 1, 2, \dots, t_y. \end{aligned} \quad (6)$$

Dėl paprastumo toliau žymima $D_{i,j} = d(x_i, y_j) = d(i, j)$.

Tikslas yra pakeisti laiko skalę taip, kad naujosios laiko eilutės X' ir Y' būtų kuo panašesnės.

Tam ieškoma tokia funkcijų pora (7), kuri atitiktų optimalų išlyginimą $D(X, Y)$ (9).

$$\begin{aligned} \phi &= (\phi_x, \phi_y), \text{ kur} \\ \phi_x(k) &\in \{1, \dots, t_x\}, \\ \phi_y(k) &\in \{1, \dots, t_y\}, \\ k &= 1, \dots, t. \end{aligned} \quad (7)$$

$$d_{\phi}(X, Y) = \sum_{k=1}^t d(\phi_x(k), \phi_y(k)) m_{\phi}(k) / M_{\phi}, \quad (8)$$

m_{ϕ} – žingsnio svoris, M_{ϕ} – normalizuojanti konstanta

$$D(X, Y) = \min_{\phi} d_{\phi}(X, Y) \quad (9)$$

Dalinis sukaupytasis minimumas

$$D(i, j) = \min_{\phi, t'} \sum_{k=1}^{t'} d(\phi_x(k), \phi_y(k)) m_{\phi}(k) / M_{\phi}, \quad (10)$$

$\phi_x(t') = i, \phi_y(t') = j,$
 m_{ϕ} – žingsnio svoris, M_{ϕ} – normalizuojanti konstanta

Siekiant išlaikyti laiko eiliškumą įvedami apribojimai, t. y. uždraudžiami žingsniai atgal.

$$\begin{aligned} \phi_x(k+1) &\geq \phi_x(k), \\ \phi_y(k+1) &\geq \phi_y(k) \end{aligned} \quad (11)$$

Galimi papildomi apribojimai.

Simetriškumo apribojimas, kuris neleidžia peršokti laiko skalės daugiau nei per 1 laiko vienetą ir nepalieka nesulygintų taškų.

$$\begin{aligned} |\phi_x(k+1) - \phi_x(k)| &\leq 1, \\ |\phi_y(k+1) - \phi_y(k)| &\leq 1. \end{aligned} \quad (12)$$

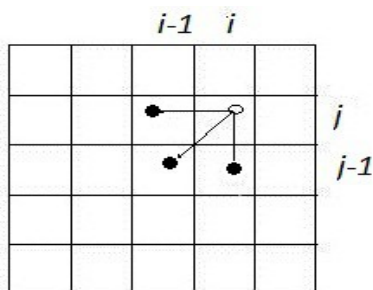
Globalaus išlyginimo apribojimas, kuris nusako, kad abiejų laiko eilučių pradžios ir pabaigos taškai turi sutapti.

$$\begin{aligned} \phi_x(1) &= \phi_y(1) = 1; \\ \phi_x(t) &= t_x; \\ \phi_y(t) &= t_y. \end{aligned} \quad (13)$$

Judėjimo šablonai – tai leidžiami kraipymo funkcijos žingsniai, reguliuojami svoriais. Tai yra galimos vertės $\phi(k+1)$, kai turima informacija apie $\phi(k), \phi(k-1)$. Judėjimo šablonai gali būti laikomi lokaliais apribojimais, kurie nurodo ar gali dubliuotis elementai, kiek taškų galima praleisti nesulyginus.

Lokalūs apribojimai įvedami apribojant $m_{\phi}(k)$:

Simetrinis judėjimo šablonas neleidžia peršokti nesulygintų taškų.



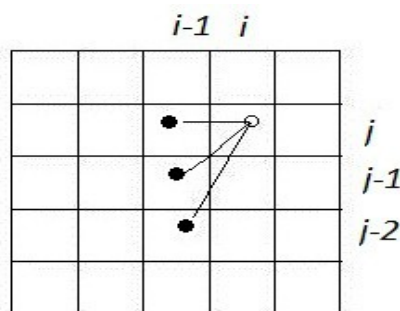
3 pav. simetrinio judėjimo šablono schema

Dalinis sukauptasis minimumas i ir j taškuose simetrinio judėjimo šablono atveju išreiškiamas formule (14).

$$D(i, j) = \min \begin{cases} D(i-1, j) + m_{\phi(k)} d(i, j) \\ D(i, j-1) + m_{\phi(k)} d(i, j) \\ D(i-1, j-1) + m_{\phi(k)} d(i, j) \end{cases} \quad (14)$$

$$m_{\phi(k)} = \begin{cases} 1, & \text{kai } (i-1, j) \text{ arba } (i, j-1) \\ 2, & \text{kai } (i-1, j-1) \end{cases} \quad (15)$$

Asimetrinis judėjimo šablonas leidžia „peršokti“ taškus.



4 pav. asimetrinio judėjimo šablono schema

Dalinis sukauptasis minimumas i ir j taškuose asimetrinio judėjimo šablono atveju išreiškiamas formule (16).

$$D(i, j) = \min \begin{cases} D(i-1, j-2) + m_{\phi(k)} d(i, j) \\ D(i-1, j-1) + m_{\phi(k)} d(i, j) \\ D(i-1, j) + m_{\phi(k)} d(i, j) \end{cases} \quad (16)$$

$$m_{\phi}(k)=1, \text{ kai } (i-1, j-2) \text{ arba } (i-1, j-1) \text{ arba } (i-1, j) . \quad (17)$$

2.2.2. Atviros pradžios ir pabaigos dinaminis laiko skalės kraipymas

Neapribotos arba atviros pradžios ir pabaigos (OBE - open-begin-end) DLSK metodas [11] sulygina tam tikrą, neturinčią trūkio, Y dalį su X.

$$Y^{(p,q)} = y_p, \dots, y_q, \\ D_{OBE}(X, Y) = \min_{1 \leq p \leq q \leq t_y} D(X, Y^{(p,q)}) \quad (18)$$

$$\phi_y(1) = p, \\ \phi_y(t) = q \quad (19)$$

Toks metodas veiksmingas tuo atveju, jei nėra tiksliai žinoma vieta, nuo kurios pradėdant ar kuria baigiant turėtų būti atliekamas išlyginimas. Tai aktualu TBK analizėje, nes T bangų pradžios ir pabaigos momentai nėra lengvai nustatomi ir priklausomai nuo naudojamo T bangų išskyrimo metodo, šie taškai gali būti pasislinkę.

Tuo atveju, kai atviros pradžios ir pabaigos DLSK metodas naudojamas kartu su asimetriniu judėjimo šablonu leidžia palikti nesulygintų taškų, tokiu būdu gali būti išgaunama panašesnė T bangų forma, nei naudojant simetrinį judėjimo šabloną.

2.2.3. Išlyginimas pikų atžvilgiu

Toks išlyginimo būdas tikslingas tuo atveju, jei dvi laiko eilutės turi aiškius pikus. T bangos turi vieną piką, kuris lengviau aptinkamas nei T bangos pradžios ir pabaigos taškai.

Tarkime turime dvi diskretaus laiko eilutes X ir Y, kurios apibrėžtos (5) formule.

Tuomet apibrėžiame X ir Y laiko indeksus atitinkamai, $i=1,2, \dots, t_x$ ir $j=1,2, \dots, t_y$.

Ieškomi maksimumo taškai kiekvienai laiko eilutei.

$$\begin{aligned} i_{max} &= \operatorname{argmax}_i x_i \\ j_{max} &= \operatorname{argmax}_j y_j \end{aligned} \quad (20)$$

Išlyginimas atliekamas tariant, kad išlygintų laiko eilučių naujieji indeksai sutampa i_{max} ir j_{max} atžvilgiu.

$$\begin{aligned} i_1^{(2)} &= i_{max} - \min(i_{max}, j_{max}) + 1, \\ j_1^{(2)} &= j_{max} - \min(i_{max}, j_{max}) + 1, \\ i_t^{(2)} &= i_{max} + \min(i_{max} - i_{max}, j_{max} - j_{max}), \\ j_t^{(2)} &= j_{max} + \min(i_{max} - i_{max}, j_{max} - j_{max}) \end{aligned} \quad (21)$$

Formuojame naujas laiko eilutes su naujaisiais indeksais.

Turėsime vienodos trukmės laiko eilutes X' ir Y', kurių maksimumo taškų indeksai sutaps.

2.2.4. Išlyginimas naudojant kryžminę koreliaciją

Kryžminę koreliaciją TBK analizėje panaudojo Burattini ir kt. [3] Jie naudoja šį metodą visų T bangų išlyginimui prieš TBK analizę, kaip pavyzdį naudojant vieną T bangą. Tačiau tuo atveju, jei T bangos iš EKG išskirtos naudojant kitą metodiką, aktualu išlyginti T bangas visuose MSVM atliekamuose palyginimo momentuose.

Kryžminė koreliacija apibrėžiama tokiu būdu

$$c_{xy}(k) = \begin{cases} \sum_{t=1}^{N-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) / N, & k=0, 1, \dots, N-1, \\ \sum_{t=1-k}^N (x_t - \bar{x})(y_{t+k} - \bar{y}) / N, & k=-1, -2, \dots, -(N-1) \end{cases} \quad (22)$$

$$r_{xy}(k) = \frac{c_{xy}(k)}{\sqrt{c_{xx}(0) \cdot c_{yy}(0)}} \quad (23)$$

Didelė kryžminės koreliacijos reikšmė, kai $k=0$ ir mažesnės reikšmės, kai $k \neq 0$ rodo, kad laiko eilutės gerai išlygintos viena kitos atžvilgiu.

Tuo atveju, kai laiko eilučių trukmė nevienoda, stengiamasi taip pastumti vieną laiko eilutę kitos atžvilgiu, kad maksimumas būtų pasiekiamas, kai $k=0$ arba kuo arčiau šio laiko momento.

2.3. Laiko eilučių panašumo nustatymo būdai

2.3.1. Čebyševio atstumas

Nearing ir Verrier [14] savo siūlomame MSVM metode naudoja (24) atstumą. Toks būdas geras tuo atveju, jei T bangos gerai išlygintos, mažai veikiamos triukšmo. TBK įvertinamas kaip maksimalus absoliutinis skirtumas tarp atitinkamų taškų.

$$\Lambda^{(1)} = \max_i |x_i - y_i| \quad (24)$$

2.3.2. Absoliutinis skirtumas tarp maksimumo taškų

Alternatyva (24) atstumui gali būti absoliutinis skirtumas tarp maksimumo taškų. Tuo atveju, jei T bangos išlygintos, bet jų pikai nežymiai pasislinkę vienas kito atžvilgiu, toks atstumas kompensuos poslinkį.

$$\Lambda^{(2)} = |\max_i x_i - \max_j y_j| \quad (25)$$

2.3.3.-2.3.4. Kosinusų panašumas ir koreliacija

Dviejų laiko eilučių panašumui vertinti gali būti naudojami kosinusų panašumas (26) ir koreliacija (27). Kuo laiko eilutės panašesnės, tuo (26) ir (27) arčiau vieneto.

T bangos kaitos atveju, toks panašumo vertinimas gali duoti informacijos, jei keičiasi T bangos forma.

$$\Lambda^{(3)} = \text{kosinusai}(X, Y) = \frac{X \cdot Y}{\|X\| \cdot \|Y\|} = \frac{\sum_{i=1}^N x_i \cdot y_i}{\sqrt{\sum_{i=1}^N x_i^2} \cdot \sqrt{\sum_{j=1}^N y_j^2}} \quad (26)$$

$$\Lambda^{(4)} = \text{koreliacija}(X, Y) = \frac{(X - \mu) \cdot (Y - \nu)}{\|X - \mu\| \cdot \|Y - \nu\|} = \frac{\sum_{i=1}^N (x_i - \mu) \cdot (y_i - \nu)}{\sqrt{\sum_{i=1}^N (x_i - \mu)^2} \cdot \sqrt{\sum_{j=1}^N (y_j - \nu)^2}} \quad (27)$$

$$\mu = \frac{1}{N} \cdot \sum_{i=1}^N x_i, \quad \nu = \frac{1}{N} \cdot \sum_{j=1}^N y_j$$

2.3.5. Tikėtinumo santykio atstumas

Tikėtinumo santykio atstumą pasiūlė Bagnal, Janacek ir Powell [7] Jie gavo gerus rezultatus pritaikę šį metodą EKG klasifikavimui esant įvairioms aritmijoms.

Šis metodas įvertina laiko eilučių panašumą pasiremdamas tuo, kokie dažniai jose vyrauja. T bangų analizėje toks metodas gali būti pritaikomas tuo atveju, jei iš EKG pakankamai gerai pašalintas triukšmas.

Laiko eilutės X, Y gali būti išreiškiamos kaip tiesinė sinusų ir kosinusų kombinacija atlikus diskrečiąją Furjė transformaciją.

$$x_t = \sum_{k=1}^N p_k \cos(2\pi w_k t) + q_k \sin(2\pi w_k t),$$

$$y_t = \sum_{k=1}^N r_k \cos(2\pi w_k t) + s_k \sin(2\pi w_k t),$$

$$t = 1, \dots, N. \quad (28)$$

Kiekvienam $i=1, \dots, N$, paskaičiuojami a_i, b_i . Prieš atliekant diskrečiąją Furjė transformaciją, iš laiko eilučių atėmus jų vidurkius, pirmasis koeficientas bus lygus 0, todėl $i=1$ galima nenaudoti.

$$a_i = p_i^2 + q_i^2$$

$$b_i = r_i^2 + s_i^2 \quad (29)$$

Kiekvieną a_i galima laikyti nepriklausomo atsitiktinio dydžio A_i , pasiskirsčiusio pagal eksponentinį skirstinį, realizacija.

$$g(a) = \frac{1}{2\alpha_i} \exp\left(\frac{-a}{2\alpha_i}\right),$$

$$i = 2, 3, \dots, N$$
(30)

Tikėtinumo funkcija bus

$$L(\mathbf{a}) = \prod_{i=2}^N \frac{1}{2\alpha_i} \exp\left(\frac{-a_i}{2\alpha_i}\right) .$$
(31)

Tikėtinumo funkcijos logaritmas

$$l(\mathbf{a}) = \sum_{i=2}^N \frac{a_i}{2\alpha_i} \log(2\alpha_i) .$$
(32)

Dvi laiko eilutės ekvivalenčios, jei A_i , B_i turi tą patį pasiskirstymą visiems i .

Formuluojamas tikėtinumo santykis λ .

$$\lambda = \frac{\prod_{i=2}^N \frac{1}{2\hat{\alpha}_i} \exp\left(\frac{-a_i}{2\hat{\alpha}_i}\right) \prod_{i=2}^N \frac{1}{2\hat{\beta}_i} \exp\left(\frac{-b_i}{2\hat{\beta}_i}\right)}{\prod_{i=2}^N \frac{1}{2\tilde{\alpha}_i} \exp\left(\frac{-a_i}{2\tilde{\alpha}_i}\right) \prod_{i=2}^N \frac{1}{2\tilde{\beta}_i} \exp\left(\frac{-b_i}{2\tilde{\beta}_i}\right)} .$$
(33)

$$\hat{\alpha}_i = \hat{\beta}_i = \frac{1}{2}(a_i + b_i),$$

$$\tilde{\alpha}_i = a_i, \quad \tilde{\beta}_i = b_i$$
(34)

Išraišius (34) išraiškas į tikėtinumo santykį (33), gaunama išraiška

$$\lambda = \prod_{i=2}^N \left(\frac{2a_i}{a_i + b_i} \right) \left(\frac{2b_i}{a_i + b_i} \right) .$$
(35)

Tada tikėtinumo santykio statistika atrodo taip

$$\Lambda = -2 \ln \lambda = 2 \sum_{i=2}^N \{2 \ln(a_i + b_i) - \ln a_i - \ln b_i - 2 \ln 2\}$$
(36)

Apibrėžiamas atstumas d_L .

$$d_L(X, Y) = 2 \sum_{i \in f_c} \{2 \ln(a_i + b_i) - \ln a_i - \ln b_i - 2 \ln 2\}$$
(37)

Galima parodyti, kad šis atstumas ekvivalentus aritmetinio ir geometrinio a_i ir b_i vidurkių santykio logaritmų sumai.

$$\begin{aligned}
d_L(X, Y) &= 2 \sum_{i \in f_c} \{2 \ln(a_i + b_i) - \ln a_i - \ln b_i - 2 \ln 2\} = \\
&= 2 \sum_{i \in f_c} \{\ln(a_i + b_i)^2 - \ln(a_i \cdot b_i) - 2 \ln 2\} = 2 \sum_{i \in f_c} \{\ln(a_i + b_i)^2 - \ln(4 a_i \cdot b_i)\} = \\
&= 2 \sum_{i \in f_c} \left\{ \ln \frac{(a_i + b_i)^2}{4 a_i \cdot b_i} \right\} = \sum_{i \in f_c} \left\{ \ln \frac{(a_i + b_i)}{2} \frac{1}{\sqrt{a_i \cdot b_i}} \right\}
\end{aligned} \tag{38}$$

f_c gali būti parenkamas įvairiai, pvz. imamas norimas skaičius pirmųjų koeficientų arba norimas skaičius didžiausių koeficientų kiekvienoje sekoje.

Jei pasirenkamas norimas skaičius didžiausių koeficientų kiekvienoje sekoje, atstumą galima apibrėžti taip:

$$\begin{aligned}
\Lambda^{(5)} = d_L^{(2)}(X, Y) &= \frac{(d_L(X, Y) + d_L(Y, X))}{2} = \\
&= \sum_{i \in f_{c,1}} \{2 \ln(a_i + b_i) - \ln a_i - \ln b_i - 2 \ln 2\} + \sum_{j \in f_{c,2}} \{2 \ln(a_j + b_j) - \ln a_j - \ln b_j - 2 \ln 2\}
\end{aligned} \tag{39}$$

Čia $f_{c,1}$ - didžiausių koeficientų indeksai X laiko eilutei, $f_{c,2}$ - didžiausių koeficientų indeksai Y laiko eilutei.

2.3.6. L_p normos apibendrinimas laiko eilutėms

L_p norma neatsižvelgia į laiko eilučių struktūrą, tai yra prarandama informacija apie laiko eilučių eiliškumą, nes koordinatinių tvarka skaičiuojant Minkowski normą nėra svarbi.

Lee ir Verleysen [8] pasiūlė L_p apibendrinimą laiko eilutėms, kai skaičiuojant įtraukiama ankstesnė ir tolimesnė koordinatės reikšmė, o ne tik konkreti koordinatė.

$$L_p^{TS} = \left(\sum_{i=1}^N (A_{i-1} + A_{i+1})^p \right)^{\frac{1}{p}}, \text{ kur} \tag{40}$$

$$A_{i-1} = \begin{cases} \frac{\tau}{2}|x_i|, & \text{jei } 0 \leq x_i x_{i-1} \\ \frac{\tau}{2} \frac{x_i^2}{|x_i| + |x_{i-1}|}, & \text{jei } 0 > x_i x_{i-1} \end{cases} \quad (41)$$

$$A_{i+1} = \begin{cases} \frac{\tau}{2}|x_i|, & \text{jei } 0 \leq x_i x_{i+1} \\ \frac{\tau}{2} \frac{x_i^2}{|x_i| + |x_{i+1}|}, & \text{jei } 0 > x_i x_{i+1} \end{cases}$$

x_0 ir x_{N+1} yra laikomi lygūs 0, τ - imties periodas.

Lee – Verleysen atstumas tarp dviejų laiko eilučių atrodo taip.

$$\Lambda^{(6)} = L_p^{TS2} = \left(\sum_{i=1}^N (B_{i-1} + B_{i+1})^p \right)^{\frac{1}{p}} \quad (42)$$

$$b_i = y_i - x_i$$

$$B_{i-1} = \begin{cases} \frac{\tau}{2}|b_i|, & \text{jei } 0 \leq b_i b_{i-1} \\ \frac{\tau}{2} \frac{b_i^2}{|b_i| + |b_{i-1}|}, & \text{jei } 0 > b_i b_{i-1} \end{cases} \quad (43)$$

$$B_{i+1} = \begin{cases} \frac{\tau}{2}|b_i|, & \text{jei } 0 \leq b_i b_{i+1} \\ \frac{\tau}{2} \frac{b_i^2}{|b_i| + |b_{i+1}|}, & \text{jei } 0 > b_i b_{i+1} \end{cases}$$

2. 4. Diskrečioji vilnelių transformacija (DBT)

2.4.1. Diskrečioji vilnelių transformacija

Vilnelės – tai tam tikrų funkcijų, vadinamų dukterinėmis vilnelėmis rinkiniai, gaunami iš konkrečios funkcijos vadinamos motinine vilnele.

Vilnelių ir diskrečiosios vilnelių transformacijos apibrėžimas pagal Adison [5].

$$\begin{aligned} \psi^{a,b}(t) &= |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), \\ \int_{-\infty}^{\infty} \psi(x) dx &= 0 \end{aligned} \quad (44)$$

Mastelio funkcija

$$\phi^{a,b}(t) = 2^{-\frac{a}{2}} \cdot \phi(2^{-a}t - b) \quad (45)$$

Vilnelių arba detalių koeficientai

$$T^{a,b} = \int_{-\infty}^{+\infty} x_t \cdot \psi^{a,b}(t) dt \quad (46)$$

Aproksimacijų koeficientai

$$S^{a,b} = \int_{-\infty}^{+\infty} x_t \cdot \phi^{a,b}(t) dt \quad (47)$$

Diskreti laiko eilutės aproksimacija m-tajam masteliui atrodo taip

$$x_t^{(m)} = \sum_{n=-\infty}^{\infty} S^{m,n} \phi^{m,n}(t) \quad (48)$$

Diskreti laiko eilutė x_t išskaidoma į

$$x_t = \sum_{n=-\infty}^{\infty} S^{m_0,n} \phi^{m_0,n}(t) + \sum_{m=-\infty}^{m_0} \sum_{n=-\infty}^{\infty} T^{m,n} \psi^{m,n}(t) \quad (49)$$

$$d_t^{(m)} = \sum_{n=-\infty}^{\infty} T^{m,n} \psi^{m,n}(t) \quad (50)$$

Signalų aproksimacija tam tikram masteliui m .

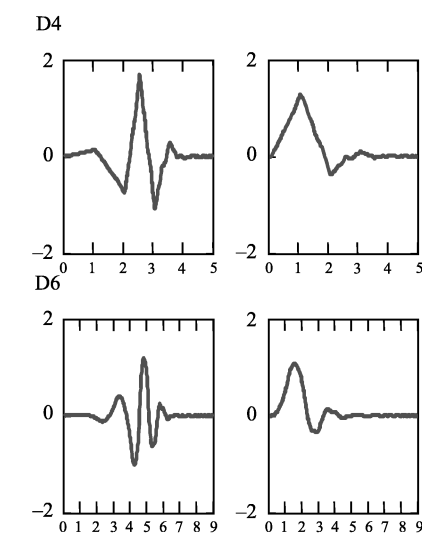
$$\begin{aligned} x_t &= x_t^{m_0} + \sum_{m=-\infty}^{m_0} d_t^m, \\ x_t^{(m)} &= x_t^{(m-1)} - d_t^{(m)} \end{aligned} \quad (51)$$

Kai signalas yra baigtinio ilgio, formules (51) galima perrašyti kaip (52). Praktiškai $S^{0,n}$

yra N ilgio, $N = 2^M$. Galimas mastelių skaičius yra $0 < m < M$, o m -tojo mastelio vilnelių koeficientų skaičius bus $N^{(m)} = 2^{M-m}$.

$$\begin{aligned} x_t^0 &= x_t^M + \sum_{m=1}^M d_t^m, \\ x_t^M &= S^{M,n} \phi_t^{M,n} \end{aligned} \quad (52)$$

$$d_t^m = \sum_{n=0}^{2^{M-m}-1} T^{m,n} \psi^{m,n}(t) \quad (53)$$



5 pav. Daubechies vilnelių D4 ir D6 motininės ir mastelio funkcijos

2.4.2. Vilnelių panaudojimas EKG triukšmo ir izoelektrinės linijos šalinimui

EKG triukšmui šalinti, izoelektrinės linijos dreifui koreguoti ir T bangoms išskirti gali būti naudojama diskreti vilnelių transformacija.

Triukšmui šalinti Alfaouri ir Daqrouq [4] naudoja Daubechies vilnelę D4.

Pirmiausia EKG signalas išskirstomas į 5 lygius, tuomet pasirinktam skaičiui lygių, (dažniausiai $m=1$ arba $m=1,2$, nes šie lygiai atstovauja didžiausio dažnio komponentes), atliekamas triukšmo šalinimas. Kiekvienam pasirinktam lygiui jo koeficientams taikomas slenkstis (54) ir atliekama transformacija (55).

$$T = \sigma \cdot \sqrt{(2 \cdot \ln(n))} \quad (54)$$

$$d(T^{j,k}) = \begin{cases} 1, & T^{j,k} \geq T \\ 0, & T^{j,k} < T \end{cases} \quad (55)$$

*T^{j,k} – j-ojo lygio k-tasis koeficientas,
d(T^{j,k}) – j-ojo lygio k-tasis koeficientas po transformacijos,
T – slenkstis*

Tuomet atliekama atvirkštinė vilnelių transformacija (ADBT) ir gaunamas EKG signalas pašalinus triukšmą.

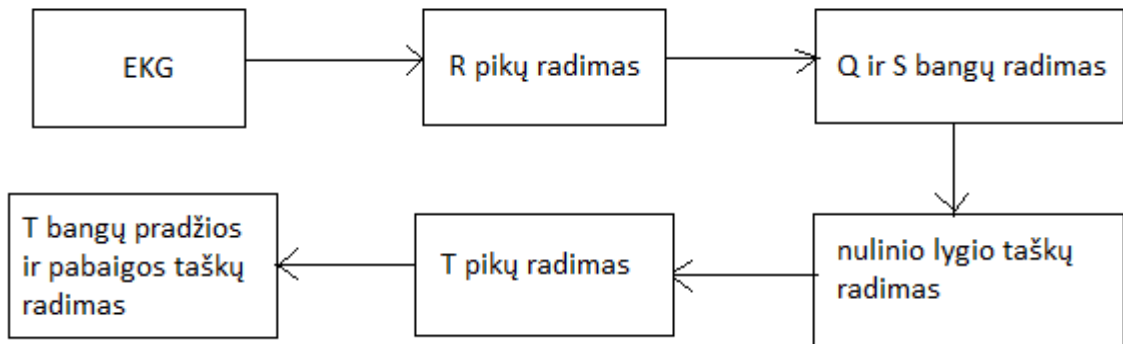
Izoelektrinės linijos dreifas yra žemo dažnio EKG komponentė. Jam pašalinti naudojama Daubechies D6 vilnelė. Signalas yra padalinamas į 8 lygius. 8-oji lygio aproksimacijų koeficientams priskiriamos nulinės reikšmės, t. y. $d(S^{8,k})=0$. Tada atliekama atvirkštinė vilnelių transformacija.

2.4.3. Vilnelių panaudojimas T bangų išskyrimui

T bangų išskyrimas yra atliekamas taip pat taikant diskrečiąją vilnelių transformaciją pagal Mahmoodabadi ir kt [6] pasiūlytą T bangų radimo algoritmą.

Dėl panašumo į EKG bangas, pasirenkama Daubechies D6 vilnelė ir signalas išskaidomas į 8 lygius. Ieškant EKG taškų ir bangų, signalas atstatomas tam tikrų lygių koeficientus paverčiant nuliais. Paliekami tie lygiai, kuriuose vyrauja dažniai atitinkantys ieškomų EKG elementų dažnį. Tokiu būdu atliekama signalo filtracija ir ieškomi elementai tampa pastebimesni.

Paieškos eigos algoritmas vienai EKG.



6 pav. T bangų išskyrimo algoritmo schema

R pikų radimas

Iš EKG signalo pašalinami žemi ir aukšti dažniai, paliekant tik komponentes $m=3, 4, 5$. Filtruotas signalas pakeliamas kvadratu.

$$y_i^{(3)} = (y_i^{(2)})^2 \quad (56)$$

$y_i^{(2)}$ – filtruotas signalas

Gautasis signalas vėl koreguojamas pašalinant mažas reikšmes.

$$y_i^{(4)} = \begin{cases} y_i^{(3)}, & \text{jei } y_i^{(3)} \geq a \cdot \max(y_i^{(3)}) \\ 0, & \text{jei } y_i^{(3)} < a \cdot \max(y_i^{(3)}) \end{cases} \quad (57)$$

a – R bangos slenkstis

Kadangi naujas R pikas neturi pasirodyti anksčiau, nei už 250 ms, Aptikus pirmąjį R piką, sekantis pikas ieškomas tik už 250 ms (500 Hz atveju, tai atitiks 125 taškus).

Tuomet fiksuojami indeksai R , kurie atitinka aptiktuosius R pikus.

Q ir S taškų aptikimas

Paliekamos komponentės $m=1, 2, 3, 4, 5$, t. y. paliekami aukšti dažniai. Gautam signalui atliekama Q ir S bangų paieška.

$$y_i^{(5)} \text{ – filtruotas signalas} \quad (58)$$

Ieškoma minimumo taškų 100 ms į kairę ir į dešinę nuo jau aptikto R piko (500 Hz atveju – ± 50 taškų). Fiksuojami indeksai Q ir S , atitinkantys rastus minimumo taškus prieš ir

po R piko.

Nulinio lygio taškų radimas

Paliekamos komponentės $m=1, 2, 3, 4, 5$, tuomet ieškomą pirmųjų nulį kertančių taškų prieš Q ir po S taškų. Vėl išsaugomi nulinio lygio taškų indeksai Z .

T bangų pikų aptikimas

Paliekamos komponentės $m=4, 5, 6, 7, 8$. Maksimumo taškas tarp po R einančių dviejų anksčiau fiksuotų nulinio lygio taškų yra T bangos pikas. Išsaugomi T bangos piką atitinkantys indeksai T_{pikas} .

T bangų pradžios ir pabaigos taškų aptikimas

Paliekamos komponentės $m=4, 5, 6, 7, 8$. Ieškoma minimumo taškų tarp pirmųjų po R pikų einančių nulinio lygio taškų ir T pikų taškų bei tarp T pikų taškų ir pusės RR intervalo ilgio atkarpos už R piko.

2.5. Statistinės analizės metodai

2.5.1. Išskirčių paieška

Išskirtys – iš kitų reikšmių toje pačioje imtyje, gautų vienodomis sąlygomis, išsiskiriančios vertės. Tokios reikšmės iškreipia visos imties analizės rezultatus. Išskirtys gali būti įvairios kilmės ir atsirasti dėl skirtingų priežasčių. Pavyzdžiai, duomenų įvedimo klaida (neteisingai atpažintas dešimtainis kablelis ar taškas), neįmanomos reikšmės (neigiamas ar labai didelis amžius), reti įvykiai (didelis TBK lygis sveikam asmeniui).

TBK analizės, kai tiriamos sveikų ir sergančių kardiologinėmis ligomis asmenys, atveju, išskirčių kilmė gali būti sunkiau nustatoma. Išsiskiriančios reikšmės sergantiesiems gali būti ir patologijos, t. y. jų ligos požymis, o gali būti ir EKG įrašymo, signalo skaitmeninimo ar triukšmo, kuris galbūt nebuvo idealiai pašalintas tuo konkrečiu atveju įtakos padarinys. Vis dėlto šalinti išskirtis ligonių grupėms yra rizikinga. Tačiau reikėtų šalinti išskirtis sveikiems žmonėms, ypač jei jos smarkiai skiriasi nuo kitų reikšmių ir tokių išskirčių yra mažai, nes TBK lygiai sveikiems žmonėms neturi būti labai aukšti dėl fiziologinės jų kilmės. Išskirtys sveikų žmonių grupei laikytinos nebūdingu šiai grupei įvykiu.

Medianų metodas

Dixon`o ar Grubbs`o metodai remiasi normalumo prielaida. Neparametrinis, nepriklausantis nuo skirstinio metodas, medianų metodas, pagrįstas medianų skaičiavimu.

Išskirčių paieška vykdoma atliekant tokius etapus [23]:

1. Duomenims skaičiuojama mediana.
2. Skaičiuojamas absoliutinis skirtumas tarp kiekvieno duomenų elemento ir medianos.
3. Skaičiuojama mediana 2 žingsnyje gautiems skirtumams.
4. Skaičiuojamas santykis tarp 2 žingsnio skirtumo ir 3 žingsnio medianos.
5. Jei šis santykis didesnis už kritinę reikšmę, tuomet duomenų elementas laikomas išskirtimi.

Sprendimo taisyklė: (konservatyvi) jei santykis > 2.5 , (švelni) jei santykis > 4.5 ar 5 .

Tarpkvartilinio pločio metodas

Imčiai skaičiuojami kvartiliai Q_1 , Q_3 ir tarpkvartilinis plotis $IQR=Q_3-Q_1$.

Reikšmė laikoma sąlygine išskirtimi, jei priklauso intervalui $[Q_1-3*IQR; Q_1-1.5*IQR]$ arba $[Q_3+1.5*IQR; Q_3+3*IQR]$; reikšmė yra išskirtis, jei ji mažesnė už $Q_1-3*IQR$ arba didesnė už $Q_3+3*IQR$.

2.5.2. Neparametriniai kriterijai [17]

Kendalo ranginis koreliacijos koeficientas

Kendalo ranginis koreliacijos koeficientas naudojamas ranginių kintamųjų ryšio stiprumui įvertinti arba tuo atveju, kai kintamųjų skirstiniai nėra žinomi. Tarkim turime kintamųjų poros

(X, Y) stebėjimus $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

(x_i, y_i) ir $(x_j, y_j), i \neq j$ - suderintosios, jei $(x_i > x_j, y_i > y_j)$ arba $(x_i < x_j, y_i < y_j)$,
nesuderintosios, jei $(x_i > x_j, y_i < y_j)$ arba $(x_i < x_j, y_i > y_j)$.

Susietosios reikšmės – reikšmės, kurioms priskiriami vienodi rangai (pasikartojančios).

Kendalo ranginis koreliacijos koeficientas τ :

$$\tau = \frac{\sum_{i < j} (\text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j))}{\sqrt{((T_0 - T_1)(T_0 - T_2))}},$$

$$T_0 = \frac{n(n-1)}{2}, T_1 = \sum_k \frac{v_k(v_k-1)}{2}, T_2 = \sum_l \frac{u_l(u_l-1)}{2},$$

v_k – kintamojo X reikšmių skaičius k – oje kintamojo X susietųjų reikšmių grupėje ;
 u_l – kintamojo Y reikšmių skaičius l – oje kintamojo Y susietųjų reikšmių grupėje

Tikrinama hipotezė

H : X ir Y nekoreliuoja,

\bar{H} : X ir Y koreliuoja.

Mano, Vitnio ir Vilksosono kriterijus

Duomenų normalumo nereikalaujantis kriterijus, skirtas homogeniškumo hipotezei tikrinti.

Stebimi du atsitiktiniai dydžiai X ir Y . X_1, X_2, \dots, X_{n_1} - atsitiktinio dydžio X paprastoji imtis, Y_1, Y_2, \dots, Y_{n_2} - atsitiktinio dydžio Y paprastoji imtis ir $n_2 \leq n_1$; X pasiskirstymo funkcija $F(x)$, Y pasiskirstymo funkcija $G(x)$.

Tikrinama hipotezė

$H : F(t) \equiv G(t), |t| < \infty$,

\bar{H} : skirtingai nevienodi.

Imtis sujungiama į vieną imtį, išrikiuojama didėjimo tvarką ir suranguojama. A.d. Y imties narių rangai bendrojoje imtyje R_1, R_2, \dots, R_{n_2} .

Mano, Vitnio ir Vilksosono kriterijaus statistika:

$$W = R_1 + R_2 + \dots + R_{n_2}$$

Van der Vardeno kriterijus

Kriterijus skirtas homogeniškumo hipotezei tikrinti.

Turimos dvi nepriklausomos imtys: X_1, X_2, \dots, X_{n_1} - atsitiktinio dydžio X paprastoji imtis, Y_1, Y_2, \dots, Y_{n_2} - atsitiktinio dydžio Y paprastoji imtis ir $n_2 \leq n_1$;

Hipotezė formuluojama taip pat kaip Manio, Vitnio ir Vilksosono kriterijui.

Kriterijaus statistika

$$S = a(R_1) + a(R_2) + \dots + a(R_{n_2}),$$
$$a(R_j) = \Phi^{-1}\left(\frac{R_j}{n+1}\right), \quad (61)$$

$\Phi(x)$ – standartinio normaliojo skirstinio pasiskirstymo funkcija

Kruskalo ir Voliso kriterijus

Kruskalo ir Voliso kriterijus taikomas norint palyginti trijų ar daugiau populiacijų skirstinius. Turime X, Y, Z, \dots - nepriklausomi tolydieji atsitiktiniai dydžiai, kuriuos stebint gautos imtys:

$$(X_1, X_2, \dots, X_{n_1}), (Y_1, Y_2, \dots, Y_{n_2}), (Z_1, Z_2, \dots, Z_{n_3}), \dots$$

Tikrinama hipotezė

H : kintamųjų skirstiniai vienodi;

\bar{H} : skirstiniai skiriasi.

Kriterijaus statistika

$$C = \frac{1}{s^2} \sum_{i=1}^r \frac{(T_i - E E_0(T_i))^2}{n_i},$$

$T_i = \sum R_j$, sumuojama pagal visus narius iš i -osios klasės,
 r – atsitiktinių dydžių skaičius; (62)

$$E_0(T_i) = n_i \bar{a}, \quad \bar{a} = \frac{1}{n} \sum_{j=1}^n R_j, \quad s^2 = \frac{1}{n-1} \sum_{j=1}^n (R_j - \bar{a})^2$$

Ženklių kriterijus

Turima imtis X_1, X_2, \dots, X_n , gauta stebint atsitiktinį dydį X .

Tikrinama hipotezė

H : X mediana lygi u_0 ,

\bar{H} : X mediana nelygi u_0 .

Kriterijau statistika

$$M = \frac{n^+ - n^-}{2}, \quad (63)$$

n^+ – skaičius reikšmių, didesnių už u_0 , n^- – skaičius reikšmių, mažesnių už u_0 , lygios u_0 reikšmės neįtraukiamos

2.5.3. Logistinė regresija

Stebimas įvykis A, kurio tikimybė priklauso nuo kovariančių x_1, x_2, \dots, x_m reikšmių. Atsitiktinis dydis Y įgyja reikšmę 0, kai neįvyksta įvykis A ir 1, kai įvykis A įvyksta. Apibrėžiama tokia tikimybė

$$\begin{aligned} \pi(\mathbf{x}) &= P(Y=1 | \mathbf{x}) = P(A | \mathbf{x}), \\ \mathbf{x} &= (x_0, x_1, \dots, x_m)^T, x_0=1 \end{aligned} \quad (64)$$

Logistinės regresijos modelis:

$$\text{logit}(\mathbf{x}) = \ln \pi \frac{(\mathbf{x})}{1 - \pi(\mathbf{x})} = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m = \boldsymbol{\beta}^T \mathbf{x} \quad (65)$$

Lygybė (56) tuomet galima užrašyti kaip

$$\pi(\mathbf{x}) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}} = \frac{e^{\boldsymbol{\beta}^T \mathbf{x}}}{1 + e^{\boldsymbol{\beta}^T \mathbf{x}}} \quad (66)$$

Turima imtis $(Y_1, \mathbf{x}^{(1)}), (Y_2, \mathbf{x}^{(2)}), \dots, (Y_n, \mathbf{x}^{(n)})$.

Tikrinamos hipotezės

$$H: \beta_1 = \beta_2 = \dots = \beta_m,$$

$$\bar{H}: \exists i, \beta_i \neq 0$$

ir

$$H: \beta_j = 0,$$

$$\bar{H}: \beta_j \neq 0.$$

2.5.4. Kelių biomarkerių apjungimas min-max kombinacija

Nors diagnostinę reikšmę gali turėti ir vienas biomarkeris ar keli biomarkeriai atskirai, tačiau tikslinga panaudoti bent kelis biomarkerius diagnostinei kokybei pagerinti. Tokiu būdu galima apjungti bent kelias skirtingas to paties biomarkerio skaičiavimo metodikas ar pakartotinius kurio nors metodo matavimus.

Populiarus biomarkerio kokybės vertinimo matas yra plotas po ROC kreive. Kuo didesnė jo vertė, tuo geriau. Tuo atveju, kai skirstinys nėra žinomas, maksimizuojama Mano, Vitnio ir Vilkoksono (Mann-Whitney U) statistika, neparametrinis ploto po ROC kreive įvertinys.

Ch. Liu, A. Liu ir Halabi pasiūlė biomarkerių apjungimą min-max kombinacija. [10]

Tariame, kad turime k biomarkerių ir $\mathbf{M}=(M_1, \dots, M_k)$ yra biomarkeriu vertės. Stebimas įvykis A. Apibrėžiamas atsitiktinis dydis D lygus 1, kai įvyksta A, t. y. stebimas ligonis ir 0, kai A neįvyksta, t. y. stebimas sveikas asmuo. $\mathbf{X}=(X_1, \dots, X_k)$ - biomarkerių rezultatų imtis sveikam asmeniui ir $\mathbf{Y}=(Y_1, \dots, Y_k)$ - sergančiam asmeniui, kur jų tankiai yra

$f=f(x_1, \dots, x_k)$ ir $g=g(y_1, \dots, y_k)$. Marginalieji tankiai i -ajam biomarkeriui yra f_i , g_i . Pasiskirstymo funkcijos atitinkamai yra F , G , F_i ir G_i .

Plotas po ROC kreive i -tajam biomarkeriui apibrėžiamas tokia tikimybe

$$AUC_i = P(Y_i - X_i > 0) \quad (67)$$

Jei turime k biomarkerių kombinaciją $h=h(\mathbf{M})$, tuomet plotas po ROC kreive šiai kombinacijai kombinacijai bus

$$AUC = P(h(\mathbf{X}) > h(\mathbf{Y})) \quad (68)$$

Kai skirstiniai yra žinomi, tikėtimumo funkcija atrodo taip:

$$L(\mathbf{M}) = \frac{g(\mathbf{M})}{f(\mathbf{M})} \quad (69)$$

Paprastesnė alternatyva yra tiesinė kombinacija:

$$l(\mathbf{M}) = \lambda_1 M_1 + \lambda_2 M_2 + \dots + \lambda_k M_k \quad (70)$$

$$AUC(\lambda) = P(\lambda_1(Y_1 - X_1) + \dots + \lambda_k(Y_k - X_k) > 0) \quad (71)$$

Min-max kombinacija apibrėžiama tokiu būdu

$$M_{max} = \max_{1 \leq i \leq k} M_i, \quad M_{min} = \min_{1 \leq i \leq k} M_i. \quad (72)$$

Jautrumas

$$\begin{aligned} P(Y_{max} > c) &\geq P(Y_i > c), \\ P(Y_{min} > c) &\leq P(Y_i > c), \\ &\forall 1 \leq i \leq k \end{aligned} \quad (73)$$

Specifiškumas

$$\begin{aligned} P(X_{max} \leq c) &\leq P(X_i \leq c), \\ P(X_{min} \leq c) &\geq P(X_i \leq c), \\ &\forall 1 \leq i \leq k \end{aligned} \quad (74)$$

Galima parodyti, kad (73) lygybės teisingos [10].

$$\begin{aligned} P(Y_{max} > c) &= 1 - P(Y_{max} \leq c) = 1 - P(Y_j \leq c, 1 \leq j \leq k) \geq 1 - P(Y_i \leq c) = P(Y_i > c), \\ P(Y_{min} > c) &= P(Y_j > c, 1 \leq j \leq k) \leq P(Y_i > c), \\ &\forall 1 \leq i \leq k \end{aligned} \quad (75)$$

Panašus ir (74) lygybių įrodymas [10].

$$\begin{aligned} P(X_{max} \leq c) &= P(X_j \leq c, 1 \leq j \leq k) \leq P(X_i \leq c), \\ P(X_{min} \leq c) &= 1 - P(X_{min} > c) = 1 - P(X_j > c, 1 \leq j \leq k) \geq 1 - P(X_i > c) = P(X_i \leq c), \\ &\forall 1 \leq i \leq k \end{aligned} \quad (76)$$

Apibrėžiama tokia tiesinė kombinacija

$$M_\lambda = M_{max} + \lambda M_{min} \quad (77)$$

Tuomet (71) formulę galima pakeiti tokia:

$$AUC(\lambda) = P((Y_{max} - X_{max}) + \lambda(Y_{min} - X_{min}) > 0) \quad (78)$$

Optimalus λ_{opt} , kuris maksimizuoja plotą po ROC kreive.

$$AUC(\lambda_{opt}) = \max_{\lambda} P((Y_{max} - X_{max}) + \lambda(Y_{min} - X_{min}) > 0) \quad (79)$$

λ_{opt} įvertinimas naudojant imtį.

Tarkime $X_i = (X_{i1}, \dots, X_{ki})$, $i = 1, 2, \dots, n$ yra biomarkerių vertės iš sveikų žmonių populiacijos, o $Y_j = (Y_{j1}, \dots, Y_{kj})$, $j = 1, 2, \dots, m$ - ligonių populiacijos.

Neturint informacijos apie skirstinius F , G , maksimizuojama Mano, Vitnio ir

Vilkoksono (Mann-Whitney U) statistika.

$$W(\lambda) = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m I(\lambda_1 Y_{1j} + \dots + \lambda_k Y_{kj} > \lambda_1 X_{1i} + \dots + \lambda_k X_{ki}) \quad (80)$$

Min-max kombinacijai maksimizuojama

$$W(\lambda) = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m I(Y_{j,max} + \lambda Y_{j,min} > X_{i,max} + \lambda X_{i,min}) \quad , \text{ kur} \quad (81)$$

$$\begin{aligned} X_{i,max} &= \max_{1 \leq l \leq k} X_{li}, & X_{i,min} &= \min_{1 \leq l \leq k} X_{li}, \\ Y_{j,max} &= \max_{1 \leq l \leq k} Y_{lj}, & Y_{j,min} &= \min_{1 \leq l \leq k} Y_{lj} \end{aligned} \quad (82)$$

Optimizuojant netinka išvestines naudojančios metodai, nes $W(\lambda)$ nėra tolydi funkcija. Galima pasinaudoti įvairiais optimizavimo metodais, pvz. pso, kuri įgyvendina R paketas *pso*).

3. Duomenų analizė

3.1. Duomenys

3.1.1. TWA duomenų bazė

Duomenų bazėje yra 100 EKG įrašų surinktų iš įvairių duomenų bazių. Jie įrašyti 500 Hz dažniu, 16 bitų rezoliucija ir ± 32 mV režiais. Kiekvienas įrašas turi 12 EKG standartinių derivacijų, tačiau kai kurie turi 2 arba 3 derivacijas.

Duomenų bazę sudaro žmonės turintys įvairių sutrikimų, tokių kaip miokardo infarktas, skilvelinės aritmijos ar yra pripažinti turintys staigios mirties riziką, taip pat sveiki individai. Duomenų bazėje yra ir 32 sintetiniai EKG signalai.

Visiems EKG suranguoti pagal TBK lygį juose, remiantis medicininę įranga ir profesionalių medikų komentarais.

3.1.2. PTB duomenų bazė

Duomenų bazėje yra surinkta sveikų ir sergančių įvairiomis ligomis asmenų EKG. Kiekviena įrašą sudaro 16 derivacijų: 12 standartinių EKG derivacijų ir 3 Franko XYZ derivacijos. Imties dažnis yra 1000 Hz, o rezoliucija - ± 16.384 mV.

Prieš analize imties dažnis sumažinamas iki 500 Hz.

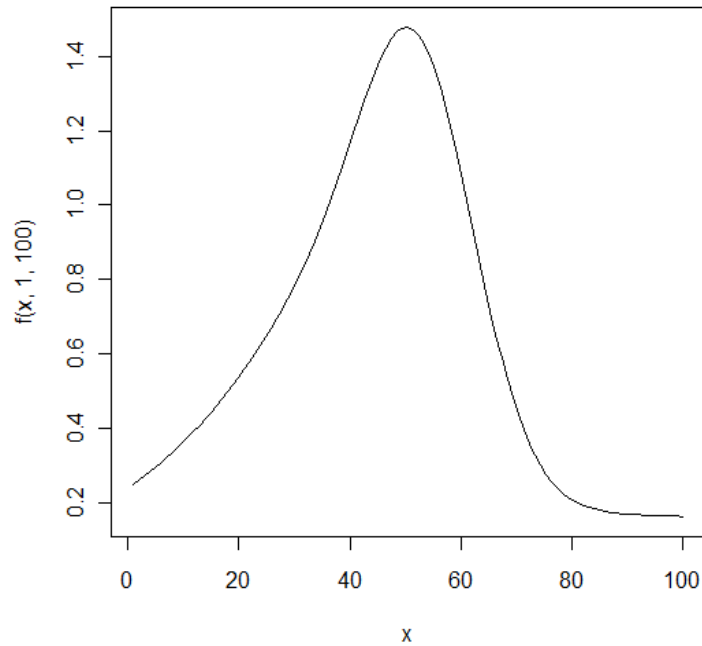
Kategorija	Kiekis
Sveikų žmonių EKG (HC)	48
Miokardo infarktas (MI)	50
Hiso pluošto kojyčių blokada (BBB)	15
Kardiomiopatija (CARDIO)	15
Aritmija (DI)	10

3.1.3. Generuoti duomenys

Pasirinktas vienas T bangos pavyzdys iš tikros EKG ir panaudojus Bernstein daugianarius (Bernstein Polynomials) [20], gauta T bangą aproksimuojanti funkcija.

$$f(z) \approx \sum_{k=1}^N C_N^k y_k^{prad} (z-a)^k (b-z)^{N-k}, \quad z \in [a, b], \quad y_k^{prad} - \text{pradinė } T \text{ banga} \quad (83)$$

$$y_{i,t} = f_i(t) \approx \sum_{k=1}^{N_i} C_{N_i}^k y_k^{prad} (t-1)^k (N_i-t)^{N_i-k}, \quad t=1, \dots, N_i \quad (84)$$



7 pav. generuota T banga

T bangos ilgiai (trukmės) keičiami naudojantis formule (84), o TBK lygiai - Hann lango funkciją (85).

$$w_i(x_i) = 0.5 \left(1 - \cos\left(\frac{2\pi x_i}{N_i - 1}\right) \right), 0 \leq x_i \leq N_i - 1, \quad (85)$$

$$w_i(x_i) \in [0, 1], \quad i = 1, \dots, K$$

T bangos amplitudė keičiama, keičiant $A^{(TBK)}$ Triukšmo lygiai keičiami generuojant triukšmą $B^{(TR)} \sim N(0, \sigma_{TR}^2)$ kaip normalų atsitiktinį dydį, keičiant σ_{TR} reikšmes.

Taip pat keičiamas T bangos ilgis $N_i^{(r)}$, kur $i = 1, \dots, k$ - T bangos numeris, $r = 1, 2, 3, 4$ - ritmo pavyzdžio numeris.

Rezultatas yra dirbtinių, simuliuotų T bangų rinkinys.

$$y_{i,k}^{(r)} = \begin{cases} y_{2 \cdot l, k}^{(r)} = f_{2 \cdot l}^{(r)}(k) + A^{(TBK)} \cdot w_{2 \cdot l}(k-1) + B_k^{(TR)}, & k = 1, \dots, N_{2 \cdot l}^{(r)}, \quad l = 1, \dots, \frac{K}{2} \\ y_{2 \cdot l - 1, k}^{(r)} = f_{2 \cdot l - 1}^{(r)}(k) + B_k^{(TR)}, & k = 1, \dots, N_{2 \cdot l - 1}^{(r)}, \quad l = 1, \dots, \frac{K}{2} \end{cases} \quad (86)$$

TBK lygiai, $A^{(TBK)}$: 0 μ V, 2 μ V, 5 μ V, 10 μ V, 20 μ V, 50 μ V, 100 μ V, 200 μ V, 500 μ V, 1000 μ V.
Triukšmo lygiai, σ_{TR} : 0 μ V, 5 μ V, 10 μ V, 20 μ V, 50 μ V, 100 μ V, 200 μ V, 300 μ V.

Ritmo pavyzdžiai, r (žr. priedą 4): 1 – vienodas intervalas; 2 – normalaus sinusinio ritmo pavyzdys; 3 – ritmo pavyzdys esant TBK, 4 – ritmo pavyzdys prieš skilvelių tachikardijos įvykį.

Kiekvienai kombinacijai generuojama $K=32$ bangos. Iš viso gaunama 320 rinkinių po 32 bangas.

3.2. Duomenų paruošimas analizei

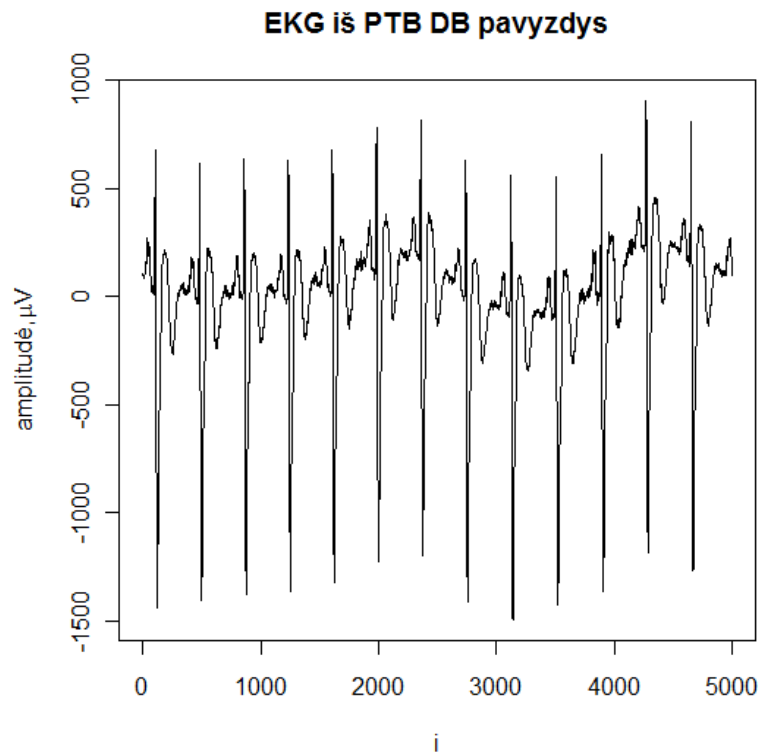
Triukšmo šalinimas ir T bangų iš visos EKG išskyrimas yra svarbi užduotis atliekant TBK analizę. Prieš ją atliekant reikia EKG duomenys atitinkamai paruošti analizei.

Išskirtini tokie etapai:

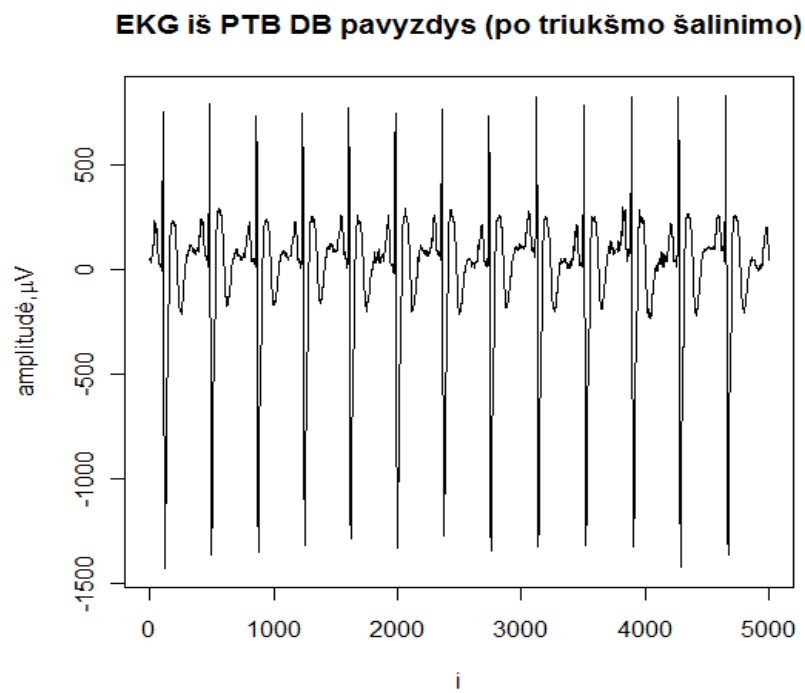
- triukšmo šalinimas;
- izoelektrinės linijos dreifo korekcija;
- T bangų išskyrimas:
 - R pikų aptikimas, QRS kompleksų aptikimas;
 - nulinių taškų aptikimas;
 - T bangų pikų aptikimas;
 - T bangos pradžios ir pabaigos taškų aptikimas;

3.2.1. Triukšmo ir izoelektrinės linijos dreifo šalinimas

EKG iš PTB ir dalis iš TWA duomenų bazių yra realių matavimų rezultatai, t. y. gali būti paveiktos triukšmo. Kiekvienai EKG pasinaudojant diskrečiąja vilnelių transformacija pašalinamas triukšmas bei izoelektrinės linijos dreifas.



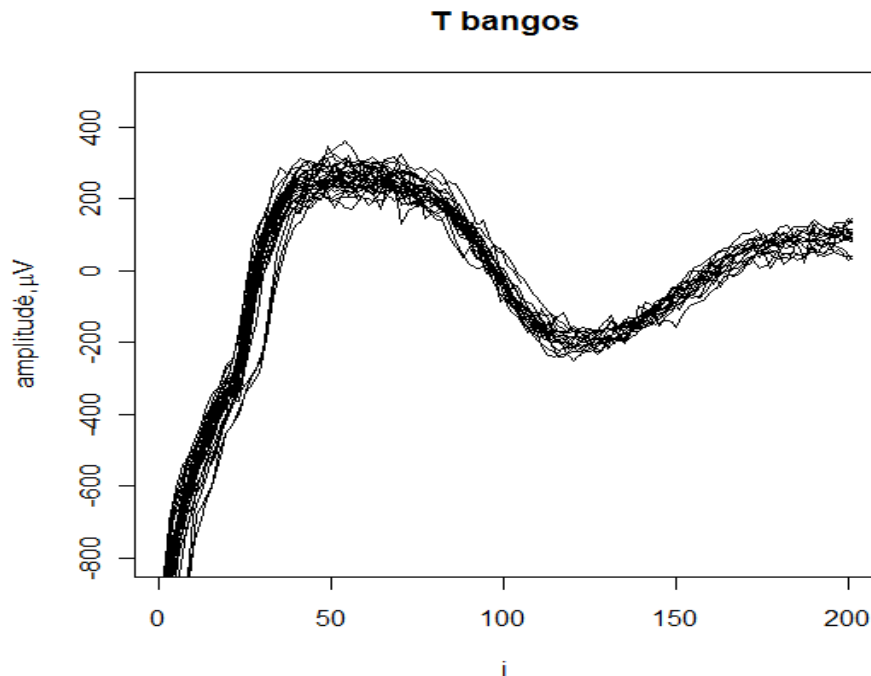
8 pav. dalis EKG V5 derivacijos prieš triukšmo ir izoelektrinės linijos šalinimą



9 pav. dalis EKG V5 derivacijos po triukšmo ir izoelektrinės linijos šalinimo

3.2.2. T bangų išskyrimas

Atlikus triukšmo šalinimą, kiekvienai EKG išskiriamos T bangos.



10 pav. T bangos išskirtos iš EKG V5 derivacijos

3.3. TWA duomenų bazės ir generuotų duomenų analizė

Kiekvienam T bangų rinkiniui taikomas modifikuotas slenkančio vidurkio metodas naudojant įvairius laiko eilučių išlyginimo ir panašumo nustatymo būdus.

Išlyginimo būdai: 1 – DLSK, 2 – atviros pradžios ir pabaigos DLSK su asimetriniu judėjimo šablonu, 3 – atviros pradžios ir pabaigos DLSK su simetriniu judėjimo šablonu, 4 – išlyginimas pikų atžvilgiu, 5 – išlyginimas naudojant kryžminę koreliaciją.

Panašumo nustatymo būdai: 1 – Čebyševio atstumas, 2 – absoliutinis skirtumas tarp maksimumo taškų, 3 – kosinusų panašumas, 4 – koreliacija, 5 – tikėtimumo santykio atstumas $f_c = 5$, 6 – tikėtimumo santykio atstumas $f_c = 10$, 7 – L_p normos apibendrinimas laiko eilutėms.

Viso gaunami 35 modifikuoto slenkančio vidurkio pakeitimai. Gautieji metodai žymimi: metodas $XY - X$ – išlyginimo metodas, Y – panašumo nustatymo metodas.

TBK verte, kiekvienam iš 35 metodų, laikoma TBK16, t. y. to metodo duodamas rezultatas, kai $k=16$, k apibėžtas (2).

TWA duomenų bazėje kiekvienai EKG yra suteiktas rangas pagal TBK lygį kiekvienoje iš jų. Visą TWA duomenų bazę galima padalinti į dvi dalis: simuliuotų EKG dalis ir tikrų EKG dalis.

Duomenys suranguojami pagal kiekvienu būdų gautus TBK vertinimus ir paskaičiuojama Kendalo koreliacija tarp žinomų rangų ir gautųjų rangų.

Metodai, kuriems koreliacijos yra didžiausios yra pasirenkami kaip galimi biomarkeriai TBK analizėje.

Generuotų duomenų rinkiniai suranguojami pagal žinomą TWA lygį. Tada kiekvienam rinkiniui paskaičiuojami TBK vertinimai ir duomenų rinkiniai suranguojami pagal kiekvieną iš jų. Žinant, kad triukšmas veikia kiekvieną iš metodų, geresniu metodu laikomas tas, kuris geriausiai išlaiko tvarką pagal TBK lygius, t. y. kuriam Kendalo koreliacija yra didžiausia.

Generuotiems duomenims geriausiai pasirodė metodas22 (koreliacija 0.7245, p-reikšmė <0.0001), metodas52 (koreliacija 0.7344, p-reikšmė <0.0001), didžiausią koreliaciją duoda metodas42 (koreliacija 0.7786, p-reikšmė <0.0001)

TWA duomenų bazės atveju tikrinta koreliacija simuliuotoms bei tikroms EKG atskirai bei bendram rinkiniui. TBK vertinimui pasirinkta V5 derivacija.

Simuliuotiems duomenims metodas22 duoda koreliaciją 0.5847, p-reikšmė <0.0001 , metodas52 (koreliacija 0.0645, p-reikšmė 0.6152), metodas42 (0.4677, p-reikšmė 0.0002). Tikriems duomenims – metodas22 (koreliacija 0.0986, p-reikšmė 0.3571, metodas52 (koreliacija 0.2204, p-reikšmė 0.0383), metodas42 (0.0831, p-reikšmė 0.4387). Bendram rinkiniui – metodas22 (koreliacija 0.0645, p-reikšmė 0.4155), metodas52 (koreliacija 0.2346, p-reikšmė 0.0029), metodas42 (-0.0277, p-reikšmė 0.7281).

Nors metodas22 tikrų bei bendrai duomenų grupėms nepasirodo pakankamai gerai, bet jis paliekamas tolesnei analizei, nes simuliuotų duomenų grupėje pasiekia didžiausią koreliaciją bei gana stiprią koreliaciją generuotiems duomenims. Metodas52 duoda neblogą koreliaciją tiek generuotiems duomenims, tiek tikrų bei bendroje grupėse iš TWA duomenų bazės. Metodas42 duoda didžiausią koreliaciją generuotiems duomenims, bet mažesnes koreliacijas nei kiti du atrinktieji metodai TWA duomenų bazės duomenims.

Metodai: metodas22 ir metodas52 pasirenkami tolesnei analizei, kuri bus atliekama su EKG iš PTB duomenų bazės.

3.4. PTB duomenų analizė

3.4.1. Pradinė analizė

PTB duomenų bazėje analizė atliekama naudojant tuos metodus, kurie geriausiai pasirodė TWA duomenų bazės ir generuotiems duomenims. Tikslas yra nustatyti ar TBK paskaičiuoti šiais metodais gali būti biomarkeriais atskiriant sveikus ir sergančius kardiologinėmis ligomis pacientus. TBK skaičiuojami V5 derivacijai.

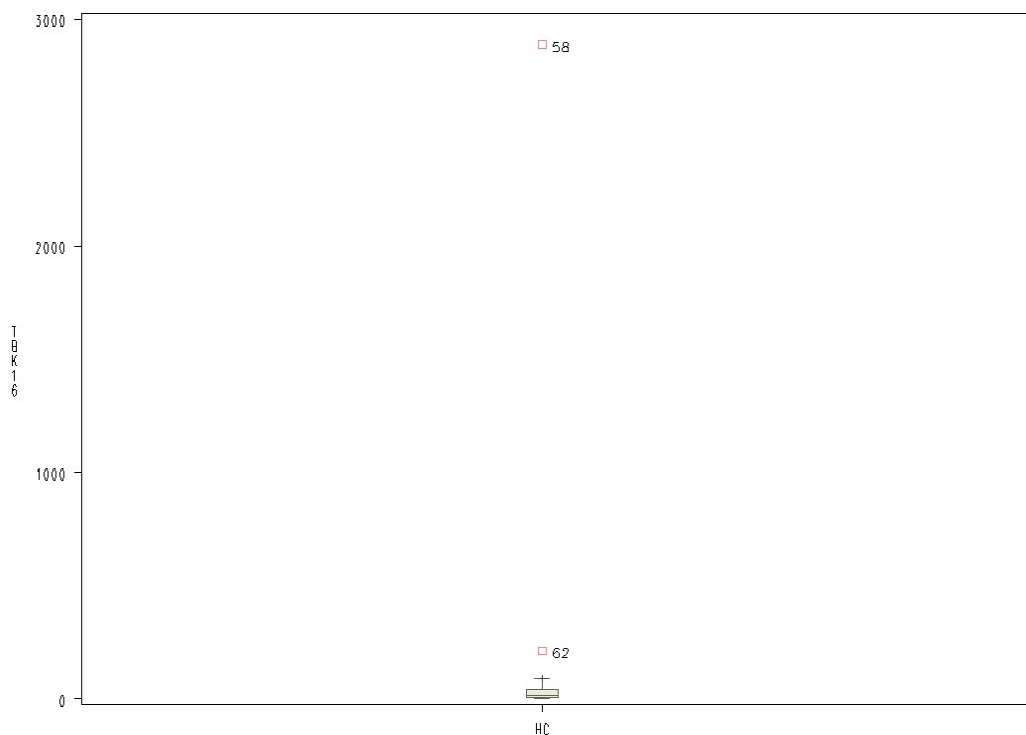
Analysis Variable : TBK16

kate gori ja	N Obs	Variance	N	Mean	Std Dev	Minimum	Maximum	Median	Lower Quartile	Upper Quartile	Quartile Range
BBB	15	264319.7601	15	231.4051	514.1204	3.8494	1767.8241	53.0054	27.2311	64.1230	36.8919
CAR DIO	15	331040.5851	15	195.3439	575.3613	2.6082	2267.1337	40.2710	9.0490	93.8089	84.7599
DI	10	13713.1486	10	82.8803	117.1032	2.0507	380.7815	32.1062	16.2262	104.9566	88.7304
HC	48	172108.2352	48	88.8550	414.8593	0.5596	2891.9282	13.7451	7.6518	42.2302	34.5783
MI	50	2837.9081	50	57.5683	53.2720	4.0787	271.4840	45.0146	30.3210	74.6888	44.3677

1 lentelė. Metodas22. TBK16 skaitinės charakteristikos.

Išskirčių paieška

Sveikam žmogui įtartinai didelė maksimali TBK16 vertė. Iš stačiakampės diagramos (10 pav.) matyti, kad yra bent viena išsiskirianti reikšmė.



11 pav. metodas22. Stačiakampė diagrama sveikų žmonių grupei

Medianų metodas siūlo išskirtimis laikyti

Obs	kategorija	nr	TBK16	sk1	mediana1	mediana2	sk2	rez
4	HC	54	81.46	67.72	13.7451	12.1413	5.578	ISSKIRTIS
6	HC	56	88.74	75.00	13.7451	12.1413	6.177	ISSKIRTIS
8	HC	58	2891.93	2878.18	13.7451	12.1413	237.056	ISSKIRTIS
12	HC	62	213.23	199.49	13.7451	12.1413	16.430	ISSKIRTIS
22	HC	72	92.05	78.30	13.7451	12.1413	6.449	ISSKIRTIS
43	HC	93	79.99	66.25	13.7451	12.1413	5.456	ISSKIRTIS

2 lentelė. Metodas22. Medianų metodas

Kvartilių metodas

Obs	kategorija	nr	TBK16	Q1	Q3	IQR	rez
8	HC	58	2891.93	7.65184	42.2302	34.5783	ISSKIRTIS
12	HC	62	213.23	7.65184	42.2302	34.5783	ISSKIRTIS

3 lentelė. Metodas22. Kvartilių metodas.

Iš imties pašalinamos šios dvi išskirtys.

Po išskirčių pašalinimo

Analysis Variable : TBK16

kate gori ja	N Obs	Variance	N	Mean	Std Dev	Minimum	Maximu m	Median	Lower Quartile	Upper Quartile	Quartile Range
BBB	15	264319.7601	15	231.4051	514.1204	3.8494	1767.8241	53.0054	27.2311	64.1230	36.8919
CAR DIO	15	331040.5851	15	195.3439	575.3613	2.6082	2267.1337	40.2710	9.0490	93.8089	84.7599
DI	10	13713.1486	10	82.8803	117.1032	2.0507	380.7815	32.1062	16.2262	104.9566	88.7304
HC	46	668.7597	46	25.2148	25.8604	0.5596	92.0457	12.9630	7.5219	34.0535	26.5315
MI	50	2837.9081	50	57.5683	53.2720	4.0787	271.4840	45.0146	30.3210	74.6888	44.3677

4 lentelė. Metodas22. TBK16 skaitinės charakteristikos po išskirčių pašalinimo

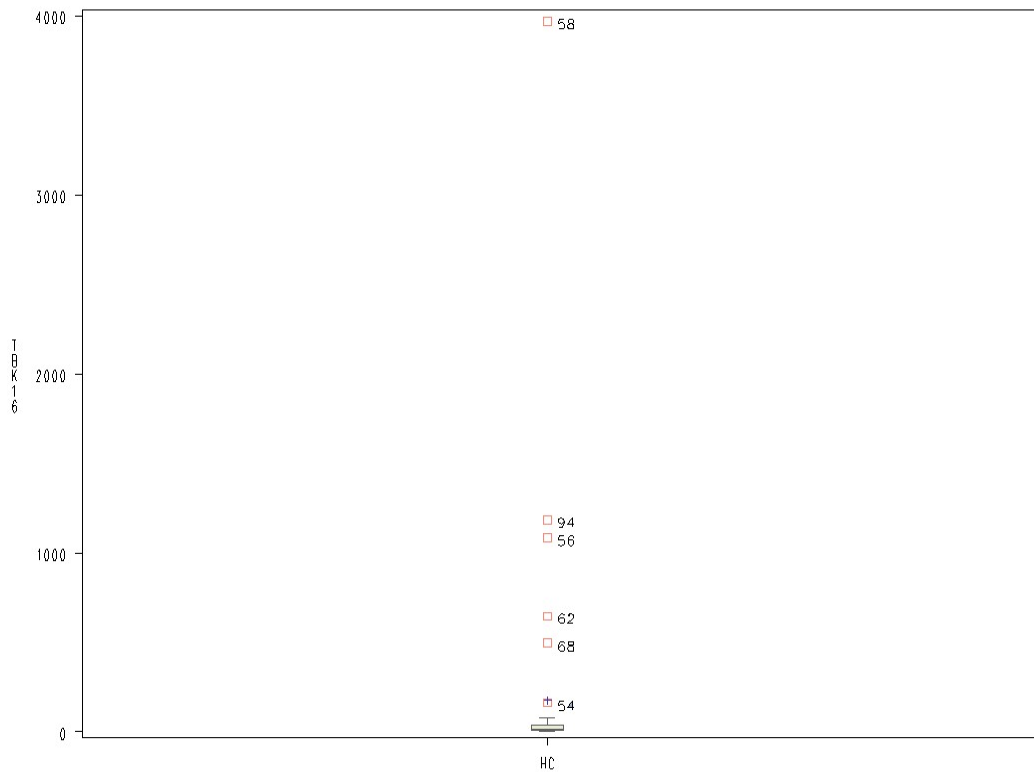
metodas52

Analysis Variable : TBK16

kategor ija	N Obs	Variance	N	Mean	Std Dev	Minim um	Maximu m	Median	Lower Quartil e	Upper Quartile	Quartile Range
BBB	15	604892.1261	15	353.1625	777.7481	11.8433	2939.4069	53.0416	25.2613	123.7434	98.4821
CARDIO	15	316205.2160	15	224.7320	562.3213	5.8684	2238.2599	87.5230	21.6899	119.3316	97.6417
DI	10	7280.1526	10	71.2816	85.3238	1.3465	215.8687	33.4014	8.7545	135.8571	127.1026
HC	48	374593.3958	48	174.5245	612.0404	0.5797	3970.4284	14.1586	7.5319	38.7440	31.2120
MI	50	4482.8064	50	39.4582	66.9538	1.0564	374.0415	23.9094	7.3811	37.7581	30.3771

5 lentelė. Metodas52. TBK16 skaitinės charakteristikos

Šiuo metodu paskaičiavus TBK16 iš stačiakampės diagramos (11 pav.) taip pat matyti, kad sveikų žmonių grupėje yra galimų išskirčių.



12 pav. metodas52. Stačiakampė diagrama sveikų žmonių grupei

Medianų metodas siūlo išskirtimis laikyti

Obs	kategorija	nr	TBK16	sk1	mediana1	mediana2	sk2	rez
4	HC	54	163.45	149.29	14.1586	9.47533	15.755	ISSKIRTIS
6	HC	56	1084.65	1070.49	14.1586	9.47533	112.977	ISSKIRTIS
8	HC	58	3970.43	3956.27	14.1586	9.47533	417.534	ISSKIRTIS
12	HC	62	646.58	632.42	14.1586	9.47533	66.744	ISSKIRTIS
17	HC	67	59.27	45.11	14.1586	9.47533	4.761	ISSKIRTIS
18	HC	68	496.57	482.41	14.1586	9.47533	50.912	ISSKIRTIS
27	HC	77	63.81	49.65	14.1586	9.47533	5.240	ISSKIRTIS
37	HC	87	61.43	47.28	14.1586	9.47533	4.989	ISSKIRTIS
43	HC	93	79.37	65.21	14.1586	9.47533	6.882	ISSKIRTIS
44	HC	94	1184.96	1170.80	14.1586	9.47533	123.563	ISSKIRTIS

6 lentelė. Metodas52. Medianų metodas

Kvartilų metodas

Obs	kategorija	nr	TBK16	Q1	Q3	IQR	rez
4	HC	54	163.45	7.53193	38.7440	31.2120	ISSKIRTIS
6	HC	56	1084.65	7.53193	38.7440	31.2120	ISSKIRTIS
8	HC	58	3970.43	7.53193	38.7440	31.2120	ISSKIRTIS
12	HC	62	646.58	7.53193	38.7440	31.2120	ISSKIRTIS
18	HC	68	496.57	7.53193	38.7440	31.2120	ISSKIRTIS
44	HC	94	1184.96	7.53193	38.7440	31.2120	ISSKIRTIS

7 lentelė. Metodas52. Kvartilų metodas

Pašalinamos trys didžiausios TBK16 reikšmės.

Po išskirčių pašalinimo

Analysis Variable : TBK16

kategorija	N Obs	Variance	N	Mean	Std Dev	Minimum	Maximum	Median	Lower Quartile	Upper Quartile	Quartile Range
BBB	15	604892.1261	15	353.1625	777.7481	11.8433	2939.4069	53.0416	25.2613	123.7434	98.4821
CARDIO	15	316205.2160	15	224.7320	562.3213	5.8684	2238.2599	87.5230	21.6899	119.3316	97.6417
DI	10	7280.1526	10	71.2816	85.3238	1.3465	215.8687	33.4014	8.7545	135.8571	127.1026
HC	45	14125.3716	45	47.4920	118.8502	0.5797	646.5777	13.6997	6.4553	30.7680	24.3126
MI	50	4482.8064	50	39.4582	66.9538	1.0564	374.0415	23.9094	7.3811	37.7581	30.3771

8 lentelė. Metodas52. TBK16 skaitinės charakteristikos po išskirčių pašalinimo

Hipotezė, kad TBK16 skirstinys kiekvienoje iš sveikų ir ligonių grupių yra normalusis atmetama su reikšmingumo lygmeniu 0.05, tiek metodo metodus22, tiek metodus52 atvejais.

Todėl toliau analizėje bus naudojami neparametriniai metodai.

Patikrinama hipotezė apie medianos lygybę 0 grupėje.

Metodo metodus22 atveju, hipotezė apie medianos lygybę nuliui atmetama visose grupėse (BBB grupė ženklų kriterijaus statistika $M=7.5$, p-reikšmė <0.0001 , CARDIO grupė – $M=7.5$, p-reikšmė <0.0001 , DI – $M=5$, p-reikšmė 0.002, HC – $M=23$, p-reikšmė <0.0001 , MI – $M=25$, p-reikšmė <0.0001). Metodo metodus52 atveju, hipotezė taip pat atmetama visose grupėse (BBB $M=7.5$, p-reikšmė <0.0001 , CARDIO grupė – $M=7.5$, p-reikšmė <0.0001 , DI – $M=5$, p-reikšmė 0.002, HC – $M=22.5$, p-reikšmė <0.0001 , MI – $M=25$, p-reikšmė <0.0001).

3.4.2. Homogeniškumo tikrinimas

metodas22

Tikrinama hipotezė, kad TBK16 skirstiniai visose penkiose grupėse yra vienodi.

**Wilcoxon Scores (Rank Sums) for Variable TBK16
Classified by Variable kategorija**

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
BBB	15	1245.0	1027.50	143.948776	83.000000
CARDIO	15	1067.0	1027.50	143.948776	71.133333
DI	10	753.0	685.00	119.937484	75.300000
HC	46	2253.0	3151.00	217.405152	48.978261
MI	50	3998.0	3425.00	221.566393	79.960000

Kruskal-Wallis Test

Chi-Square	17.9158
DF	4
Pr > Chi-Square	0.0013

9 lentelė. Metodas22. Proc npar1way rezultatas

Kruskalo ir Voliso statistika $C=17.9158$, p -reikšmė 0.0013. Hipotezė apie skirtinių vienodumą atmetama su reikšmingumo lygmeniu 0.05. Iš vidutinių rangų matyti, kad mažiausios reikšmės yra sveikiems žmonėms (HC), o didžiausios Hiso kojųčių blokados (BBB) ir miokardo infarkto (MI) grupėse.

Taip pat tikrinama hipotezė, kad TBK16 sveikų žmonių grupėje ir kurioje nors kitoje grupėje skirstiniai sutampa.

**Wilcoxon Scores (Rank Sums) for Variable TBK16
Classified by Variable kategorija**

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
HC	46	1696.0	2231.0	136.351262	36.869565
MI	50	2960.0	2425.0	136.351262	59.200000

Wilcoxon Two-Sample Test

Statistic	1696.0000
Normal Approximation	
Z	-3.9200
One-Sided Pr < Z	<.0001
Two-Sided Pr > Z	<.0001
t Approximation	
One-Sided Pr < Z	<.0001
Two-Sided Pr > Z	0.0002
Z includes a continuity correction of 0.5.	

**Van der Waerden Scores (Normal) for Variable TBK16
Classified by Variable kategorija**

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
HC	46	-18.597572	0.0	4.720291	-0.404295
MI	50	18.597572	0.0	4.720291	0.371951

Van der Waerden Two-Sample Test

Statistic	-18.5976
Z	-3.9399
One-Sided Pr < Z	<.0001
Two-Sided Pr > Z	<.0001

10 lentelė. Metodas22. Proc npar1way rezultatas HC ir MI grupėms

Hipotezė, kad sveikų žmonių grupėje ir miokardo infarkto grupėje TBK16 skirstiniai sutampa atmetama (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=-3.92$, p-reikšmė < 0.001 , Van der Vardeno statistika $S=-18.5976$, asimptotinė Van der Vardeno statistika $Z=-3.9399$, p-reikšmė < 0.0001).

Hipotezė, kad sutampa TBK16 skirstiniai sveikų žmonių ir Hiso kojyčių blokados grupėse atmetama (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=2.8891$, p-reikšmė 0.0039 , asimptotinė Van der Vardeno statistika $Z=2.8861$, p-reikšmė 0.0039).

Hipotezės, kad sutampa TBK16 skirstiniai sveikų žmonių ir kardiopatijos grupėse Mano, Vitnio ir Vilkoksono kriterijus neatmeta (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=1.6665$, p-reikšmė 0.0956), tačiau šią hipotezę atmeta Van der Vardeno kriterijus (asimptotinė Van der Vardeno statistika $Z=1.9871$, p-reikšmė 0.0469).

Mano, Vitnio ir Vilkoksono kriterijus neatmeta hipotezės apie TBK16 skirstinių lygybę sveikų žmonių ir aritmijų grupėse (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=1.9147$, p-reikšmė 0.0555), tačiau šią hipotezę atmeta Van der Vardeno kriterijus (asimptotinė Van der Vardeno statistika $Z=2.1879$, p-reikšmė 0.0287).

metodas52

Tikrinama hipotezė, kad TBK16 skirstiniai visose penkiose grupėse yra vienodi.

**Wilcoxon Scores (Rank Sums) for Variable TBK16
Classified by Variable kategorija**

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
BBB	15	1458.0	1020.0	142.828569	97.200000
CARDIO	15	1399.0	1020.0	142.828569	93.266667
DI	10	739.0	680.0	119.023807	73.900000
HC	45	2504.0	3060.0	214.242853	55.644444
MI	50	3080.0	3400.0	219.469056	61.600000

Kruskal-Wallis Test

Chi-Square	20.6742
DF	4
Pr > Chi-Square	0.0004

11 lentelė. Metodas52. Proc npar1way rezultatas

Naudojant metodą metodus52, Kruskalo ir Voliso statistika $C=20.6742$, p-reiškė 0.0004. Hipotezė atmetama su reikšmingumo lygmeniu 0.05. Mažiausias TBK16 reikšmes turi sveiki žmonės (HC), o didžiausias – Hiso kojų blokados (BBB) ir miokardo infarkto (MI) grupėms priklausantys žmonės.

Tikrinamos hipotezės apie TBK16 skistinių sutapimą tarp sveikų ir kitos grupės, kai TBK16 gaunamas naudojant metodą metodus52.

**Wilcoxon Scores (Rank Sums) for Variable TBK16
Classified by Variable kategorija**

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
HC	45	2051.0	2160.0	134.164079	45.577778
MI	50	2509.0	2400.0	134.164079	50.180000

Wilcoxon Two-Sample Test

Statistic	2051.0000
Normal Approximation	
Z	-0.8087
One-Sided Pr < Z	0.2093
Two-Sided Pr > Z 	0.4187
t Approximation	
One-Sided Pr < Z	0.2104
Two-Sided Pr > Z 	0.4207
Z includes a continuity correction of 0.5.	

Van der Waerden Scores (Normal) for Variable TBK16

Classified by Variable kategorija

kategorija	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
HC	45	-3.096338	0.0	4.691837	-0.068808
MI	50	3.096338	0.0	4.691837	0.061927

Van der Waerden Two-Sample Test

Statistic	-3.0963
Z	-0.6599
One-Sided Pr < Z	0.2546
Two-Sided Pr > Z 	0.5093

12 lentelė. Metodas52. Proc npar1way rezultatas

Hipotezė, kad sveikų žmonių grupėje ir miokardo infarkto grupėje TBK16 skirstiniai sutampa neatmetama (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=-0.8087$, p-reikšmė 0.4187, asimptotinė Van der Vardeno statistika $Z=-0.6599$, p-reikšmė 0.5093).

Hipotezė, kad skirstiniai sutampa taip pat neatmetama sveikų žmonių bei aritmijos grupėse (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=1.1456$, p-reikšmė 0.2519, asimptotinė Van der Vardeno statistika $Z=1.1387$, p-reikšmė 0.2548).

Hipotezė, kad sutampa TBK16 skirstiniai grupėse atmetama sveikų žmonių ir Hiso kojyčių blokados grupėms (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=3.5850$, p-reikšmė 0.0003, asimptotinė Van der Vardeno statistika $Z=3.5763$, p-reikšmė 0.0003) bei sveikų žmonių ir kardiopatija sergančiųjų grupėms (asimptotinė Mano, Vitnio ir Vilkoksono statistika $Z=3.1241$, p-reikšmė 0.0018), tačiau šią hipotezę atmeta Van der Vardeno kriterijus (asimptotinė Van der Vardeno statistika $Z=3.1032$, p-reikšmė 0.0019).

3.4.3. Logistinė regresija

PTB duomenų bazės duomenims mėginama pritaikyti ir logistinės regresijos modelį. Bandomi tokie atvejai: pirmasis - įvykiu laikoma priklausymas kažkuriai iš kardiologinių problemų grupių, neįvykiu – priklausymas sveikų žmonių grupei; kiti atvejai – įvykiu laikoma priklausymas konkrečiai sergančiųjų grupei, neįvykiu – priklausymas sveikų žmonių grupei.

Į modelį įtraukiama TBK16 vertė pacientui ir papildoma informacija: paciento lytis ir amžius.

Logistinės regresijos modelis, kai TBK16 apskaičiuotas naudojant metodą metodas22 yra tinkamas (hipotezė, kad modelis yra netinkamas, atmetama). Ir kovariančių TBK16 (p-reikšmė 0.0015) ir amžius (p-reikšmė <0.0001) įtaka statistiškai reikšminga.

Logistinės regresijos modelis tinka sveikiems ir miokardo infarktą patyrusiems (kovariančių TBK16 (p-reikšmė 0.0061) ir amžius (p-reikšmė <0.0001) įtaka statistiškai reikšminga), sveikiems ir Hiso kojyčių blokadą patiriantiems pacientams (kovariančių TBK16 (p-reikšmė 0.0286) ir amžius (p-reikšmė 0.0468) įtaka statistiškai reikšminga), sveikiems ir sergantiems aritmijomis (kovariančių TBK16 (p-reikšmė 0.0197) ir amžius (p-reikšmė 0.0069) įtaka statistiškai reikšminga), sveikiems ir sergantiems kardiopatija (statistiškai reikšminga tik kovariantė amžius (0.0069), kovariantė TBK16 (p-reikšmė 0.0715) neturi įtakos).

Logistinės regresijos modelis, kai TBK16 apskaičiuotas naudojant metodą metodas52 yra tinkamas, tačiau kovariantės TBK16 (p-reikšmė 0.0015) įtaka statistiškai nereikšminga, statistiškai reikšminga tik kovariantės amžius (p-reikšmė <0.0001) įtaka.

Logistinės regresijos modelis tinka sveikiems ir miokardo infarktą patyrusiems, tačiau kovariantės TBK16 (p-reikšmė 0.2256) įtaka statistiškai nereikšminga, statistiškai reikšminga tik kovariantės amžius (p-reikšmė <0.0001) įtaka), analogiška situacija sveikiems ir Hiso kojyčių blokadą patiriantiems pacientams (kovariančių TBK16 (p-reikšmė 0.1241) ir amžius (p-reikšmė 0.0384)) , sveikiems ir sergantiems aritmijomis logistinės regresijos modelis nėra tinkamas, sveikiems ir sergantiems kardiopatija (statistiškai reikšminga tik kovariantė amžius (p-reikšmė 0.0072), o TBK16 nėra statistiškai reikšminga (p-reikšmė 0.1088)).

3.4.4. Kelių biomarkerių apjungimas

Tuo atveju, kai TBK vertinta naudojant metodą metodus²², sveiki ir sergantys žmonės gali būti atskiriami taikant logistinės regresijos modelį, tačiau lyginant sveikus ir tam tikra konkrečia liga sergančius žmones, TBK16 nedavė papildomos informacijos.

Patikrinama ar apjungus metodus min-max kombinacija, gali būti pagerinamas sveikų ir konkrečia liga sergančiųjų atskyrimas. Bandomi tokie deriniai: TBK16 ir TBK15, TBK16 ir kTBK16, kTBK16 ir kTBK15, TBK16 ir vidTBK, TBK16 ir vidkTBK, kTBK16 ir vidkTBK (TBK15, TBK16 ir vidTBK – 15-oji, 16-oji ir vidutinė TBK vertės, paskaičiuotos metodu metodus²², kTBK15, kTBK16 ir vidkTBK – 15-oji, 16-oji ir vidutinė TBK vertės, paskaičiuotos metodu metodus⁵²).

Maksimizavus išraišką (81), gautos tokios min-max kombinacijos:

$$\text{minmaxkomb1} = \max(\text{TBK16}, \text{TBK15}) + 0.0505 * \min(\text{TBK16}, \text{TBK15});$$

$$\text{minmaxkomb2} = \max(\text{TBK16}, \text{kTBK16}) + 1.5699 * \min(\text{TBK16}, \text{kTBK16});$$

$$\text{minmaxkomb3} = \max(\text{kTBK16}, \text{kTBK15}) + 54.6338 * \min(\text{kTBK16}, \text{kTBK15});$$

$$\text{minmaxkomb4} = \max(\text{TBK16}, \text{vidTBK}) + 150.2663 * \min(\text{TBK16}, \text{vidTBK});$$

$$\text{minmaxkomb5} = \max(\text{TBK16}, \text{vidkTBK}) + 1.7621 * \min(\text{TBK16}, \text{vidkTBK});$$

$$\text{minmaxkomb6} = \max(\text{kTBK16}, \text{vidkTBK}) + 3.6903 * \min(\text{kTBK16}, \text{vidkTBK}).$$

Įtraukus atskirai po vieną iš šių kombinacijų, amžių ir lytį į logistinės regresijos modelį, gauta, kad sveikus nuo sergančiųjų bei sveikus ir kuriai nors konkrečiai sergančiųjų grupei priklausančius geriausiai skiria kombinacijos: minmaxkomb1 ir minmaxkomb4.

Skaitinės charakteristikos

Analysis Variable : minmaxkomb1											
kategorija	N Obs	Variance	N	Mean	Std Dev	Minimum	Maximum	Median	Lower Quartile	Upper Quartile	Quartile Range
BBB	15	291173.3731	15	244.5714	539.6048	6.9332	1857.0992	55.1080	28.3165	83.8100	55.4935
CARDIO	15	373312.8259	15	214.0364	610.9933	4.2839	2414.6239	54.6989	9.3982	98.5462	89.1480
DI	10	14117.1434	10	93.5246	118.8156	15.6324	398.4455	45.7517	21.2244	109.0141	87.7897
HC	44	769.4509	44	30.1146	27.7390	0.5943	110.7880	19.5365	11.7560	38.7371	26.9811
MI	50	3084.0478	50	65.5044	55.5342	5.2669	285.1939	55.9473	33.1675	86.0922	52.9247

13 lentelė. Minmaxkomb1 skaitinės charakteristikos

Analysis Variable : minmaxkomb4											
kategorija	N Obs	Variance	N	Mean	Std Dev	Minimum	Maximum	Median	Lower Quartile	Upper Quartile	Quartile Range
BBB	15	6087941055.0	15	34455.5040	78025.2591	587.0167	267532.8862	4758.5045	3557.4524	9756.6085	6199.1561
CARDIO	15	7589114632.2	15	29478.1201	87115.5246	362.1123	343164.6301	6101.5948	702.9842	14206.6532	13503.6691
DI	10	303533315.74	10	11977.4130	17422.2075	321.9282	56046.5472	4412.5673	2460.5435	15984.6252	13524.0817
HC	44	8242289.1961	44	3008.4844	2870.9387	87.5231	12378.2395	1832.7439	1162.4365	4451.2385	3288.8020
MI	50	43474477.202	50	6552.1992	6593.5178	530.5444	41087.9859	5534.7967	2271.0669	8515.6132	6244.5462

Analysis Variable : minmaxkomb4		
kategorija	N Obs	Quartile Range
BBB	15	6199.1561
CARDIO	15	13503.6691
DI	10	13524.0817
HC	44	3288.8020
MI	50	6244.5462

14 lentelė. Minmaxkomb4 skaitinės charakteristikos

Iš skaitinių charakteristikų matyti, kad minmaxkomb1 yra lengviau interpretuojamas.

Logistinės regresijos modelis, į kurį įtraukta kombinacija minmaxkomb1, amžius ir lytis, yra tinkamas duomenims, hipotezė apie logistinės regresijos netinkamumą atmetama, kovariančių minmaxkomb1 (p-reikšmė 0.001) ir amžius (p-reikšmė <0.0001) įtaka statistiškai reikšminga.

Logistinės regresijos modelis taip pat tinka sveikiems ir miokardo infarktą patyrusiems (kovariančių minmaxkomb1 (p-reikšmė 0.0036) ir amžius (p-reikšmė <0.0001) įtaka statistiškai reikšminga), sveikiems ir Hiso kojų blokadą patiriantiems pacientams (kovariančių minmaxkomb1 (p-reikšmė 0.0239) ir amžius (p-reikšmė 0.0205) įtaka statistiškai reikšminga), sveikiems ir sergantiems aritmijomis (kovariančių minmaxkomb1 (p-reikšmė 0.0093) ir amžius (p-reikšmė 0.0270) įtaka statistiškai reikšminga), sveikiems ir sergantiems kardiopatija (kovariančių minmaxkomb1 (p-reikšmė 0.0357) ir amžius (p-reikšmė 0.0048) įtaka statistiškai reikšminga).

min-max kombinacija ($\text{minmaxkomb1} = \max(\text{TBK16}, \text{TBK15}) + 0.0505 * \min(\text{TBK16}, \text{TBK15})$) skiria sveikus nuo sergančiųjų tiek bendrai, tiek lyginant su kiekviena sergančiųjų grupe, todėl gali būti laikoma biomarkeriu širdies sutrikimų analizėje.

Išvados ir rekomendacijos

Šiame darbe buvo siekiama panagrinėti T bangos kaitą, pakeitus tradicinį modifikuotą slenkančio vidurkio metodą, t. y. pasinaudojant įvairiais laiko eilučių išlyginimo bei panašumo nustatymo būdais ir TBK, kaip galimo biomarkerio širdies sutrikimų diagnostikoje galimybes.

T bangų išlyginimui naudoti įprasto dinaminio laiko skalės kraipymo, atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo (su asimetriniu ir su simetriniu judėjimo šablonu), kryžminės koreliacijos ir išlyginimo pagal maksimumo taškus metodai, o panašumui vertinti – Čebyševio atstumas, absoliutinis skirtumas tarp maksimumo taškų, kosinusų panašumas ir koreliacija, tikėtinumo santykio atstumas ir Lee – Verleysen atstumas (Minkowski'o normos apibendrinimas laiko eilutėms).

Skirtingų metodų derinių efektyvumas vertinant TBK patikrintas su generuotais duomenimis, kur vienam T bangos pavyzdžiui buvo keičiama amplitudė, T bangos trukmė, pridodamas atsitiktinis triukšmas bei su TWA duomenų bazės su pagal TBK lygį suranguotais simuliuotais bei tikrais EKG. Geriausiai tvarka pagal TBK lygį buvo išlaikyta TBK vertinant modifikuotu slenkančio vidurkio metodu, kuriame bangos prieš palyginimą išlyginamos atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo su asimetriniu judėjimo šablonu arba kryžminės koreliacijos metodu, o panašumas įvertinamas skaičiuojant absoliutinį skirtumą tarp maksimumo taškų.

MSVM su šiais pakeitimais patikrinti kaip galimi biomarkeriai diagnozuojant širdies sutrikimus pasinaudojant PTB duomenų baze. Sveikus ir sergančiuosius geriausiai skyrė metodas su atviros pradžios ir pabaigos dinaminio laiko skalės kraipymu su asimetriniu judėjimo šablonu ir panašumo vertinimu absoliutiniu skirtumu tarp maksimumo taškų. Šis metodas neskyrė tik sveikų ir sergančiųjų kardiopatija grupių. Šis trūkumas pašalintas suradus min-max kombinaciją tarp dviejų paskutinių TBK reikšmių.

TBK vertinimas gali būti veikiamas EKG apdorojimo prieš analizę, t.y. triukšmo šalinimo ir T bangų išskyrimo kokybės, todėl šios srities tobulinimas yra aktualus uždavinys. Pastebėta, kad laiko eilučių metodai, kurie leidžia T bangos pradžios ir pabaigos poslinkį, davė geresnius rezultatus, nei tie, kurie išlaiko pradžios ir pabaigos taškus fiksuotus. Be to, metodai, kurie išlygino bangas atsižvelgdami į jų formą (DSK metodai), pasirodė geriau už metodus, kurie keitė tik vienos bangos poslinkį kitos atžvilgiu.

Geriausius rezultatus davė tie panašumo vertinimui naudoti metodas, kuris atsižvelgė į

maksimumo taškus. Kosinusų panašumas ir koreliacija T bangų panašumo vertinime nedavė gerų rezultatų, taip gali būti dėl T bangų formos panašumo. Lee – Verleysen atstumas nerodo gerų rezultatų, tad galima daryti prielaidą, kad sveikiems ir sergantiesiems gaunami nepakankamai skirtingi TBK įverčiai. Šis atstumas sumažina atstumą tuo atveju, jei sekančiuose taškuose skirtumas keičia ženklą, t. y. gali būti nepakankamai įvertinama tokia TBK, kai amplitudės pasikeitimai yra ne iš eilės einančiuose taškuose, ar jei skiriasi tik maža, nors ir vientisa, bangų dalis. Pastebėtina, kad tam, kad būtų galima naudoti tikėtinumo santykio atstumą, turėtų būti kuo geriau pašalinamas triukšmas, nes būtent triukšmo komponentė gali paveikti atstumo skaičiavimą, jei triukšmo dažnis T bangoje bus vyraujantis.

TBK analizei buvo pasinaudota esamais, laisvai prieinamos programos R, paketais, taip pat parašytos funkcijos realizuojančios panaudotus laiko eilučių išlyginimo ir panašumo nustatymo metodus. Todėl galima tolesnė šių metodų tinkamumo kitiems laiko eilučių uždaviniams spęsti analizė.

Santrauka

Šiame darbe analizuotos įvairių laiko eilučių išlyginimo ir panašumo nustatymo metodų pritaikymo galimybės T bangų kaitos (TBK) analizėje, pagerinant Nearing ir Verrier pasiūlytą, modifikuoto slenkančio vidurkio metodą.

Pasinaudojant TWA duomenų baze ir generuotais duomenimis surasti labiausiai TBK analizei tinkami išlyginimo ir panašumo vertinimo metodai.

TBK, paskaičiuoto naudojant modifikuotą slenkančio vidurkio metodą papildytą TBK analizei tinkamiausiais laiko eilučių išlyginimo ir panašumo nustatymo metodais, tinkamumas širdies ligų diagnostikai patikrintas su duomenimis iš PTB duomenų bazės.

Pasiremiant PTB duomenimis rastas galimas biomarkeris širdies ligų diagnostikoje, paskutinių dviejų TBK įverčių, gautų taikant MSVM su atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo su asimetriniu judėjimo šablonu išlyginimą ir panašumą vertinant kaip absoliutinį skirtumą tarp maksimumo taškų, min-max kombinaciją.

Summary

T Wave Alternans Analysis Using a Modified Moving Average Method: a Comparison of Various Time Series Alignment and Similarity Detection Techniques

T wave alternans (TWA) is a beat-to-beat change in the amplitude or shape of T wave. TWA is one of potential biomarkers for ventricular arrhythmias and can be a sign of serious heart disease. Because there is no gold standard in TWA measuring, modifications of existing methods and new solutions are possible. Modified moving average method, proposed by Nearing and Verrier, is one of mostly used in medical practise, but can give misleading results then T waves is not properly aligned or T wave length and morphology changes because of heart rate variability. It is known, that some ventricular arrhythmias can cause heart rate variability, so this type of error is unwanted, because online TWA measuring can become one of sudden ventricular arrhythmias predictors in the near future.

In this work, various time series alignment and similarity detection techniques were used to improve TWA measuring and this measure capabilities in heart disease diagnostic were analyzed.

TWA analysis with simulated and real data from ECG databases was performed and potential biomarker was found by using biomarkers combining method, proposed by Liu, Liu and Halabi.

Literatūros sarakšas

1. D. Cuesta-Frau, M. Biaget, T-wave Alternans Analysis Improvement by Means of Curve Alignment Prior to Distance Calculation, *Engineering in Medicine and Biology Society, 2007.*, 29th Annual International Conference of the IEEE, 2007, p. 690 – 693.
2. D. Cuesta-Frau. et al. Enhanced modified moving average analysis of T-wave alternans using a curve matching method: a simulation study, *Medical & Biological Engineering & Computing*, 2009, 47, p.323-331.
3. L. Burattini, S. Bini, R. Burattini, Correlation method versus enhanced modified moving average method for automatic detection of T-wave alternans. *Computer Methods and Programs in Biomedicine* , 2010, 98, p. 94-102.
4. M. Alfaouri, K. Daqrouq, ECG Signal Denoising By Wavelet Transform Thresholding, *American Journal of Applied Sciences*, 2008, 5, p. 276-281.
5. P. S. Adison, *The illustrated wavelet transform handbook*, IOP Pub., 2002.
6. S. Z. Mahmoodabbadi, A. Ahmadian, M.D. Abolhasani, ECG Feature Extraction using Daubechies Wavelets, *Proceedings of the Fifth IASTED International Conference on VISUALIZATION, IMAGING, and IMAGE PROCEEDINGS, Sep 7-9, 2005 Bedirom, Spain.*
7. A. J. Bagnal, G. J. Janacek, M. Powell, A likelihood ratio distance measure for the similarity between the fourier transform of time series, *Advances in Knowledge Discovery and Data Mining*, 2005, 3518, p. 737-743.
8. J. A. Lee, M. Verleysen, Generalization of the Lp norm for time series and its application to self-organizing maps, *5th Workshop on Self-Organizing Maps, Paris (France)*, 2005, p. 733-740.
9. A. Diery, Novel Applications of the wavelet transform for analysis of P waves in clinical ECG recordings, PhD thesis, 2008, <http://www4.gu.edu.au:8080/adt-root/uploads/approved/adt-QGU20090910.114910/public/02Whole.pdf>.
10. Chunling Liu, Aiyi Liu, Susan Halabi, A min-max combination of biomarkers to improve diagnostic accuracy, *Statistics in Medicine*, 2011, 30 (27), p. 2005-2014.
11. T. Giorgino, Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package, *Journal of Statistical Software*, 2009, 31(7), p.1-24.
12. T. Giorgino, P. Tormente, dtw: Dynamic Time Warping Algorithms. R Package. <http://CRAN.R-project.org/package=dtw>, 2009
13. L. Rabiner, B.H. Juang, *Fundamentals of Speech Recognition*, Prentice-Hall, Upper

Saddle River, NJ, USA, 1993.

14. B. D. Nearing, R. L. Verrier, Modified moving average analysis of T-wave alternans to predict ventricular fibrillation with high accuracy, *J Appl Physiol*, 2002, 92, p.541-549.
15. D. M. Bloomfield, S. H. Hohnloser, R. J. Cohen, Interpretation and classification of microvolt T wave alternans test, *Journal of cardiovascular electrophysiology*, 2002, 13 (5), p.502-512.
16. B. Huang, W. Kinsner, ECG frame classification using dynamic time warping, *Proceedings of the 2002 IEEE Canadian Conference on Electrical & Computer Engineering*, 2002, 2, p.1105-1110.
17. R. Levulienė, *Statistikos taikymai naudojant SAS*, Vilnius:Vilniaus universiteto leidykla, 2009.
18. J. S. Rugienius, *Klinikinė elektrokardiologija*, Vilnius, 2004.
19. Z. O. Markienė, *Klinikinė elektrokardiografija*. Vilnius, 2000, www.ekg.lt.
20. W. Qiana, , M. D. Riedela, , I. Rosenbergb, Uniform approximation and Bernstein polynomials with coefficients in the unit interval, *European Journal of Combinatorics*, 2011, 32 , p.448–463.
21. E.V. Garcia, N. Samesima, P.H.G. Filho ir kt., Comparison of Quantitative T-Wave Alternans Profiles of Healthy Subjects and ICD Patients, *Ann Noninvasive Electrocardiol*, 2009, 14(2), p.108–118.
22. R. High, Dealing with 'Outliers': How to Maintain Your Data's Integrity, 2000, <http://rfd.uoregon.edu/files/rfd/StatisticalResources/outl.txt>.

Priedai

1. TBK analizei naudoti R paketai

DTW paketas [12] skirtas dinaminio laiko skalės kraipymo metodui

Pagrindinė funkcija *dtw*.

Naudojimas

```
DSKM<-dtw(pirmoji_laiko_eilutė, antroji_laiko_eilutė, step.pattern= judėjimo_šablonas)
```

step.pattern pavyzdžiai: *symmetric2* – įprastas simetrinis judėjimo šablonas, *asymmetric* – nesimetrinis judėjimo šablonas.

Rezultatas:

```
DSKM$rezultato_elementas
```

rezultato_elementas gali būti nurodomas toks:

query – pirmoji laiko eilutė;

reference – antroji laiko eilutė;

stepPattern – naudotas judėjimo šablonas;

N – pirmosios laiko eilutės ilgis;

M – antrosios laiko eilutės ilgis;

distance – minimalus sukauptasis atstumas (nenormuotas);

normalizeDistance – minimalus sukauptasis atstumas (normuotas);

index1 – pirmosios laiko eilutės kraipymo funkcijos reikšmės (indeksai);

index2 – antrosios laiko eilutės kraipymo funkcijos reikšmės (indeksai);

directionMatrix – trajektorijos matrica;

jmin – pabaigos taškas daliniams palyginimams.

Norint gauti išlyginimą imamos pirmosios laiko eilutės reikšmės indeksuose *index1* ir antrosios laiko eilutės reikšmės indeksuose *index2*.

Vaizdavimo būdai

```
plot(DSKM, type="twoway")
```

Vaizduojama lokalaus atstumo matrica ir sukaupytojo atstumo matricos, sujungti taškai rodo trajektoriją.

```
plot(DSKM, type="threeway")
```

Vaizduojama trajektorija, kairėje ir apačioje yra lyginamosios laiko eilutės.

waveslim

waveslim paketas naudojamas diskrečiai vilnelių transformacijai bei atvirkštinei diskrečiai vilnelių transformacijai atlikti.

dwt – atliekama diskrečioji vilnelių transformacija;

idwt – atliekama atvirkštinė vilnelių transformacija.

```
DBT<- dwt(laiko_eilutė, vilnelė, dekompozicijos_lygiai)
```

laiko_eilutė – laiko eilutė, kuriai bus atlikti diskrečioji vilnelių transformacija;

vilnelė – naudojama vilnelė, pvz. "d4", "d6";

dekompozicijos lygiai – į kiek komponentų bus skaidoma laiko eilutė.

Rezultatai

DBT\$d1 – 1-ojo lygio detalių koeficientai;

...

DBT\$d"n" – n-ojo lygio detalių koeficientai;

DBT\$s"n" - n-ojo lygio aproksimacijų koeficientai;

vietoje n reikia įrašyti n-ojo lygio numerį.

idwt(dwt_objektas) – bus atlikta atvirkštinė diskrečioji vilnelių transformacija.

2. Sukurtų funkcijų komentarai

Funkcijos skirtos EKG paruošimui T bangų analizei:

dreifas(ecg, iki) – ecg – EKG laiko eilutė, iki – kiek taškų analizuoti; rezultatas – EKG pašalinus izoelektrinės linijos dreifą;

triuksmas2(ecg, iki, kiek) – ecg – EKG laiko eilutė, iki – kiek taškų analizuoti; kiek –

keliuose lygiuose šalinti triukšmą; rezultatas – EKG pašalinus triukšmą;

aptikimasbangeliu5(ecgr, ecg2, iki, iki2, Rbanguslenkstis) – ecgr – EKG, kuriai reikia išskirti T bangas, ecg2 – EKG, kuria naudosime T bangų aptikimui (tuo atveju, jei ieškome T bangų vienoje derivacijoje, o analizuosime kitos derivacijos ar kitais atvejais), iki – parametras pagalbinėms funkcijoms, iki2 – kiek taškų analizuosime, Rbanguslenkstis – parametras skirtas apriboti galimai R bangų amplitudei; rezultatas – T bangų rinkinys;

Rezultatas T pateiktas tokiu formatu:

T[[1]] – 1-osios T bangos amplitudės reikšmės;

T[[2]] – 2-osios --- " ---;

ir taip toliau. Indeksas atitinka T bangos eilės numerį analizuojamoje EKG.

Pagalbinės funkcijos:

Riesko(ecgr, iki)

QSiesko2(ecgr, iki)

PTiesko(ecgr, iki)

nuliai(ecgr, iki)

nuliai2(ecgr, iki)

Šios funkcijos realizuoja 2.4.3. skyrelyje aprašytą filtravimą; ecgr – analizuojamas EKG, iki – kiek taškų bus naudojama.

Funkcijos realizuojančios laiko eilučių išlyginimo metodus:

warpingas2(bang1, bang2) – įprastas dinaminio skalės kraipymo metodas;

warpingas4(bang1, bang2) – atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo metodas su asimetrišku judėjimo šablonu;

warpingas5(bang1, bang2) – atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo metodas su simetrišku judėjimo šablonu;

kkoreliacija(bang1, bang2) – kryžminės koreliacijos metodas;

pagalmaks(bang1, bang2) – išlyginimas pagal maksimalius taškus;

Funkcijos, įrašius norimas išlyginti laiko eilutes, pateikia išlygintas laiko eilutes tokiu būdu LE[[1]] – pirmoji laiko eilutė; LE[[2]] – antroji laiko eilutė.

Funkcijos realizuojančios įvairius laiko eilučių panašumo nustatymo būdus:

kosinusai(bang1,bang2)

korpan(bang1,bang2)

dLa(bang1,bang2,kiek)

Lpatstumas2(bang1,bang2,tau,p)

Funkcijos realizuojančios MSVM keitimus:

funkcijos pavadinimo struktūra – EMMAMxxxxyyy, kur

xxx - išlyginimo_metodas;

yyy - panašumo_nustatymo_metodas;

pvz. EMMAMmodifl3 – modifikuotas slenkančio vidurkio metodas taikant atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo metodą su asimetrišku judėjimo šablonu ir maksimalių kiekvienos laiko eilutės reikšmių absoliutaus skirtumo metodą panašumui tarp laiko eilučių vertinimui.

Išlyginimo metodai: nieko neįrašoma – įprastas dinaminio skalės kraipymo metodas; modifl - atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo metodas su asimetrišku judėjimo šablonu, modif2 - atviros pradžios ir pabaigos dinaminio laiko skalės kraipymo metodą su simetrišku judėjimo šablonu; kkoreliacija – išlyginimas naudojant kryžminės koreliacijos metodą; pagalmaks – išlyginimas atsižvelgiant į maksimumo taškus.

Panašumo nustatymo metodai: 2 – didžiausias absoliutinis skirtumas tarp atitinkamų laiko eilučių taškų reikšmių; 3 - maksimalių kiekvienos laiko eilutės reikšmių absoliutinis skirtumas; dL – tikėtino santykio atstumas; Lpdtw2 – Lp normos apibendrinimas laiko eilutėms, kos – kosinusų panašumas; kor – koreliacija.

Funkcijoms reikia įrašyti T bangų komplekto pavadinimą. Pvz.

EMMAMmodifl3(bangos)

bangos – funkcijos *aptikimasbangeliu5* rezultatas ar jo dalis;

Rezultatas šiuo konkrečiu atveju $\Lambda_k^{(2)}$, $k=1, \dots, \frac{N}{2}$, kur N – T bangų skaičius rinkinyje.

Kitų funkcijos veikia analogiškai, tačiau yra tokios išimtys.

Funkcijos, kuriose naudojamas dL atstumas vykdomos taip:

EMMAMxxxdL(bangos, kiek)

bangos – analogiškai kitiems atvejams;
kiek – kiek didžiausių koeficientų reikšmių imtį;

Funkcijos, kuriose naudojamas Lp atstumas vykdomos taip:

$EMMAM_{xxx}Lp_{dtw2}(bangos, \tau, p)$

bangos – analogiškai kitiems atvejams;
 τ ir p yra (41) ir (42) funkcijų parametrai;

biomarkerių apjungimas min-max kombinacijai vykdomas optimizuojant funkciją
 $minmaxkomb(\lambda, X1, X2, Y1, Y2)$ – $X1, X2$ – biomarkerių įgyjamos reikšmės sveikiems žmonėms, $Y1, Y2$ – biomarkerių įgyjamos reikšmės sergantiesiems.

Pastaba. Optimizavimas atliekamas su pso paketu.

```
opt<-psoptim(rep(NA,1), fn=minmaxkomb,X1=X1,X2=X2,Y1=Y1,Y2=Y2)
```

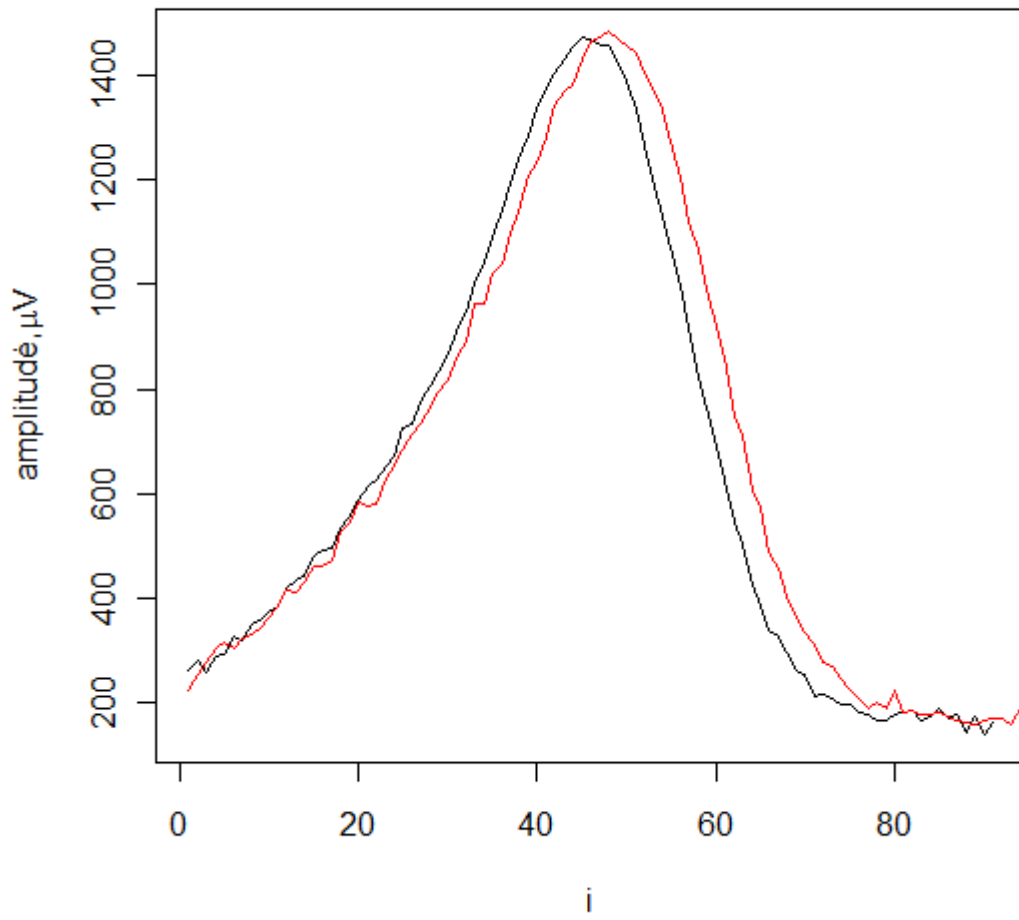
$-(opt\$value)$ – įgyjamas maksimumas;

$opt\$par$ – parametro įvertis;

2. Laiko eilučių išlyginimo pavyzdžiai

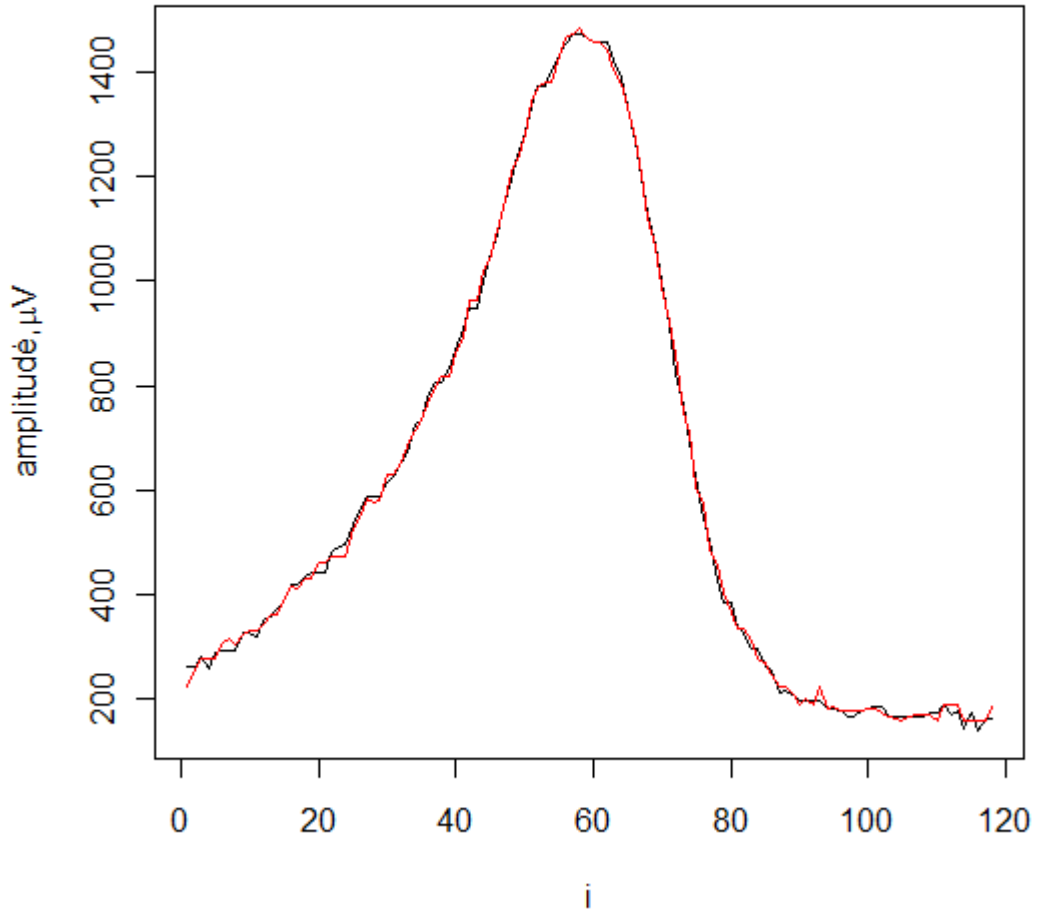
Išlygiavimo pavyzdžiai: 1 pav. - netaikius išlyginimo metodų; 2 pav. - taikant įprastą dinaminio laiko skalės kraipymo metodą; 3 pav. - taikant atviros pradžios-pabaigos dinaminio laiko skalės kraipymo metodą su asimetriniu judėjimo šablonu; 4 pav. - taikant atviros pradžios-pabaigos dinaminio laiko skalės kraipymo metodą su asimetriniu judėjimo šablonu; 5 pav. - taikant kryžminės koreliacijos metodą; 6 pav. - taikant išlyginimą pagal maksimumo taškus.

Išlyginimo pavyzdys



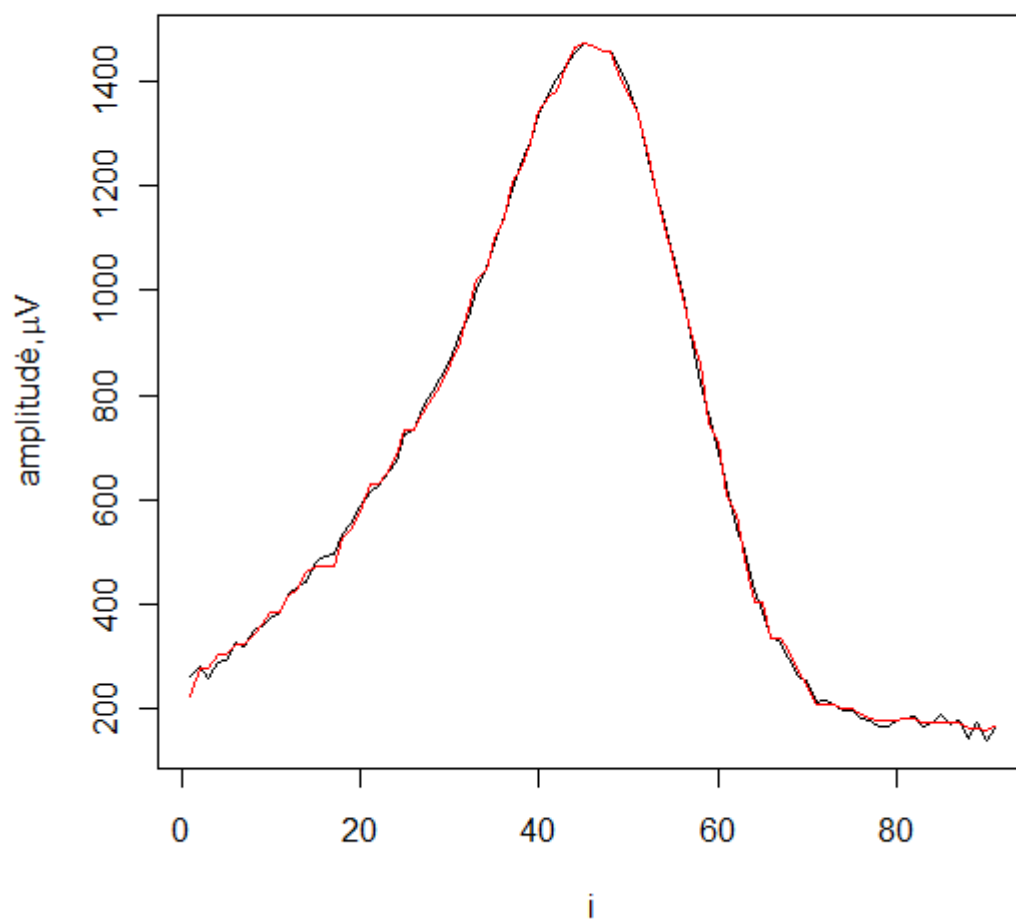
1 pav. netaikius išlyginimo

Išlyginimo pavyzdys 1



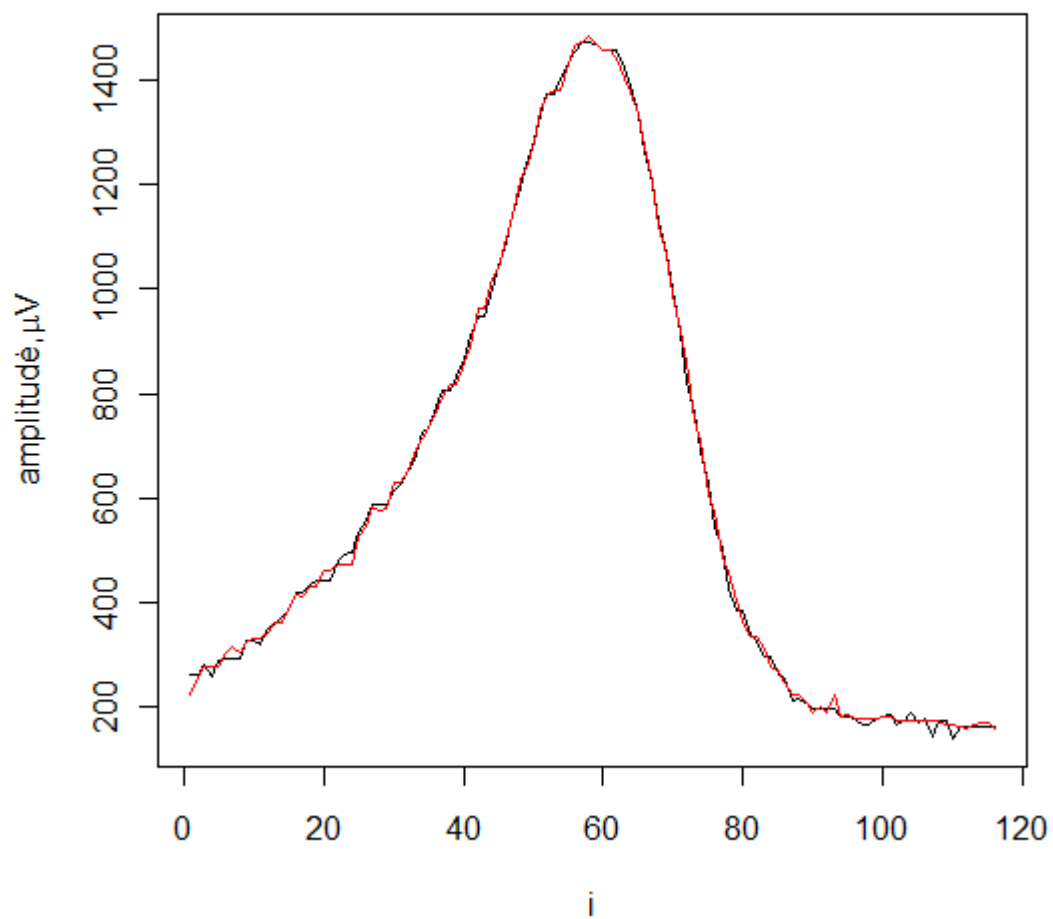
2 pav. išlyginta taikant įprastą DLSK

Išlyginimo pavyzdys 2



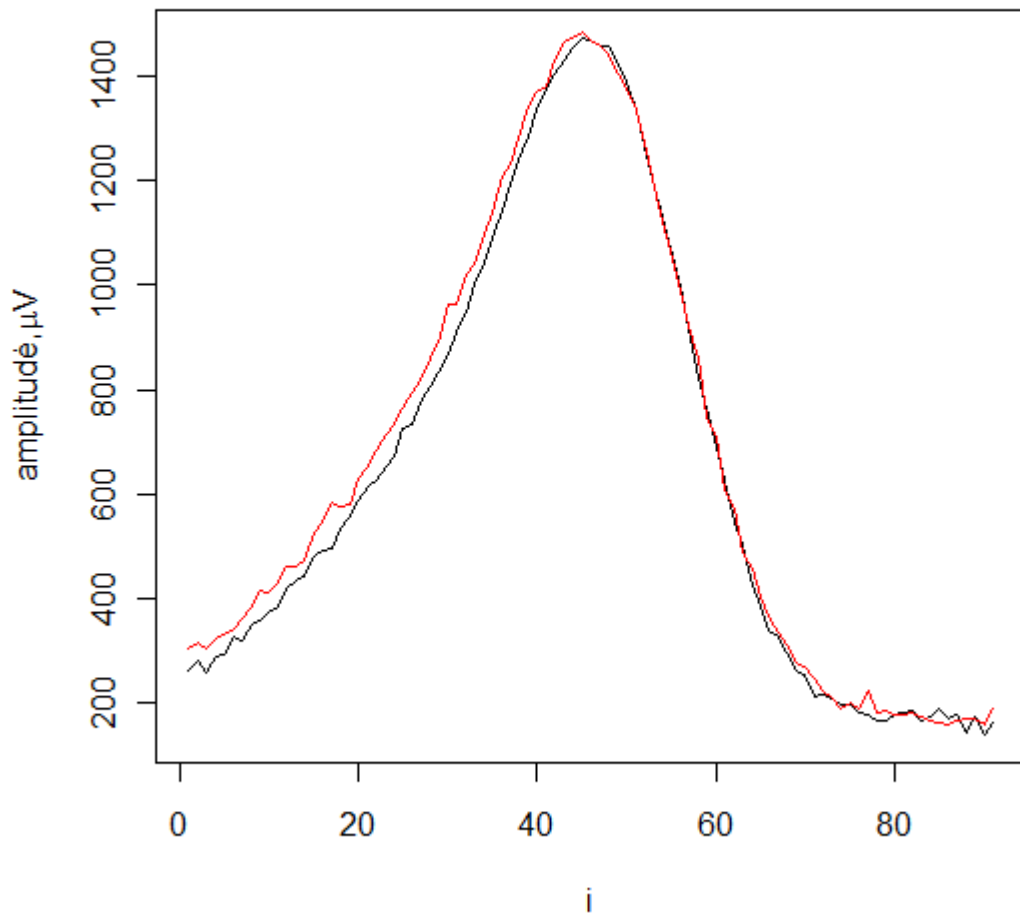
3 pav. išlyginta taikant APP-DLSK su simetriniu judėjimo šablonu

Išlyginimo pavyzdys 3



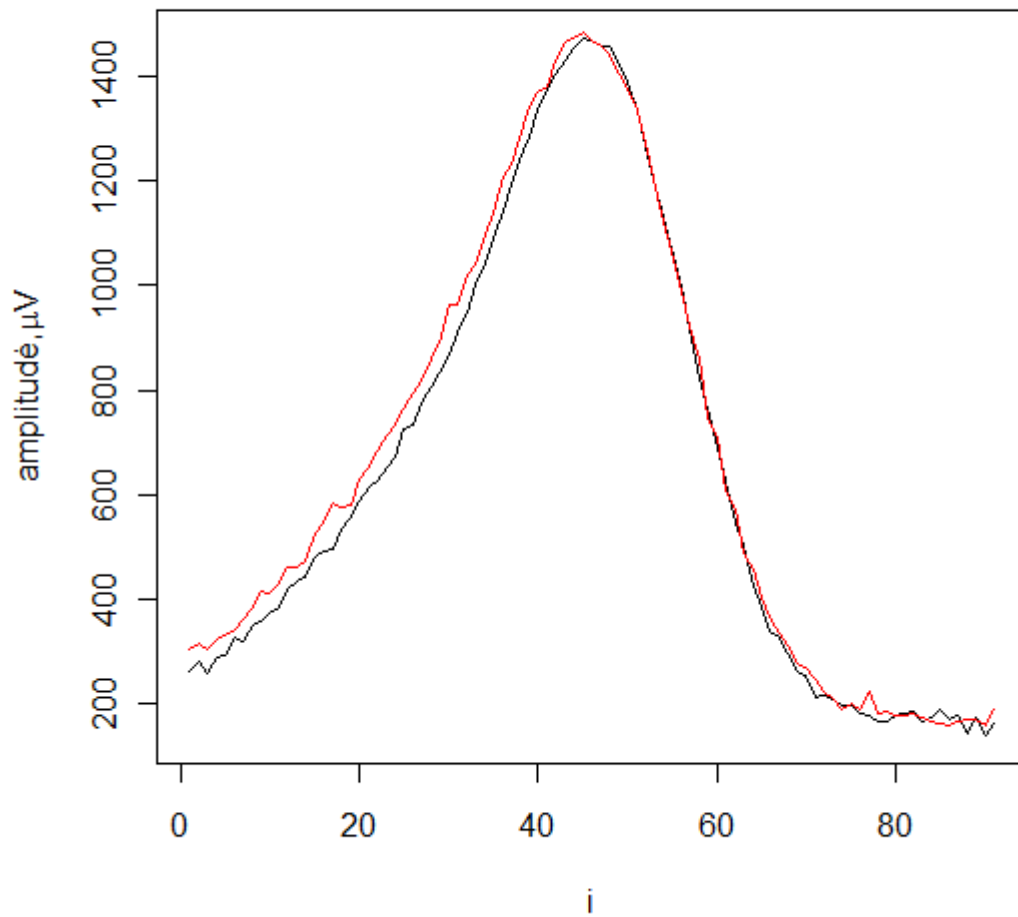
4 pav. išlyginta taikant APP-DLSK su asimetriniu judėjimo šablonu

Išlyginimo pavyzdys 4



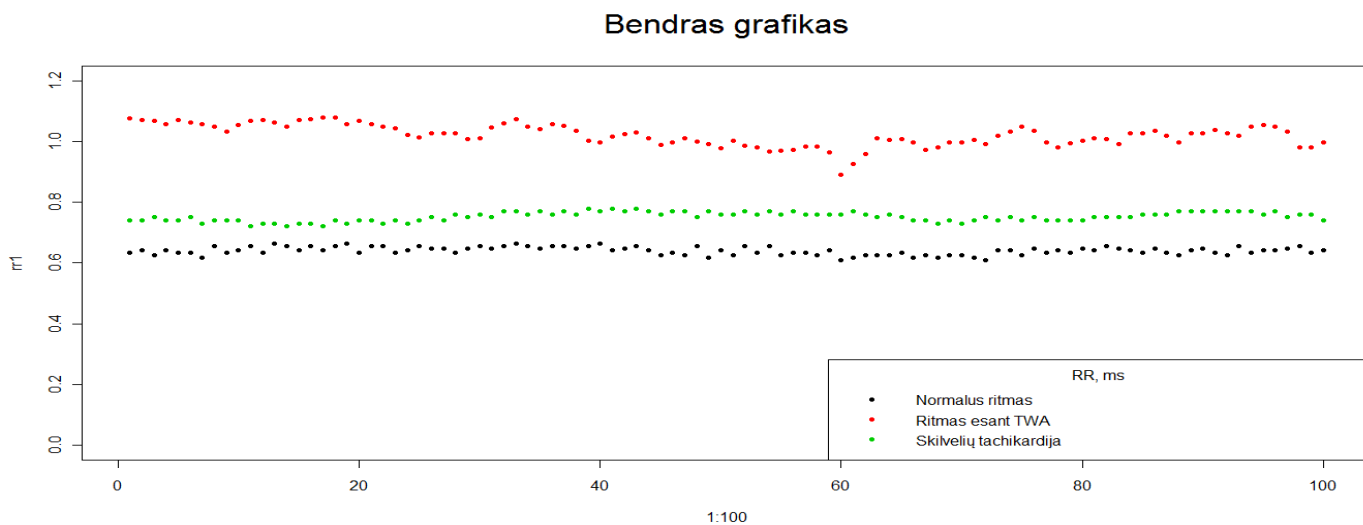
5 pav. išlyginta taikant kryžminės koreliacijos metodą

Išlyginimo pavyzdys 5



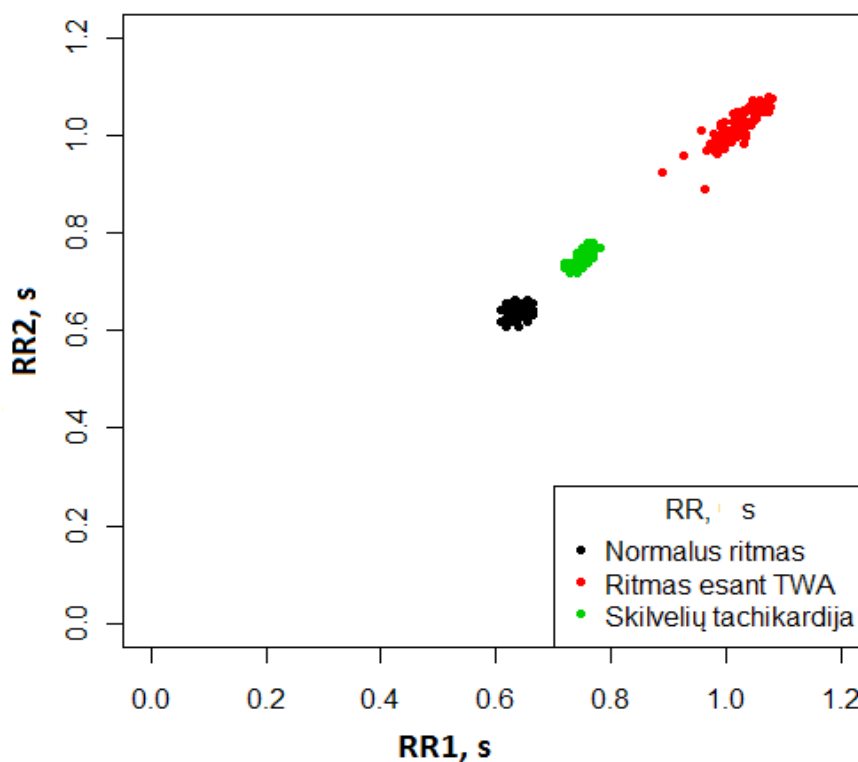
6 pav. išlyginta taikant išlyginimą pagal maksimumo taškus

4. Ritmo pavyzdžiai



7 pav. Kaip keičiasi RR intervalo trukmė (s) esant skirtingoms situacijoms, juoda – normalaus ritmo atvejis, raudona – esant TBK, žalia – skilvelių tachikardijos atvejis.

Bendras grafikas



8 pav. RR intervalo trukmės (s) Puankarė grafikas, kuriame vaizduojama RR2 – esamasis RR

intervalas, RR1 – ankstesnis RR intervalas

5. Kendalo koreliacijos generuotiems duomenims ir TWA duomenų bazės duomenims

Pateikiama koreliacija generuotiems duomenims tarp TBK lygio ir apskaičiuotų 16 TBK kiekvienam rinkiniui vidurkio, maksimalaus TBK, minimalaus TBK, medianos, TBK 16-osios reikšmės.

Pirmoji eilutė koreliacija, antroji p-reikšmė, įrašas 0 reiškia, kad p-reikšmė <0.0001.

metodas	vid	maks	min	mediana	TWA16
metodas11	0.4612	0.3886	0.5116	0.4698	0.5018
	0	0	0	0	0
metodas12	0.6898	0.6294	0.7145	0.6927	0.7123
	0	0	0	0	0
metodas13	-0.422	-0.4906	-0.3247	-0.4425	-0.5016
	0	0	0	0	0
metodas14	-0.4051	-0.4706	-0.3078	-0.4251	-0.4814
	0	0	0	0	0
metodas15	0.5708	0.5351	0.5879	0.5724	0.5492
	0	0	0	0	0
metodas16	0.5775	0.5368	0.5747	0.5664	0.5445
	0	0	0	0	0
metodas17	0.446	0.4622	0.4241	0.4438	0.465
	0	0	0	0	0
metodas21	0.4948	0.4338	0.5255	0.4954	0.5106
	0	0	0	0	0
metodas22	0.696	0.6364	0.7126	0.7015	0.7245
	0	0	0	0	0
metodas23	-0.3943	-0.4351	-0.3476	-0.4046	-0.4448
	0	0	0	0	0
metodas24	-0.3936	-0.4286	-0.351	-0.4032	-0.4365
	0	0	0	0	0
metodas25	0.5666	0.5266	0.5871	0.5564	0.5556
	0	0	0	0	0

metodas26	0.5488	0.5051	0.568	0.5449	0.548
	0	0	0	0	0
metodas27	0.5038	0.4683	0.5202	0.507	0.5258
	0	0	0	0	0
metodas31	0.4609	0.3908	0.5066	0.4666	0.4972
	0	0	0	0	0
metodas32	0.6927	0.6339	0.7148	0.6945	0.7189
	0	0	0	0	0
metodas33	-0.4253	-0.4934	-0.3293	-0.4482	-0.5033
	0	0	0	0	0
metodas34	-0.4043	-0.4666	-0.3112	-0.4249	-0.4755
	0	0	0	0	0
metodas35	0.5689	0.5369	0.5901	0.5715	0.5617
	0	0	0	0	0
metodas36	0.5743	0.5567	0.5782	0.5509	0.5456
	0	0	0	0	0
metodas37	0.4514	0.4618	0.4276	0.4489	0.4647
	0	0	0	0	0
metodas41	0.4462	0.3368	0.546	0.4578	0.5352
	0	0	0	0	0
metodas42	0.7203	0.6478	0.7477	0.7277	0.7786
	0	0	0	0	0
metodas43	-0.1747	-0.2904	-0.1068	-0.2043	-0.2893
	0	0	0.0067	0	0
metodas44	-0.1735	-0.2909	-0.1083	-0.1991	-0.2816
	0	0	0.0059	0	0
metodas45	0.0854	0.0532	0.2083	0.1078	0.1681
	0.0296	0.1754	0	0.006	0
metodas46	0.0321	0.0134	0.0795	0.0496	0.0798
	0.4135	0.7338	0.0428	0.2066	0.0419
metodas47	0.5384	0.4376	0.632	0.548	0.6221
	0	0	0	0	0
metodas51	0.4843	0.3692	0.5778	0.4993	0.5681
	0	0	0	0	0

metodas52	0.7157	0.6279	0.7698	0.7246	0.7344
	0	0	0	0	0
metodas53	-0.2279	-0.3387	-0.1473	-0.2588	-0.3373
	0	0	2e-04	0	0
metodas54	-0.2294	-0.3333	-0.1502	-0.2576	-0.3282
	0	0	1e-04	0	0
metodas55	0.0873	0.0377	0.2191	0.106	0.1967
	0.0261	0.3367	0	0.0069	0
metodas56	0.049	0.0205	0.0845	0.0581	0.1238
	0.2115	0.6017	0.0314	0.1388	0.0016
metodas57	0.5786	0.4786	0.6644	0.5947	0.6557
	0	0	0	0	0

Pateikiama koreliacija tarp TWA duomenų bazėje apibrėžto rango ir apskaičiuotų 16-os TBK kiekvienam EKG, maksimalaus TBK, minimalaus TBK, medianos, TBK 16-osios reikšmės.

TWA sintetiniai duomenys

	vid	maks	min	mediana	TWA 16
metodas11	0.2339	0.1734	0.2298	0.2339	0.2621
	0.0622	0.1681	0.0669	0.0622	0.0364
metodas12	0.6734	0.6288	0.6008	0.6815	0.5524
	0	0	0	0	0
metodas13	-0.1331	-0.2661	-0.0645	-0.1411	-0.0605
	0.2919	0.0336	0.6152	0.2632	0.6382
metodas14	-0.1411	-0.2621	-0.0766	-0.1573	-0.0645
	0.2632	0.0364	0.5485	0.2118	0.6152
metodas15	0.1855	0.2056	0.1694	0.2177	0.0685
	0.14	0.1015	0.1783	0.0827	0.5926
metodas16	0.1008	0.0363	0.2056	0.3226	0.1734
	0.4268	0.7828	0.1015	0.0099	0.1681
metodas17	0.1653	0.1169	0.1774	0.1855	0.121
	0.189	0.3553	0.1583	0.14	0.3387
metodas21	0.1653	0.121	0.2707	0.1613	0.2339
	0.189	0.3387	0.031	0.2002	0.0622
metodas22	0.6573	0.6343	0.6113	0.6694	0.5847
	0	0	0	0	0
metodas23	-0.1371	-0.2419	-0.125	-0.1573	-0.121
	0.2773	0.0536	0.3226	0.2118	0.3387

metodas24	-0.1492	-0.2581	-0.129	-0.1411	-0.1129
	0.2365	0.0394	0.307	0.2632	0.3724
metodas25	0.1411	0.2056	0.1452	0.1452	0.1169
	0.2632	0.1015	0.2496	0.2496	0.3553
metodas26	0.2944	0.1613	0.246	0.2984	0.1008
	0.0187	0.2002	0.0497	0.0171	0.4268
metodas27	0.0847	0.1089	0.2339	0.0806	0.1653
	0.5061	0.3901	0.0622	0.5271	0.189
metodas31	0.2339	0.1734	0.2298	0.2339	0.2621
	0.0622	0.1681	0.0669	0.0622	0.0364
metodas32	0.6734	0.6288	0.6008	0.6815	0.5524
	0	0	0	0	0
metodas33	-0.1371	-0.2863	-0.0645	-0.1492	-0.0766
	0.2773	0.0222	0.6152	0.2365	0.5485
metodas34	-0.1452	-0.254	-0.0766	-0.1573	-0.0484
	0.2496	0.0427	0.5485	0.2118	0.7092
metodas35	0.1734	0.2056	0.1452	0.1734	0.0766
	0.1681	0.1015	0.2496	0.1681	0.5485
metodas36	0.1008	0.0363	0.2097	0.2984	0.1331
	0.4268	0.7828	0.0949	0.0171	0.2919
metodas37	0.1734	0.125	0.1815	0.1895	0.1331
	0.1681	0.3226	0.1489	0.1315	0.2919
matodas41	0.4234	0.2759	0.4839	0.4516	0.3427
	7e-04	0.0284	1e-04	3e-04	0.0061
metodas42	0.5927	0.541	0.4919	0.6129	0.4677
	0	0	1e-04	0	2e-04
metodas43	-0.2097	-0.2159	-0.1573	-0.2563	-0.1169
	0.0949	0.0856	0.2118	0.041	0.3553
metodas44	-0.2379	-0.2159	-0.1653	-0.2621	-0.1089
	0.0578	0.0856	0.189	0.0364	0.3901
metodas45	0.1935	0.1371	0.4476	0.2339	0.2823
	0.1234	0.2773	3e-04	0.0622	0.0242
metodas46	0.2339	0.1935	0.2218	0.2097	0.1169
	0.0622	0.1234	0.0771	0.0949	0.3553
metodas47	0.4153	0.3105	0.4677	0.4476	0.3992
	9e-04	0.0131	2e-04	3e-04	0.0014
metodas51	-0.0484	-0.0121	0.1756	-0.0605	0.0645
	0.7092	0.9354	0.1631	0.6382	0.6152
metodas52	0.2137	0.1371	0.6016	0.2258	0.0524
	0.0886	0.2773	0	0.0719	0.6852

metodas53	0.0484	-0.004	0	0.0968	-0.0121
	0.7092	0.9871	1	0.446	0.9354
metodas54	0.0887	0.0081	-0.0403	0.1573	-0.0202
	0.4856	0.9612	0.758	0.2118	0.884
metodas55	-0.129	-0.1169	0.1169	-0.1452	0.0927
	0.307	0.3553	0.3553	0.2496	0.4655
metodas56	-0.1331	-0.0968	0.0887	-0.1331	0.1815
	0.2919	0.446	0.4856	0.2919	0.1489
metodas57	-0.0685	-0.0363	0.1694	-0.0685	0.0524
	0.5926	0.7828	0.1783	0.5926	0.6852

TWA tikri duomenys

	vid	maks	min	mediana	TWA 16
metodas11	0.165	0.1296	0.1694	0.1274	0.1229
	0.1214	0.2248	0.1117	0.2328	0.2497
metodas12	0.1717	0.2337	0.0897	0.1074	0.0476
	0.107	0.028	0.4025	0.3151	0.6603
metodas13	0.165	0.1362	0.1672	0.165	0.1429
	0.1214	0.2017	0.1165	0.1214	0.1804
metodas14	0.1229	0.0698	0.1229	0.103	0.0808
	0.2497	0.5164	0.2497	0.3356	0.4511
metodas15	-0.1229	-0.0986	-0.1141	-0.1451	-0.134
	0.2497	0.3571	0.2858	0.1737	0.2092
metodas16	-0.1406	-0.0986	-0.165	-0.1894	-0.1096
	0.1873	0.3571	0.1214	0.0752	0.3051
metodas17	0.0543	0.0388	0.1362	0.0698	0.072
	0.6154	0.722	0.2017	0.5164	0.503
metodas21	0.1694	0.1894	0.1561	0.1805	0.0986
	0.1117	0.0752	0.1429	0.09	0.3571
metodas22	0.1141	0.1473	0.0365	0.0919	0.0986
	0.2858	0.1671	0.7377	0.3908	0.3571
metodas23	0.1783	0.1008	0.1783	0.1384	0.1606
	0.094	0.3463	0.094	0.1944	0.1318
metodas24	0.1052	0.0609	0.1451	0.0963	0.1229
	0.3252	0.572	0.1737	0.3681	0.2497
metodas25	-0.1185	-0.1096	-0.1584	-0.1185	-0.103
	0.2673	0.3051	0.1373	0.2673	0.3356
metodas26	-0.1118	-0.1141	-0.1207	-0.0897	-0.1207
	0.2953	0.2858	0.2584	0.4025	0.2584

metodas27	0.0897	0.0609	0.1008	0.0698	0.0233
	0.4025	0.572	0.3463	0.5164	0.8342
metodas31	0.1473	0.1761	0.0498	0.0786	-0.0055
	0.1671	0.0982	0.6452	0.4638	0.9666
metodas32	0.0808	0.1761	0.0742	0.0565	0.0476
	0.4511	0.0982	0.4897	0.6008	0.6603
metodas33	0.247	0.2337	0.2137	0.2757	0.309
	0.0202	0.028	0.0445	0.0094	0.0036
metodas34	0.1495	0.1451	0.1296	0.1783	0.2159
	0.1608	0.1737	0.2248	0.094	0.0423
metodas35	-0.1406	-0.1384	-0.1539	-0.1783	-0.1761
	0.1873	0.1944	0.1487	0.094	0.0982
metodas36	-0.165	-0.1096	-0.2403	-0.1672	-0.2137
	0.1214	0.3051	0.0238	0.1165	0.0445
metodas37	0.0144	0.0233	0.0498	0.0144	-0.0676
	0.9001	0.8342	0.6452	0.9001	0.5301
metodas41	0.165	0.0853	0.2447	0.2381	0.1849
	0.1214	0.4264	0.0213	0.0251	0.0823
metodas42	0.2625	0.2691	0.1805	0.2713	0.0831
	0.0135	0.0113	0.09	0.0107	0.4387
metodas43	0.0388	-0.0122	0.0055	-0.021	-0.0277
	0.722	0.9167	0.9666	0.8506	0.8017
metodas44	-0.1872	-0.1805	-0.1783	-0.1495	-0.1827
	0.0787	0.09	0.094	0.1608	0.0861
metodas45	0.1141	0.0653	0.1672	0.1473	0.1628
	0.2858	0.5439	0.1165	0.1671	0.1265
metodas46	0.0498	-0.0078	0.0963	0.0853	0.1827
	0.6452	0.9499	0.3681	0.4264	0.0861
metodas47	0.1717	0.1761	0.1274	0.1982	0.0831
	0.107	0.0982	0.2328	0.0625	0.4387
metodas51	0.247	0.2425	0.2359	0.2625	0.165
	0.0202	0.0225	0.0265	0.0135	0.1214
metodas52	0.2182	0.2492	0.0986	0.1739	0.2204
	0.0402	0.0191	0.3571	0.1026	0.0383
metodas53	-0.1096	-0.0565	-0.1384	-0.0831	-0.0764
	0.3051	0.6008	0.1944	0.4387	0.4767
metodas54	-0.1672	-0.0831	-0.2403	-0.1141	-0.1694
	0.1165	0.4387	0.0238	0.2858	0.1117
metodas55	0.1739	0.1827	0.1561	0.1694	0.0986
	0.1026	0.0861	0.1429	0.1117	0.3571

metodas56	0.1185	0.1362	0.0853	0.1717	0.1894
	0.2673	0.2017	0.4264	0.107	0.0752
metodas57	0.2381	0.2447	0.1916	0.2492	0.103
	0.0251	0.0213	0.0719	0.0191	0.3356

TWA bendri duomenys

	vid	maks	min	mediana	TWA 16
metodas11	-0.1769	-0.1942	-0.1135	-0.1877	-0.107
	0.025	0.0139	0.1509	0.0174	0.1757
metodas12	-0.0054	-0.0747	0.2577	-0.0133	0.0609
	0.9489	0.346	0.0011	0.8692	0.4422
metodas13	0.285	0.2144	0.2966	0.2714	0.2238
	3e-04	0.0066	2e-04	6e-04	0.0046
metodas14	0.2238	0.138	0.2425	0.2216	0.1532
	0.0046	0.0806	0.0021	0.005	0.0524
metodas15	-0.2065	-0.1755	-0.1856	-0.2404	-0.2
	0.0089	0.0262	0.0187	0.0023	0.0113
metodas16	-0.0724	0.0544	-0.1596	-0.195	-0.1611
	0.3603	0.4926	0.0432	0.0135	0.0413
metodas17	-0.2368	-0.2353	-0.2101	-0.2281	-0.2043
	0.0027	0.0029	0.0078	0.0038	0.0096
metodas21	-0.1575	-0.1481	-0.12	-0.1517	-0.0926
	0.0461	0.0607	0.1288	0.0547	0.2416
metodas22	0.0241	-0.0263	0.1078	0.0299	0.0645
	0.7627	0.7419	0.1728	0.7076	0.4155
metodas23	0.28	0.2447	0.2663	0.2728	0.2764
	4e-04	0.0019	7e-04	5e-04	5e-04
metodas24	0.2303	0.1791	0.2332	0.2382	0.2151
	0.0035	0.0233	0.0031	0.0025	0.0064
metodas25	-0.0407	0.0011	-0.1193	-0.0616	0.0054
	0.6084	0.9927	0.1312	0.4368	0.9489
metodas26	0.355	0.3088	0.3204	0.36	0.2973
	0	1e-04	0	0	2e-04
metodas27	-0.1964	-0.187	-0.156	-0.2029	-0.1582
	0.0128	0.0178	0.0481	0.0101	0.0451
metodas31	-0.1777	-0.1647	-0.1733	-0.2022	-0.1676
	0.0244	0.037	0.0281	0.0104	0.0338
metodas32	-0.0407	-0.0963	0.164	-0.0378	0.0371
	0.6084	0.2237	0.0378	0.6343	0.6408

metodas33	0.2807	0.239	0.2822	0.2829	0.2627
	4e-04	0.0025	3e-04	3e-04	9e-04
metodas34	0.2065	0.182	0.2216	0.2115	0.195
	0.0089	0.0211	0.005	0.0074	0.0135
metodas35	-0.1755	-0.1308	-0.1841	-0.2216	-0.1748
	0.0262	0.0977	0.0197	0.005	0.0268
metodas36	0.0241	0.1604	-0.1777	-0.1049	-0.1741
	0.7627	0.0423	0.0244	0.1847	0.0275
metodas37	-0.2159	-0.1921	-0.2382	-0.2166	-0.2505
	0.0062	0.015	0.0025	0.0061	0.0015
metodas41	-0.2029	-0.2594	-0.1128	-0.1726	-0.1611
	0.0101	0.001	0.1535	0.0288	0.0413
metodas42	-0.0227	-0.1017	0.2	-0.0112	-0.0277
	0.7767	0.1986	0.0113	0.8908	0.7281
metodas43	0.3305	0.315	0.3225	0.3049	0.3247
	0	1e-04	0	1e-04	0
metodas44	0.2425	0.2573	0.2418	0.262	0.2555
	0.0021	0.0011	0.0022	9e-04	0.0012
metodas45	-0.267	-0.2822	-0.1647	-0.2562	-0.2108
	7e-04	3e-04	0.037	0.0012	0.0076
metodas46	-0.231	-0.1813	-0.2238	-0.2613	-0.2231
	0.0034	0.0217	0.0046	9e-04	0.0047
metodas47	-0.2166	-0.2396	-0.1596	-0.1935	-0.187
	0.0061	0.0024	0.0432	0.0142	0.0178
metodas51	0.1524	0.2144	-0.0151	0.1423	0.1993
	0.0536	0.0066	0.8512	0.0715	0.0116
metodas52	0.2238	0.195	0.2551	0.1863	0.2346
	0.0046	0.0135	0.0012	0.0183	0.0029
metodas53	-0.0703	0.2425	-0.1243	-0.0032	-0.1135
	0.3749	0.0021	0.1156	0.9708	0.1509
metodas54	-0.2447	0.1315	-0.3355	-0.1344	-0.236
	0.0019	0.0959	0	0.0888	0.0028
metodas55	0.2677	0.244	0.2238	0.2605	0.311
	7e-04	0.002	0.0046	0.001	1e-04
metodas56	0.285	0.2858	0.231	0.3023	0.391
	3e-04	3e-04	0.0034	1e-04	0
metodas57	0.1366	0.1834	-0.0105	0.115	0.1784
	0.0838	0.0201	0.8981	0.1458	0.0238

