

Klasterizavimo algoritmai dideles apimties medicinos duomenims

Ataskaita už IV-ąjį doktorantūros kursą
(2024/2025 metų I pusmetis)

Doktorantas: Roma Purnaitė

Vadovas: prof. dr. Audronė Jakaitienė

Doktorantūros pradžios ir pabaigos metai: 2017-2025

2025 m. kovas

Tyrimo objektas, tikslas ir uždaviniai

- Tyrimo objektas
 - Didelės apimties medicinos duomenys, administracinio pobūdžio sveikatos duomenų rinkiniai
 - Klasterizavimo algoritmai
- Tyrimo tikslas
 - Pasiūlyti metodą didelės apimties medicinos duomenims klasterizuoti, atsižvelgiant į duomenų dinamikos laike savybes.
- Uždaviniai
 - Ištirti dideles apimties medicinos duomenų klasterizavimui dažniausiai taikomus klasterizavimo metodus.
 - Pasiūlyti klasterizavimo algoritmą ar esamo metodo patobulinimą, kuris atsižvelgtų į dinamiką laike.
 - Pritaikyti atrinktus algoritmus ir pasiūlytą sprendimą realiems medicinos duomenų rinkiniams.
 - Pasiūlyti algoritmo integravimo į sveikatos priežiūros įstaigos informacinę sistemą modelį.

Visų doktorantūros studijų planas

1 lentelė Visų studijų planas.

Studijų metai	Egzaminai	
	Planas	Ivykdyta
I (2017/2018)	2	2
II (2018/2019)	2	0
III (2021/2022)	2 (skola iš II metų)	2 (skola iš II metų)
IV (2024/2025*)	0	0
Iš viso:	4	4

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citav. rodikliu			Be citav. rodikliu		
	Planas	Ivykdyta	Planas	Ivykdyta	Planas	Ivykdyta	Būklė	Planas	Ivykdyta	Būklė
I (2017/2018)		(1)				1	publikuota		1	publikuota
II (2018/2019)		1 + (3)					publikuota		1	publikuota
III (2021/2022)	1	2			1	1	publikuota		1	publikuota
IV (2024/2025*)	1	2** + (1)		1	1	1	publikuota			
Iš viso:		5 + (5)		1	2	3			3	

*- Po akademinių atostogų. (2019 10 16–2021 09 30; 2023 03 13–2024 02 13 akademinėse atostogose)

** - viena įvyko, į kitą priimta (konferencija vyks 2025 metų gegužę). Skliaustuose – susiję su disertacija pranešimai, kurias pristatė bendraautoriai.

Ataskaita už IV mokslo metus, I pusmetį

- Egzaminai

Egzaminai		
Planas	Įvykdyta	Būklė
Nenumatyta	-	-

- Publikacijos

Publikacijos			
Planas	Įvykdyta	Būklė	Publikacijos tipas
1 publikacija (žurnale su Web of Science IF)	Ramanauskaitė, Dovilė; Puronaitė, Roma ; Jakaitienė, Audronė; Glaveckaitė, Sigita. Prevalence of multimorbidity in Lithuania: insights from national health insurance fund data // Journal of cardiovascular development and disease. Basel : MDPI. eISSN 2308-3425. 2025, vol. 12, iss. 2, art. no. 47, p. [1-18]. DOI: 10.3390/jcdd12020047 .	publikuota	Web of Science IF 2.4

Ataskaita už IV mokslo metus, I pusmetį

- Konferencijos

Dalyvavimas konferencijose		
2023 <i>ISCB44: 44th annual conference of the International Society for Clinical Biostatistics; Joint conference with the Italian Region of the International Biometric Society,</i> 2023 m. rugpjūčio mėn. 27-31 d. Milanai, Italija (nuotoliu)	Puronaitė, Roma ; Ramanauskaitė, Dovilė; Švaikevičienė, Kristina; Tarutytė, Gabrielė; Trinkūnas, Justas; Burneikaitė, Greta; Kazėnaitė, Edita; Glaveckaitė, Sigita; Jakaitienė, Audronė. Using nonlinear models to identify breakpoints in disease prevalence among patients with multimorbidity // ISCB44: 44th annual conference of the International Society for Clinical Biostatistics; Joint conference with the Italian Region of the International Biometric Society, Milan, Italy, 27-31 August 2023 : final programme & abstract book. Milano : ISCB. 2023, abstract no. MP117, p. 306. Prieiga per internetą: https://iscb.international/wp-content/uploads/2023/08/Book-of-Abstract-ISCB44_25-AGO-2.pdf . Pranešėja: Roma Puronaitė	Tarptautinė
2025 <i>29th Nordic-Baltic Congress of Cardiology</i> 2025 2025 m. gegužės mėn. 22-24 d. Kopenhaga, Danija (gyvai)	2025 gegužės mėn. NBCC konferencija (priimta) Kilaitė, Lorian; Puronaitė, Roma ; Jankauskienė, Augustina; Burneikaitė, Greta; Jakaitienė, Audronė Age-related breakpoints in the prevalence of multimorbidity and common comorbidities: a population-based study in Lithuania Pranešėja: Roma Puronaitė	Tarptautinė

Ataskaita už IV mokslo metus, I pusmetį

- Konferencijos

	Kiti susiję pranešimai	
2023 <i>Dirbtinio intelekto technologijos medicinoje: tyrimai ir diagnostika</i> 2023 m. spalio 20 d. Vilnius, Lietuva (gyvai)	Šavareikaitė, Alicija; Puronaitė, Roma ; Burneikaitė, Greta; Švaikevičienė, Kristina; Ramanauskaitė, Dovilė; Trinkūnas, Justas; Kazėnaitė, Edita; Jakaitienė, Audronė. Modeliai gebantys prognozuoti nerimą ir depresiją multimorbidškiems pacientams. Tolima ateitis ar realybė? // Dirbtinio intelekto technologijos medicinoje: tyrimai ir diagnostika, 2023 spalio 20 d. : konferencijos tezės. Vilnius : [Lietuvos dirbtinio intelekto asociacija]. 2023, p. 9. Pranešėjos: Roma Puronaitė, Alicija Šavareikaitė	Vietinė
2024 <i>Evolutionary medicine: how evolutionary thinking can contribute to the medical and health sciences: the 6th international conference</i> 2024 m. birželio mėn. 18-21 d. Vilnius, Lietuva (gyvai)	Švaikevičienė, Kristina; Puronaitė, Roma ; Kasiulevičius, Vytautas; Jakaitienė, Audronė. Chronic disease prevalence and clusters in Lithuanian diabetics // Evoliucinė medicina: kaip evoliucinis mąstymas gali prisidėti prie medicinos ir sveikatos mokslų raidos: šeštoji tarptautinė konferencija, 2024 m. birželio mėn. 18-21 d., Vilniaus universitetas, Lietuva = Evolutionary medicine: how evolutionary thinking can contribute to the medical and health sciences: the 6th international conference, 18th-21st of June, 2024, Vilnius University, Lithuania. Vilnius. 2024, p. 65. Prieiga per internetą: https://www.evomed.mf.vu.lt/images/EvolMed_2024_Abstract_book.pdf . Pranešėja: Kristina Švaikevičienė	Tarptautinė

Mokslinių tyrimų ir disertacijos rengimo etapai

Darbo pavadinimas		Atlikimo terminai	Pastabos
1	Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje): Anotuotos bibliografijos sudarymas. Mokslinės literatūros apžvalga. Egzistuojančių metodų taikymo medicinos duomenims analizavimas.	2017 m. spalio mėn. – 2018 m. birželio mėn.	Naudojant sisteminės literatūros apžvalgos metodą, atlikta literatūros analizė šiomis temomis: Sisteminė literatūros apžvalga – dauginio ligotumo modelių klasterizavimo metodai (išnagrinėta 151 publikacija, identifikuoti šioje tematikoje naudojami klasterizavimo metodai) ir Sisteminė literatūros apžvalga – netiesiniai statistikos modeliai: dauginio ligotumo analizės tyrimai naudojant didelės apimties medicinos duomenų masyvus (išnagrinėta 206 publikacijos, identifikuoti šioje tematikoje naudojami klasterizavimo metodai) Apžvelgtos antriniu tikslu naudojamos administracinės duomenų bazės analizės strategijos
2	Mokslinio tyrimo vykdymas: 2.1. Tyrimo metodikos sudarymas: 1. Disertacijos tikslo formulavimas. 2. Disertacijos uždavinių formulavimas.	2018 m. birželio mėn. – 2018 m. lapkričio mėn.	Atlikus literatūros analizę nustatyta, kad tokio tipo duomenys (elektroninių sveikatos įrašų, administracinių duomenų bazių) dažniausiai analizuojami taikant skerspjūvio tipo metodus taikant klasterizavimo metodus duomenims tam tikrame stebėjimo taške ar apibendrintai tam tikro laikotarpio informacijai, neatsižvelgiant į dinamiką. Nuspręsta nagrinėti nuo laiko priklausomų sąryšių įvertinimo galimybes. Disertacijos tikslas: Pasiūlyti klasterizavimo algoritmų papildymą didelės apimties medicinos duomenims, atsižvelgiant į duomenų dinamikos laike savybes. Disertacijos uždaviniai: 1. Ištirti didelės apimties medicinos duomenų klasterizavimui dažniausiai taikomus klasterizavimo metodus. 2. Pasiūlyti klasterizavimo algoritmą ar esamo metodo patobulinimą, kuris atsižvelgtų į dinamiką laike. 3. Pritaikyti atrinktus algoritmus ir pasiūlytą sprendimą realiems medicinos duomenų rinkiniams. 4. Pasiūlyti algoritmo integravimo į sveikatos priežiūros įstaigos informacinę sistemą modelį.

	2.2. Teorinis tyrimas: Matematinio modelio sudarymas. Algoritmų konstravimas ir tobulinimas.	2018 m. lapkričio mėn. – 2019 m. birželio mėn.	Nagrinėjami ir išbandomi nagrinėtoje literatūroje pasiūlyti metodai atsižvelgiant į duomenų dinamiką. Empiriniui tyrimui atrinkta: • Skerspjūvio tipo analizė • Temporalinio fenotipavimo metodai (PARAFAC2, neneigiamas matricos faktorizavimas) • Laiko dimensijos transformavimas (siūloma transformacijai naudoti Poincare grafikus (indeksus), laiko eilučių išlyginimą, netiesinius statistinius modelius) • Kiti: gilusis mokymas, gilieji rekurentiniai neuroniniai tinklai, arba giliojo mokymo teksto analizės metodai.
	2.3. Empirinis tyrimas: 1. Siūlomų algoritmų pritaikymas medicinos duomenims. 2. Siūlomų algoritmų tobulinimas, atsižvelgiant į gautus rezultatus.	2018 m. birželio mėn. – 2020 m. vasario mėn. 2018 m. birželio mėn. – 2022 m. vasario mėn.*	Dauginio ligotumo duomenys pagal atitinkamus ligų kodus buvo suskaidyti į pacientų turinčių tam tikrą ligą pogrupius: • Diabetas • Lėtinė obstrukcinė plaučių liga • Depresija ir nerimas Kiekvieno paciento atveju buvo sudaryti dvinariai kintamieji atitinkantys ligos (išskirta 30 lėtinių ligų) buvimą ar nebuvimą atitinkamo paciento ligos istorijoje per tyrimo laikotarpį. Šiuose pogrupiuose pritaikius hierarchinį klasterizavimą (angl. HCA) gauti ligų klasteriai, kurie buvo įvertinti gydytojų specialistų, dalis šių klasterių įvertinti kaip turintys klinikinį paaiškinimą. Taip pat pritaikyta tiriamoji faktorinė analizė (angl. EFA). Koreliacijos matrica sudaryta naudojant tetrachorinę koreliaciją tarp dvinarių kintamųjų porų. Atlikta vizuali daugiamačių dauginio ligotumo duomenų analizė Pradėta: • Temporalinio fenotipavimo metodai (PARAFAC2, neneigiamas matricos faktorizavimas) Numatoma: • Laiko dimensijos transformavimas (siūloma transformacijai naudoti Poincare grafikus (indeksus), laiko eilučių išlyginimą, netiesinius statistinius modelius) • Kiti: gilusis mokymas, gilieji rekurentiniai neuroniniai tinklai, arba giliojo mokymo teksto analizės metodai.

Pastaba. (2019 10 16–2021 09 30; 2023 03 13–2024 02 13 akademinėse atostogose)

2.4. Gautų duomenų analizė, apibendrinimas, išvadų parengimas: 1. Algoritmų tikslumo įvertinimas, palyginimas su kitų autorių metodais, atrinktais remiantis išanalizuota mokslinė literatūra. 2. Gautų rezultatų apibendrinimas. 3. Išvadų parengimas.	2020 m. vasario mėn. – 2020 m. liepos mėn. 2022 m. vasario mėn. – 2022 m. liepos mėn.*	2021/2022 metais I pusmetis Planuojamos publikacijos (pradėta rengti): Depresijos ir nerimo prognozavimas remiantis administracinio pobūdžio duomenimis, pritaikant Poincare plot metodą ir apskaičiuotus indeksus (Health Informatics tematikos žurnale) Pastaba. Sustabdyta dėl iškilusių metodologinių klausimų susijusių su neapibrėžtumu atsirandančiu dėl skirtingų analitinių sprendimų pasirinkimų. Toliau gilinamasi į neapibrėžtumo šaltinius, rezultatų skirtumus atsirandančius dėl skirtingų analitinių sprendimų (duomenų šaltinio, duomenų paruošimo algoritmų, apimančių vertinamą laiko intervalą, lėtinių ligų įrašų pasikartojimą duomenų rinkinyje, dauginio ligotumo apibrėžimą). Literatūros apžvalga ir iššūkiai susiję su lėtinių ligų vertinimu ir naudojimu pristatyti konferencijoje (IBC 2022) 2021/2022 metais II pusmetis <u>2022 metų liepą pranešimas konferencijoje.</u> Indėlis: metodologinis pagrindas, literatūros apžvalgos rezultatų sisteminimas ir analizė, vizualizacijos, pristatymas konferencijoje (žodinis pranešimas). Puronaite, Roma, Ramanauskaitė, Dovilė, Burneikaitė, Greta, Švaikevičienė, Kristina, Šavareikaitė, Alicija, Vaitkute, Samanta, Jakaitienė, Audronė, Dambrauskas, Laimis, Jurevičienė, Elena, Trinkūnas, Justas, Kasiulevičius, Vytautas, Kazėnaitė, Edita, „Challenges of modeling depression and anxiety risk using data from large healthcare databases: systematic review and situation analysis“, <i>31st International Biometric Conference</i>, 2022 m. liepos 10-15 d., Ryga, Latvija Pranešimas pristatytas 2021/2022 metais II pusmetis Vykdomi: Prognostinių modelių depresijai prognozuoti eksperimentai. Pastaba. Išplėsta į apskritai pagrindinių dauginio ligotumo tyrimų uždavinių rezultatų stabilumo esant skirtingiems analitiniams sprendimams tyrimą. Nagrinėjami šie uždaviniai: (1) paplitimo vertinimo, (2) sveikatos priežiūros resursų naudojimo modelių konstravimo, (3) amžiaus lūžio taškų, kuriuose pakinta dauginio ligotumo paplitimas amžiaus atžvilgiu, identifikavimo.
---	---	---

Pastaba. (2019 10 16–2021 09 30; 2023 03 13–2024 02 13 akademinėse atostogose)

	2022/2023 (akademinės atostogos 2023 03 13–2025 02 13)	2023 metais pristatyti tyrimo, kurio metu palyginti netiesiniai modeliai lėtinių ligų paplitimo vertimui identifikuojant amžiaus lūžio taškus, kai įvyksta paplitimo pokytis, rezultatai. Indėlis: duomenų paruošimas analizei, lėtinių ligų paplitimo vertinimas, netiesiniai modeliai (paplitimo modeliavimas identifikuojant amžiaus lūžio taškus, kai įvyksta paplitimo pokytis), vizualizacijos, pristatymas konferencijoje (stendinis pranešimas). Puronaitė, Roma; Ramanauskaitė, Dovilė; Švaikevičienė, Kristina; Tarutytė, Gabrielė; Trinkūnas, Justas; Burneikaitė, Greta; Kazėnaitė, Edita; Glaveckaitė, Sigita; Jakaitienė, Audronė. Using nonlinear models to identify breakpoints in disease prevalence among patients with multimorbidity // ISCB44: 44th annual conference of the International Society for Clinical Biostatistics; Joint conference with the Italian Region of the International Biometric Society, Milan, Italy, 27-31 August 2023 : final programme & abstract book. Milano : ISCB. 2023, abstract no. MP117, p. 306. Prieiga per internetą: <https://iscb.international/wp-content/uploads/2023/08/Book-of-Abstract-ISCB44_25-AGO-2.pdf>. <u>2025 metų pradžioje</u> publikuotas straipsnis. Analizei naudotas atnaujintas duomenų rinkinys. Indėlis: duomenų paruošimas analizei, dauginio ligotumo paplitimo vertinimas, netiesiniai modeliai (sveikatos priežiūros resursų naudojimas), vizualizacijos, prisidėjimas prie publikacijos teksto rengimo pagal kompetenciją. Ramanauskaitė, Dovilė; Puronaitė, Roma; Jakaitienė, Audronė; Glaveckaitė, Sigita. Prevalence of multimorbidity in Lithuania: insights from national health insurance fund data // Journal of cardiovascular development and disease. Basel : MDPI. eISSN 2308-3425. 2025, vol. 12, iss. 2, art. no. 47, p. [1-18]. DOI: 10.3390/jcdd12020047. <u>2025 gegužės mėn.</u> NBCC konferencijoje numatoma pristatyti pranešimą, kuriame pristatomi naujausi rezultatai. Indėlis: duomenų paruošimas analizei, lėtinių ligų paplitimo vertinimas, netiesiniai modeliai (paplitimo modeliavimas identifikuojant amžiaus lūžio taškus, kai įvyksta paplitimo pokytis), vizualizacijos, pristatymas konferencijoje (stendinis pranešimas). Kilaitė, Lorian; Puronaitė, Roma; Jankauskienė, Augustina; Burneikaitė, Greta; Jakaitienė, Audronė Age-related breakpoints in the prevalence of multimorbidity and common comorbidities: a population-based study in Lithuania
--	--	--

		2024/2025 II pusmečio planai	Atsižvelgiant į tyrimo metu identifikuotus svarbius mokslinio tyrimo klausimus susijusius su neapibrėžtumu atsiradusiu dėl pasirinktų analitinių sprendimų, detaliau nagrinėtas analitinių sprendimų poveikis pagrindiniams dauginio ligotumo tyrimų uždavinių rezultatams (dauginio ligotumo paplitimo įverčiams, sveikatos priežiūros resursų naudojimo modeliams, identifikuotiems amžiaus lūžio taškams). Rezultatus planuojama pristatyti konferencijoje <u>2025 metų rugpjūtį</u> (International Vilnius Conference on Statistics and Its Applications). Santraukų priėmimo terminas: 2025-04-01. <u>2024/2025 metų I pusmečio pab. / 2024/2025 metų II pusmečio pr. (kovas-balandis)</u> Tyrimo metu nagrinėjant literatūrą nustatyta, kad nėra vieningo sutarimo dėl dauginio ligotumo apibrėžimo, o analitinių sprendimų poveikio dauginio ligotumo tyrimų rezultatams, kai analizuojami didelės apimties administraciniai duomenis, tyrimai tebėra fragmentiški, apimantys siaurą sritį (pvz., tik duomenų šaltinių ar tik laiko intervalo, ligų pasikartojimo duomenyse skaičiaus skirtumus). Todėl šiuo metu atliekami ir planuojami užbaigti iki balandžio pabaigos eksperimentai, kuriuose pritaikant daugiakriterinius sprendimo priėmimo metodus siekiama identifikuoti analitinius sprendimus, kurie labiausiai atitinka dauginio ligotumo, kaip vertinamo rodiklio, tikslus ir konceptą. Planuojama ištirti geriausiai įvertintų analitinių sprendimų pritaikomumą realiuose uždaviniuose (rezultatų stabilumą vertinant dauginio ligotumo paplitimo įverčius, sveikatos priežiūros resursų naudojimo modeliams, identifikuotiems amžiaus lūžio taškams). Rezultatus planuojama apibendrinti ir pateikti į sveikatos informatikos arba į medicinos (sveikatos politikos / visuomenės sveikatos / epidemiologijos) krypties žurnalą iki 2025 metų birželio mėn. [Šiuo metu parengtumas yra apie 20 proc., didesnė dalis įvado, metodikos ir dalis rezultatų.]
3	Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų, ir kt.) parengimas: 1. Tyrimų apžvalga ir analizė. 2. Tyrimo metodikos sudarymas. 3. Teorinis tyrimas. 4. Empirinis tyrimas. 5. Gautų duomenų analizė, apibendrinimas. 6. Išvados, įvadas, literatūros sąrašas.	2020 m. rugsėjo mėn. – 2021 m. gegužės mėn. 2025 m. vasaris mėn. – 2025 m. gegužės mėn.*	Disertacijos dalių baigtumas: 1. Tyrimų apžvalga ir analizė – 30%. 2. Tyrimo metodikos sudarymas – 30%. 3. Teorinis tyrimas – 30%. 4. Empirinis tyrimas – 30%. 5. Gautų duomenų analizė, apibendrinimas – 20%. 6. Išvados (0%), įvadas (0%), literatūros sąrašas (20%). 2025 balandžio mėn. numatomas pirmasis juodraštis.
4	Daktaro disertacijos parengimas ir svarstymas padalinyje	2021 m. birželio mėn. 2025 m. birželio mėn.*	
5	Daktaro disertacijos gynimas	2021 m. rugsėjo mėn. 2025 m. rugsėjo mėn.*	

Trumpas įvadas

Dauginis ligotumas, duomenys

Medicinos srities problematika

Dažniausiai sutinkamas **dauginio ligotumo (multimorbidiškumo, MM)** apibrėžimas:

Bent dvi lėtinės ligos pasireiškiančios kartu vienam individui.

Multimorbidiškumo „diagnozavimas“ pagal 2+ apibrėžimą nėra didelė problema, jei yra žinoma kokiomis lėtinėmis ligomis pacientas serga. Tačiau:

- Mokslininkai nesutaria dėl apibrėžimo (ligų sąrašo, kiek ligų skaičiuoti, ar nereikėtų išvesti indekso ir kt.).
- Sudėtinga nuspręsti kada MM reikalauja specifinio dėmesio ir kompleksinės priežiūros, o kada ne.

Epidemiologinė problematika

Dauginio ligotumo paplitimo vertinimas Lietuvos gyventojų populiacijoje.

Vertinimo būdai: reprezentatyvi statistinė gyventojų apklausa, el. sveikatos įrašai (e-sveikata), **administraciniai sveikatos duomenys (ASD)**. Tačiau:

- Gyventojų suvokimas apie savo turimas ligas subjektyvus.
- E-sveikatos duomenų apdorojimą sunkina tai, kad dauguma duomenų nestruktūruoti.
- ASD – pirminis tikslas paslaugų apskaita, vis dėlto šis duomenų šaltinis turi ilgiausią prieinamą laikotarpį.

Informatikos, statistikos problematika

Dauginio ligotumo nustatymo DB algoritmas ASD duomenų rinkinyje

Reikalingi sprendimai.

- Duomenų transformacija iš pirminio formato.
- Skirtingų analitinių sprendimų poveikis rezultatų stabilumui vertinimas.
- Algoritmo parinkimas MM nustatymui, kai nėra vieningo apibrėžimo.
- Algoritmo tinkamumo dauginio ligotumo tyrimams vertinimas.

Trumpas įvadas

Duomenų bazės struktūra

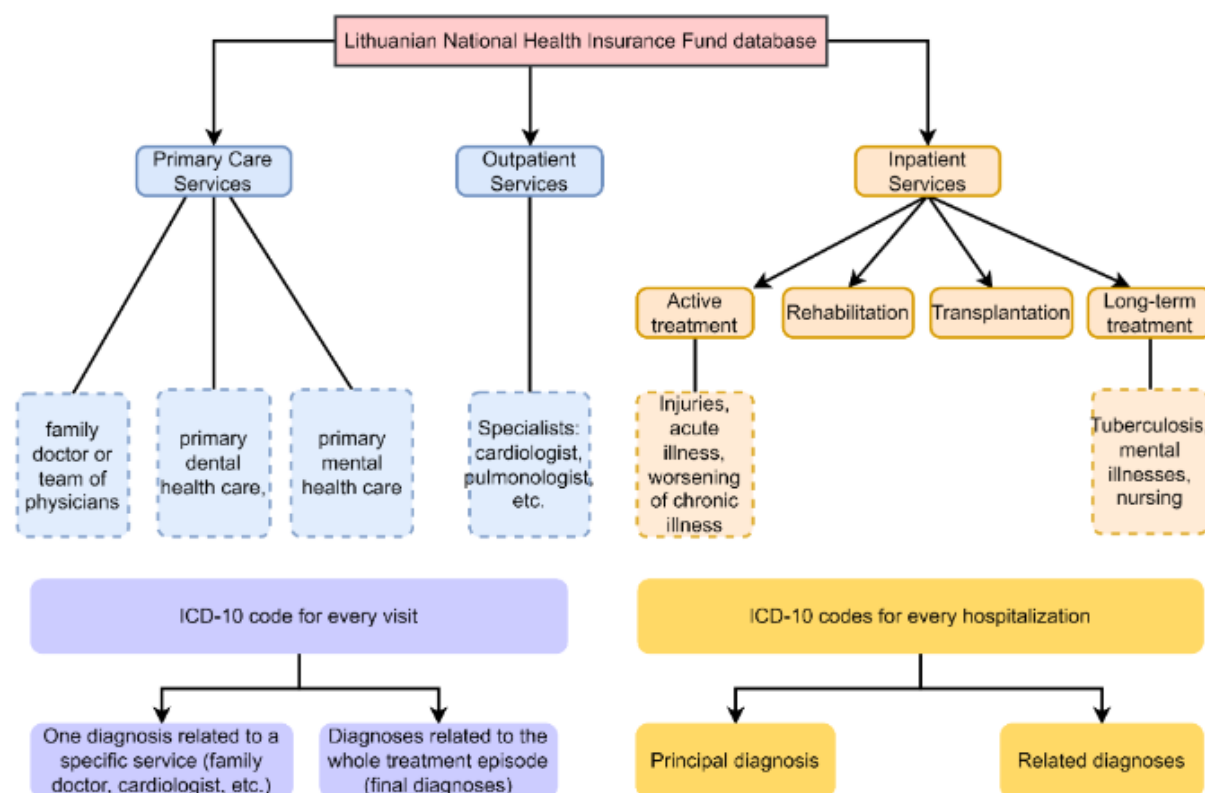


Fig. 1 Structure of Lithuanian NHIF primary database.

Valstybinės ligonių kasos DB SVEIDRA

Pagrindinės dalys:

- Šeimos gydytojo ir jo komandos paslaugos (angl. *primary care*).
- Gydytojų specialistų konsultacijos (angl. *outpatient care*).
- Stacionarus gydymas (ligoninėje) (angl. *inpatient services*)

Dėl sveikatos priežiūros paslaugų pobūdžio skirtumų, duomenų struktūra kiekvienoje iš dalių skiriasi.

Liga	Diagnozė
Chronic condition	ICD-10-AM diagnostic codes
Cancer	C00-C96
Anemia	D50
Hypothyroidism	E02, E03, E89.0
Diabetes	E10.0 - E10.9, E11.01 - E11.9

Pagrindinė dalis: analitiniai sprendimai

Duomenų šaltinis	Laiko intervalas	Ligos patvirtinimui pasirinktas atsikartojimo DB skaičius	Dauginio ligotumo apibrėžimas
S1: Šeimos medicina (P) S2: Specialistai (O) S3: Stacionaras (I) S4: P + O S5: P + I S6: O + I S7: P + O + I	T1: Tik analizuojami metai* T2: +1 metai atgal** T3: +2 metai atgal T4: +3 metai atgal T5: +4 metai atgal T5: +5 metai atgal	N1: Bent 1 kartą per analizuojamą laikotarpį* N2: Bent 2 kartus per analizuojamą laikotarpį N3: Bent 3 kartus per analizuojamą laikotarpį	M1: Bent dvi lėtinės ligos* M2: Bent trys lėtinės ligos M3: Bent dvi lėtinės ligos iš bent dviejų specialybių** M4: Bent trys lėtinės ligos iš bent dviejų specialybių M5: Bent viena fizinė ir bent viena psichiatrinė diagnozė
	*pvz., tikslas įvertinti dauginio ligotumo būklę 2019 metais ** tada įvertinsime paplitimą 2019 metais, remdamiesi tuo ar asmuo gydėsi dėl lėt. ligų 2018-2019 metais	* Pasirinktam šaltiniui ir vertinimo laikotarpiui	*pvz., hipertenzija ir išeminė širdies liga ** pvz., hipertenzija ir cukrinis diabetas (kardiologija, endokrinologija)
MacRae et al. (2023A), Ho et al. (2021)	Simard et al. (2024)	Simard et al. (2024), Beaney et al. (2024)	Simard et al. (2024), MacRae et al. (2023B), Ho et al. (2021), Chua et al. (2021)

Literatūra

Analitiniai sprendimai

M. Simard, E. Rahme, M. Dubé, V. Boiteau, D. Talbot, and C. Sirois, 'Multimorbidity prevalence and health outcome prediction: assessing the impact of lookback periods, disease count, and definition criteria in health administrative data at the population-based level', *BMC Med Res Methodol*, vol. 24, no. 1, Art. no. 1, May 2024, doi: [10.1186/s12874-024-02243-0](https://doi.org/10.1186/s12874-024-02243-0).

T. Beaney, J. Clarke, T. Woodcock, A. Majeed, M. Barahona, and P. Aylin, 'Effect of timeframes to define long term conditions and sociodemographic factors on prevalence of multimorbidity using disease code frequency in primary care electronic health records: retrospective study', *bmjmed*, vol. 3, no. 1, Art. no. 1, Feb. 2024, doi: [10.1136/bmjmed-2022-000474](https://doi.org/10.1136/bmjmed-2022-000474).

C. MacRae *et al.*, 'Impact of data source choice on multimorbidity measurement: a comparison study of 2.3 million individuals in the Welsh National Health Service', *BMC Med*, vol. 21, no. 1, Art. no. 1, Aug. 2023, doi: [10.1186/s12916-023-02970-z](https://doi.org/10.1186/s12916-023-02970-z). (A)

C. MacRae *et al.*, 'Age, sex, and socioeconomic differences in multimorbidity measured in four ways: UK primary care cross-sectional analysis', *Br J Gen Pract*, vol. 73, no. 729, Art. no. 729, Apr. 2023, doi: [10.3399/BJGP.2022.0405](https://doi.org/10.3399/BJGP.2022.0405). (B)

I. S.-S. Ho *et al.*, 'Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies', *The Lancet Public Health*, vol. 6, no. 8, Art. no. 8, Aug. 2021, doi: [10.1016/S2468-2667\(21\)00107-9](https://doi.org/10.1016/S2468-2667(21)00107-9).

Y. P. Chua, Y. Xie, P. S. S. Lee, and E. S. Lee, 'Definitions and Prevalence of Multimorbidity in Large Database Studies: A Scoping Review', *IJERPH*, vol. 18, no. 4, Art. no. 4, Feb. 2021, doi: [10.3390/ijerph18041673](https://doi.org/10.3390/ijerph18041673).

Pagrindinė dalis

Dažniausi dauginio ligotumo tyrimų uždaviniai

Dauginio ligotumo tyrimų uždaviniai:

(1) paplitimo vertinimas,

(2) sveikatos priežiūros resursų
naudojimo modeliai,

(3) amžiaus lūžio taškų, kuriuose
pakinta dauginio ligotumo
paplitimas amžiaus atžvilgiu,
identifikavimo.

1. Ramanauskaitė, **Puronaite**, Jakaitienė, Glaveckaitė
Prevalence of multimorbidity in Lithuania: insights from
national health insurance fund data DOI:
10.3390/jcdd12020047. Straipsnis.

2. **Puronaite**; Ramanauskaitė, Švaikevičienė, et al.
Using nonlinear models to identify breakpoints in
disease prevalence among patients with multimorbidity
// ISCB 2023. Konferencija

3. Kilaitė, **Puronaite**, Jankauskienė, Burneikaitė, Jakaitienė
Age-related breakpoints in the prevalence of
multimorbidity and common comorbidities: a population-
based study in Lithuania // NBCC 2025. Konferencija

Publikacija: trumpa apžvalga

Tikslas

Įvertinti dauginio ligotumo paplitimą 2019 metais vertinamą keliais skirtingais apibrėžimais. Įvertinti dauginio ligotumo būklės ir pacientams suteiktų paslaugų lygio ryšius.

Populiacija

Pilnamečiai (18+) Lietuvos gyventojai.
Asmenys turintys bent vieną lėtinę ligą.

Duomenų šaltiniai

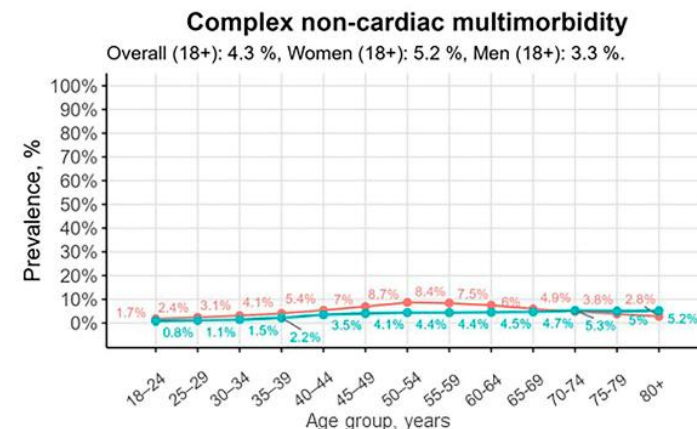
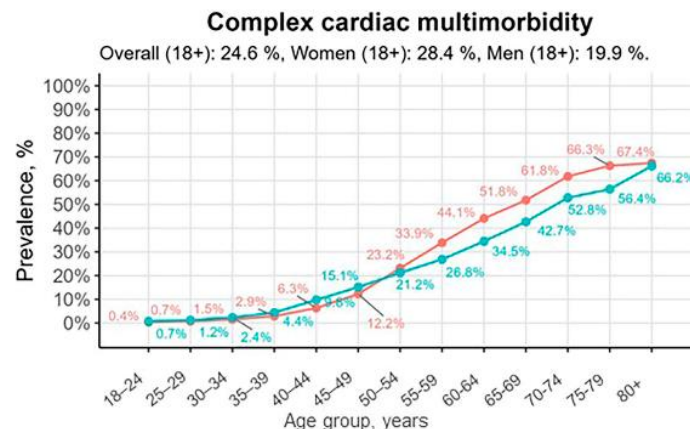
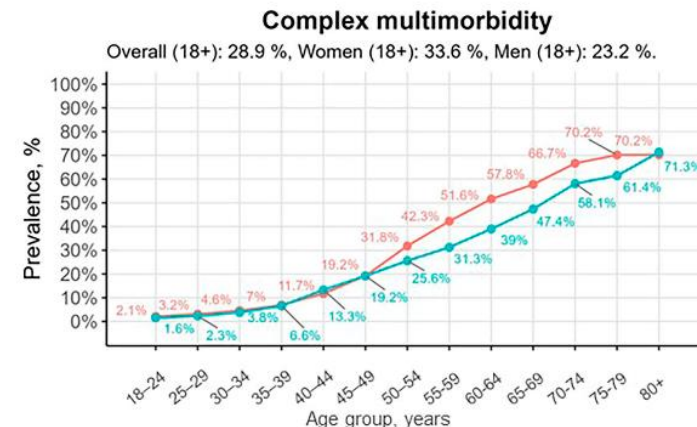
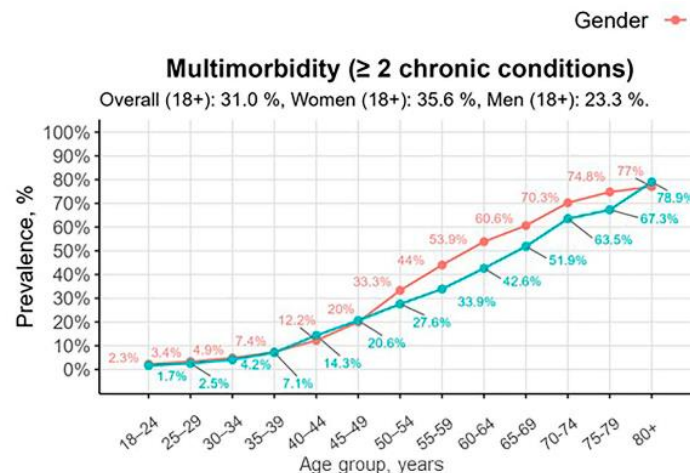
Valstybės duomenų agentūra – Lietuvos gyventojų skaičius 2019 metų pradžioje, skirstant pagal amžių ir lytį.

Valstybinės ligonių kasos informacinės sistemos SVEIDRA duomenų bazė – 2019 metais asmenims, kurie bent kartą 2019 metais lankėsi dėl lėtinės ligos, suteiktos paslaugos ir su kiekviena jų susietos diagnozės (TLK-10-AM kodai).

Pagrindiniai rezultatai

Dauginio ligotumo apibrėžimą atitinka (pagal 2+ apibrėžimą) 31 proc. 18 ir vyresnių Lietuvos gyventojų, 18–24 metų amžiuje tokių asmenų yra apie 2 proc. toliau kyla iki 77,5% (85 ir daugiau).

Kas trečias dauginį ligotumą turintis žmogus bent kartą per metus hospitalizuojamas, o kardiovaskulinį dauginį ligotumą turintys pacientai turi didesnius šansus būti hospitalizuoti, pakartotinai hospitalizuoti per 30 dienų, dažniau lankytis pas šeimos gydytojus lyginant su asmenimis sergančiais tik viena lėtine liga (ŠS: 2,23, 1,60, 4,24, $p < 0.001$).



Pav. 2. Dauginio ligotumo paplitimas. Keturi dauginio ligotumo apibrėžimai: tradicinis (2+ lėtinės ligos), kompleksinis (2+ lėtinės ligos iš 2+ specialybių (pvz. kardiologija ir endokrinologija)), kompleksinis kardiologinis (kompleksinis, kai būtinai turima bent viena kardiologinė liga), kompleksinis ne kardiologinis (kompleksinis, kai galimo visos kitos ligos išskyrus kardiologines).

ISCB 2023 konferencija: trumpa apžvalga

Tikslas

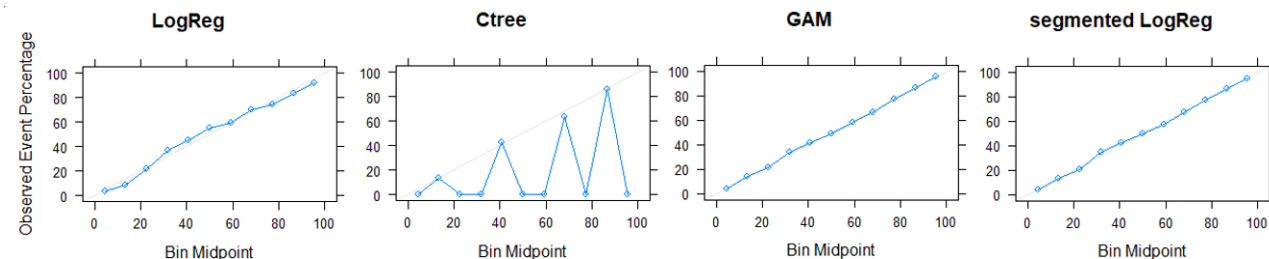
Palyginti netiesinių modelių tinkamumą klasifikuoti kelias lėtines ligas turinčius pacientus į tuos, kurie turi/neturi širdies nepakankamumą ir turi/neturi cukrinį diabetą. Identifikuoti amžiaus grupes, ties kuriomis pakinta modeliuojama ligos tikimybė.

Populiacija

Pilnamečiai (18+) Lietuvos gyventojai turintys bent dvi lėtines ligas (asmenys, kurie atitinka dauginio ligotumo apibrėžimą).

Duomenų šaltiniai

Valstybinės ligonių kasos informacinės sistemos SVEIDRA duomenų bazė – 2012.01-2014.06 metais asmenims, kurie bent kartą šiame laikotarpyje lankėsi dėl bent dviejų lėtinių ligų, suteiktos paslaugos ir su kiekviena jų susietos diagnozės (TLK-10-AM kodai).



Pav. 3. Lyginamų modelių kalibravimo kreivės.

Lentelė. Lyginamų modelių diskriminacinės savybės.

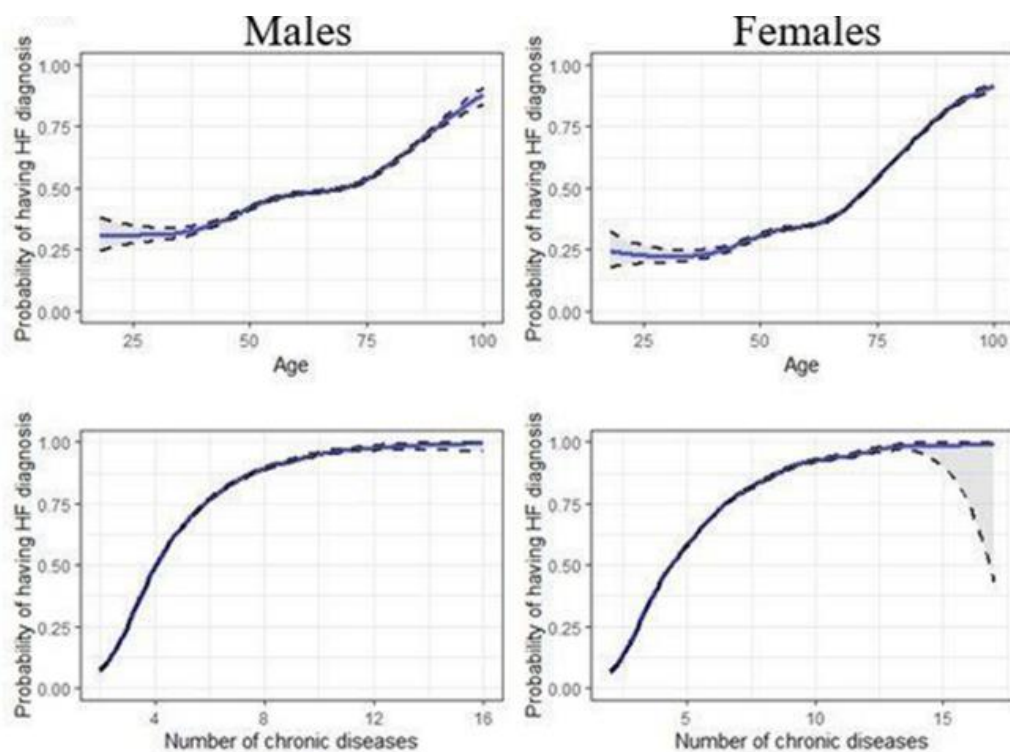
Metodas	AUC (95% PI)	Jautrumas	Specifiškumas
Širdies nepakankamumas (HF), vyrai			
Logistinė regresija	0,828 (0,823-0,832)	0,772	0,722
Klasifikavimo medis	0,801 (0,796-0,805)	0,721	0,756
Apibendrintasis adityvusis modelis (GAM)	0,829 (0,825-0,833)	0,758	0,737
Segmentinė regresija	0,829 (0,824-0,833)	0,764	0,729
Moterys (HF), vyrai			
Logistinė regresija	0,820 (0,817-0,824)	0,757	0,728
Klasifikavimo medis	0,771 (0,767-0,775)	0,771	0,653
Apibendrintasis adityvusis modelis (GAM)	0,825 (0,821-0,828)	0,772	0,714
Segmentinė regresija	0,825 (0,821-0,828)	0,751	0,735
Cukrinis diabetas (DM), vyrai			
Logistinė regresija	0,607 (0,600-0,613)	0,595	0,553
Klasifikavimo medis	0,598 (0,592-0,603)	0,834	0,307
Apibendrintasis adityvusis modelis (GAM)	0,627 (0,620-0,633)	0,619	0,554
Segmentinė regresija	0,626 (0,620-0,633)	0,613	0,558
Cukrinis diabetas (DM), moterys			
Logistinė regresija	0,585 (0,580-0,591)	0,613	0,505
Klasifikavimo medis	0,583 (0,578-0,588)	0,804	0,329
Apibendrintasis adityvusis modelis (GAM)	0,616 (0,611-0,621)	0,591	0,576
Segmentinė regresija	0,616 (0,611-0,621)	0,568	0,598

ISCB 2023 konferencija

Trumpa apžvalga

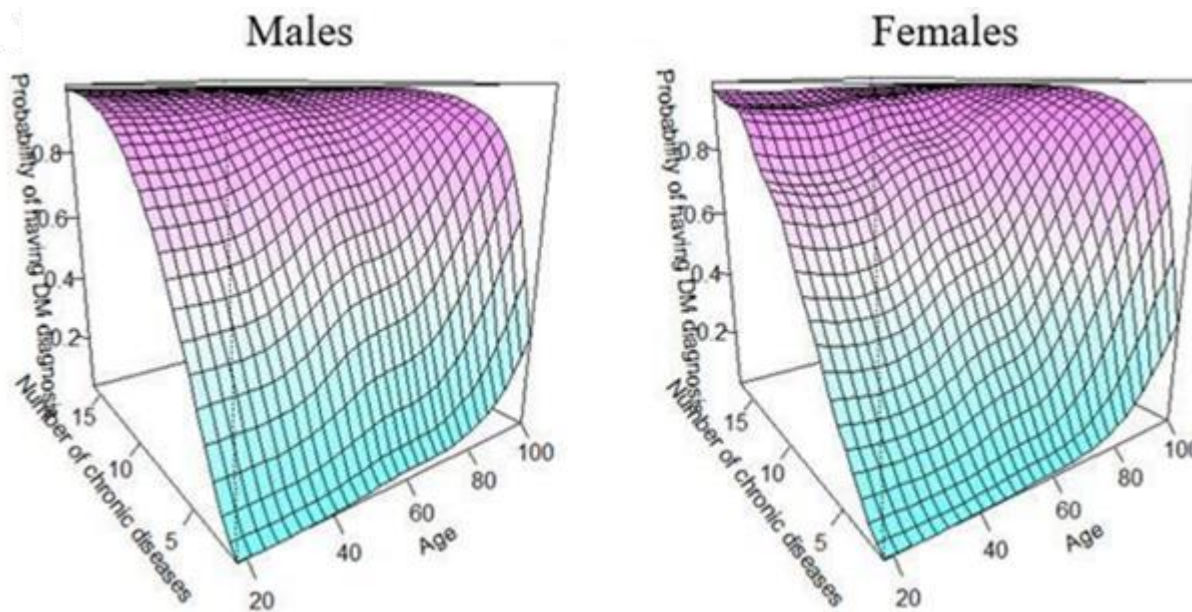
A

GAM



B

GAM



Pav. 4. Širdies nepakankamumo (HF) tikimybės modeliuotos taikant GAM modelį asmenims turintiems dvi ar daugiau lėtinių ligų (dauginis ligotumas).
A – 2D, B – 3D atvaizdavimas.

Analitiniai sprendimai ir dauginio ligotumo tyrimai

Dauginio ligotumo paplitimas

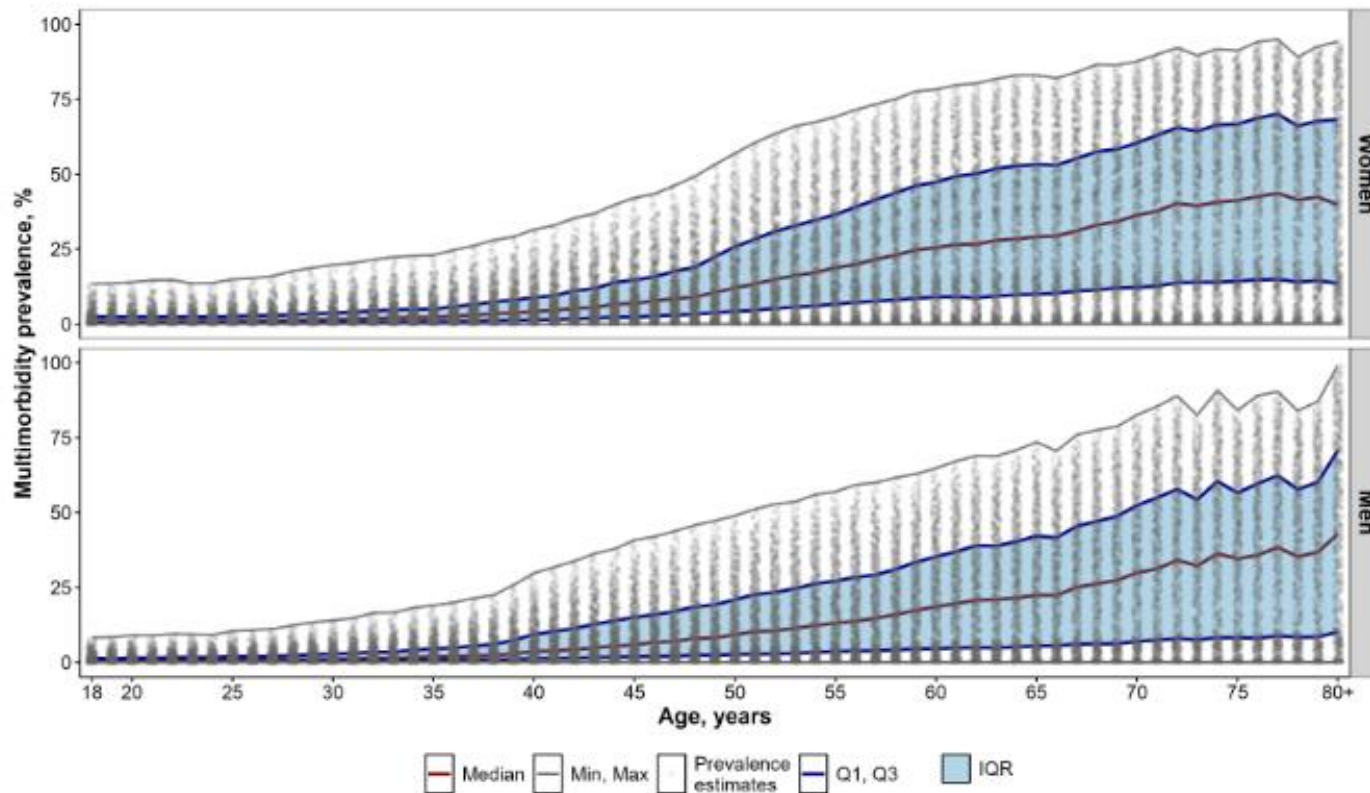


Table 2 The list of variants for four themes used for experiments.

Source of data	Timeframe	Number of appearances per timeframe to confirm	MM definition (Number of confirmed conditions)
S1: Primary care (P)	T1: Year of interest	N1: At least 1 time	M1: At least 2
S2: Outpatient care (O)	T2: +1 year before	N2: At least 2 time	M2: At least 3
S3: Inpatient care (I)	T3: +2 years before	N3: At least 3 time	M3: At least 2 from 2 specialties
S4: P + O	T4: +3 years before		M4: At least 3 from 2 specialties
S5: P + I	T5: +4 years before		M5: At least 1 physical + mental
S6: O + I	T6: +5 years before		
S7: P + O + I			

- Analitinių sprendimų variantų skaičius: 630.
- Paplitimas jaunesniame amžiuje svyruoja mažesniame intervale.
- Vyriausiems asmenims svyravimai dideli.

Pav. 5. Dauginio ligotumo paplitimas Lietuvoje 2019 metais, skaičiavimams naudojant 630 skirtingų analitinių sprendimų.

Analitiniai sprendimai ir dauginio ligotumo tyrimai

Dauginio ligotumo paplitimas

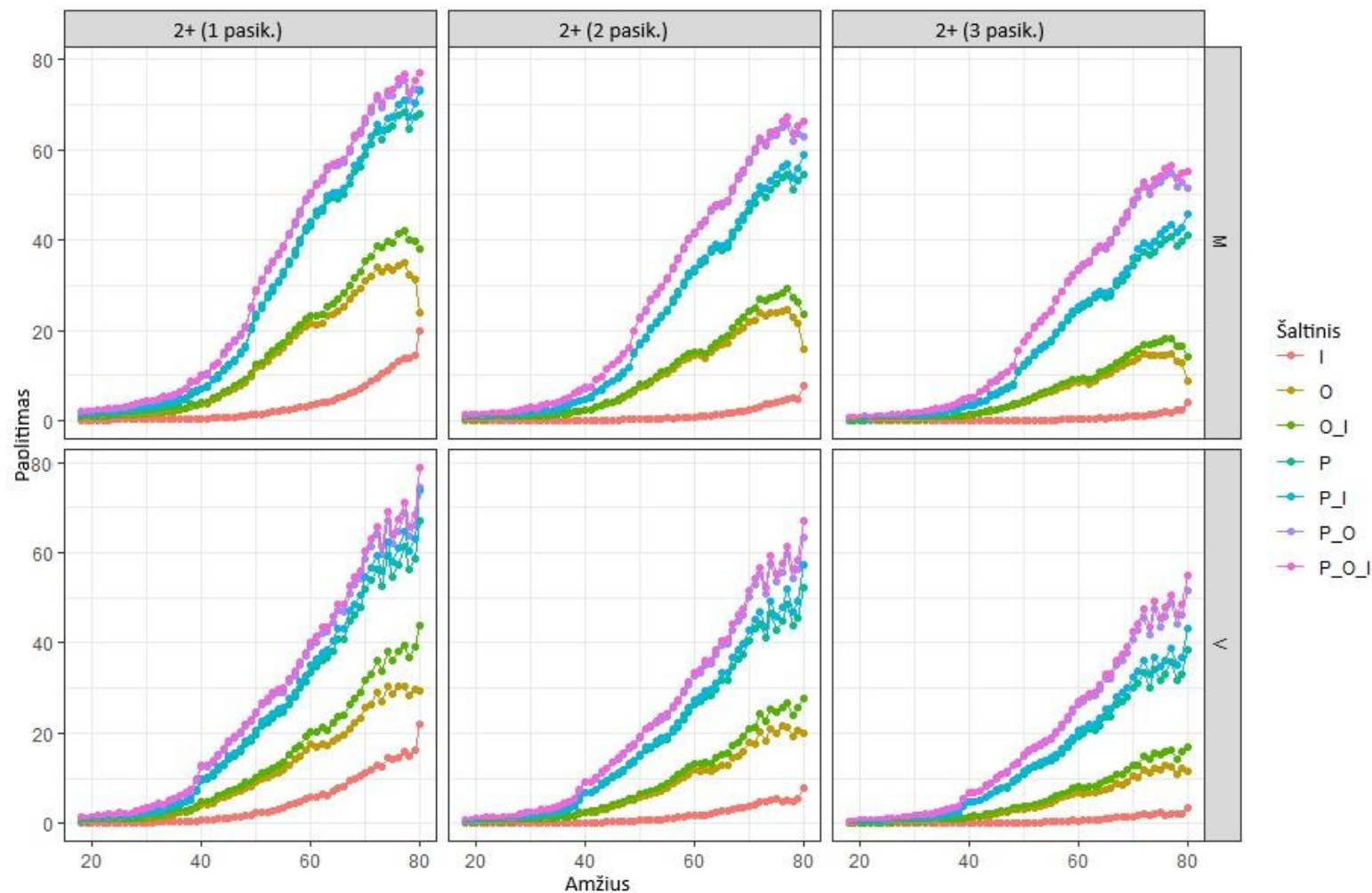


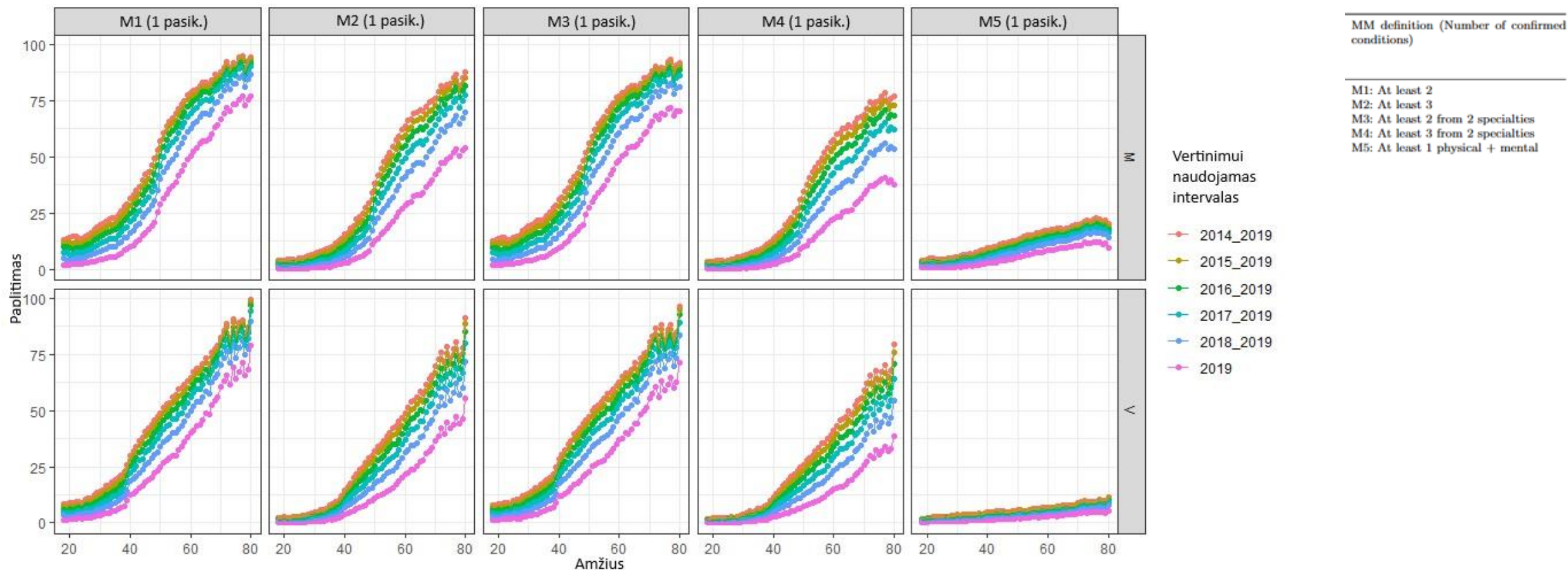
Table 2 The list of variants for four themes used for experiments.

Source of data	Timeframe	Number of appearances per timeframe to confirm	MM definition (Number of confirmed conditions)
S1: Primary care (P)	T1: Year of interest	N1: At least 1 time	M1: At least 2
S2: Outpatient care (O)	T2: +1 year before	N2: At least 2 time	M2: At least 3
S3: Inpatient care (I)	T3: +2 years before	N3: At least 3 time	M3: At least 2 from 2 specialties
S4: P + O	T4: +3 years before		M4: At least 3 from 2 specialties
S5: P + I	T5: +4 years before		M5: At least 1 physical + mental
S6: O + I	T6: +5 years before		
S7: P + O + I			

Pav. 6. Dauginio ligotumo paplitimas Lietuvoje 2019 metais. Apibrėžimai: [S1-S7, T1, N1-N3, M1]

Analitiniai sprendimai ir dauginio ligotumo tyrimai

Dauginio ligotumo paplitimas



Pav. 6. Dauginio ligotumo paplitimas Lietuvoje 2019 metais. Apibrėžimai: [S7, T1-T6, N1, M1-M5]

Apibendrinimas

DABAR

Dauginio ligotumo tyrimų uždaviniams naudojant dažniausiai taikomą analitinių sprendimų variantą:

- (1) paplitimo vertinimas – matomos bendros tendencijos, palygintina su panašų apibrėžimą naudojančiais tyrimais.
- (2) sveikatos priežiūros resursų naudojimo modeliai – rezultatai atitinka bendrą tendenciją, kad MM turintys asmenys dažniau naudojami sveikatos priežiūros resursais.
- (3) amžiaus lūžio taškų, kuriuose pakinta dauginio ligotumo paplitimas amžiaus atžvilgiu, identifikavimas – naudojant netiesinius statistikos metodus galima identifikuoti taškus, ties kuriais pasikeičia dauginio ligotumo paplitimas.

KITĄ PUSMETĮ

Dauginio ligotumo tyrimų uždaviniams naudojant geriausius sprendimų variantus:

- (1) paplitimo vertinimas – objektyvizuotas apibrėžimo pasirinkimas.
- (2) sveikatos priežiūros resursų naudojimo modeliai – MM grupės, kurios labiau atitinka specifinio dėmesio ir kompleksinės priežiūros poreikį.
- (3) amžiaus lūžio taškų, kuriuose pakinta dauginio ligotumo paplitimas amžiaus atžvilgiu, identifikavimas – amžiaus taškai, kurie labiau atitinka paplitimą įvertintą objektyvizuotu būdu.

Kito pusmečio planas

- Užbaigti analitinių sprendimų analizę.

Daugiakriteriniai sprendimų priėmimo metodai analitiniams sprendimams pasirinkti. Kriterijai suformuoti remiantis dauginio ligotumo konceptu (sunkumas, didelis sveikatos priežiūros resursų naudojimas, kompleksiskumas): reikalavimai mirtingumo lygiui MM grupėje, „sunkių pacientų“ (angl. *heavy users*) daliai, lėtinių ligų tarpusavio ryšio stiprumas (ligų tinklo tankis).

Geriausi sprendimai ir rezultatų stabilumas pagrindiniams dauginio ligotumo uždaviniams.

- Susisteminti doktorantūros metu atliktus tyrimus.

Susikoncentruojant į esminius rezultatus, atsisakant perteklinės informacijos.

- Užbaigti rengiamas disertacijos dalis.

2025 balandžio mėn. numatomas pirmasis juodraštis.