

ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

I ПОДАЦИ О КОМИСИЈИ		
1. Датум и орган који је именовано комисију: 26.04.2023. Наставно-научно веће Медицинског факултета Универзитета у Новом Саду		
2. Састав комисије у складу са <i>Правилима докторских студија Универзитета у Новом Саду</i> :		
1. Проф. др Александар Реџек	редовни професор	Хирургија са анестезиологијом (Кардиохирургија), 01.08.2020.
презиме и име	звање	ужа научна област и датум избора
Медицински факултет Универзитета у Новом Саду		председник комисије
установа у којој је запослен-а		функција у комисији
2. Проф. др Голуб Самарџија	ванредни професор	Патологија, 13.4.2022.
презиме и име	звање	ужа научна област и датум избора
Медицински факултет Универзитета у Новом Саду		члан комисије
установа у којој је запослен-а		функција у комисији
3. Проф. др Ненад Филиповић	редовни професор	Примењена механика и Примењена информатика и рачунарско инжењерство, 27.5.2010.
презиме и име	звање	ужа научна област и датум избора
Факултет инжењерских наука Универзитета у Крагујевцу		члан комисије
установа у којој је запослен-а		функција у комисији
II ПОДАЦИ О КАНДИДАТУ		
1. Име, име једног родитеља, презиме: Мила (Сава) Главашки		
2. Датум рођења, општина, држава: 22.03.1985., Сента, Србија		
3. Назив факултета, назив претходно завршеног нивоа студија и стечени стручни/академски назив:		

Фармацеутски факултет Универзитета у Београду, дипломске основне студије, дипломирани фармацеут—медицински биохемичар
Медицински факултет и Природно-математички факултет Универзитета у Новом Саду, мастер академске студије, мастер ембриологије

4. Година уписа на докторске студије и назив студијског програма докторских студија:
2016. године, Докторске академске студије—Претклиничка истраживања

III НАСЛОВ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Утврђивање повезаности генотипа и фенотипа хипертрофичне кардиомиопатије применом машинског учења

IV ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Навести кратак садржај са назнаком броја страница, поглавља, слика, схема, графикана и сл.

Докторска дисертација је написана систематично, прегледно, јасним и разумљивим стилем, обухвата 428 страница и садржи неопходне елементе научног рада. Дисертација је организована у 7 поглавља:

1. УВОД — написан на 76 страна, укључујући укупно 2 слике и 5 табела.
 2. ЦИЉЕВИ ИСТРАЖИВАЊА И ХИПОТЕЗЕ— приказани су циљеви постављени на основу података приказаних у уводном поглављу, радне хипотезе су проистекле из постављених циљева.
 3. МАТЕРИЈАЛИ И МЕТОДЕ — у овом поглављу налази се детаљан опис начина избора и величине узорка, протокола и метода истраживања као и статистичке обраде података. Поглавље је написано на 24 страна, укључујући и 1 слику.
 4. РЕЗУЛТАТИ — резултати су приказани систематично уз текстуални опис. Поглавље је написано на 249 страна, укључујући укупно 50 слика, 149 графикана и 126 табеле.
 5. ДИСКУСИЈА — написана је детаљна дискусија добијених резултата и поређење са релевантним публикованим студијама. Поглавље је написано на 11 страна.
 6. ЗАКЉУЧАК — изведен на основу добијених резултата дисертације. Приказан је на 2 стране. 1-6. Проширени сажетак на српском језику обухвата сва претходно наведена поглавља. Написан је на српском језику на 25 страна.
 7. ЛИТЕРАТУРА — цитирано је 349 библиографских јединица у складу са важећим правилима за цитирање Ванкуверским цитатним стилем.
- У посебном делу докторске дисертације дат је списак свих скраћеница коришћених у рукопису дисертације и наведен је на 8 страна.

У дисертацији је представљена 131 табела, 53 слике и 149 графикана. На почетку докторске дисертације дата је Кључна документацијска информација на српском и енглеском језику са изводима, а на крају, иза текста дисертације, налази се План третмана података.

V ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Наслов докторске дисертације „Утврђивање повезаности генотипа и фенотипа хипертрофичне кардиомиопатије применом машинског учења“ је јасно и прецизно формулисан и у потпуности у складу са изнетим садржајем истраживања.

Комисија сматра да је наслов докторске дисертације јасно и прецизно формулисан и адекватно осликава садржај и тематику докторске дисертације.

Увод докторске дисертације је написан систематично, са детаљним и садржајним приказом актуелних сазнања о изучаваној проблематици и јасно указује на комплексност уоченог проблема и потребе за спровођењем истраживања. Ово поглавље је подељено у 8 мањих целина са подцелинама. У првој целини дат је приказ хипертрофичне кардиомиопатије, а у другој приказ кардиомиопатија генерално. Пошто се хипертрофична кардиомиопатија понекад описује као „болест саркомере“, након тога је детаљно представљена саркомера и њене компоненте. Након тога су приказани генотип, затим фенотип хипертрофичне кардиомиопатије, а потом повезаности генотипа и фенотипа хипертрофичне кардиомиопатије. Након тога, описан је значај хипертрофичне кардиомиопатије у савременој

кардиологији и кардиохирургији. У последњој мањој целини увода дат је осврт на машинско учење и његову примењивост у биомедицини.

У уводном делу кандидат је јасно, детаљно, свеобухватно, систематично и уз употребу актуелних и релевантних литературних извора указао на комплексност ученог проблема и оправдао сврсисходност спровођења истраживања.

Циљеви истраживања су јасно дефинисани и омогућавају доношење конкретних закључака. Циљеви истраживања који су наведени у дисертацији засновани су на резултатима ранијих истраживања објављених у релевантним научним публикацијама. Циљеви спроведеног истраживања су били следећи:

Циљ 1. Идентификација субфенотипова хипертрофичне кардиомиопатије применом кластеризације

Циљ 2. Утврђивање корелације између генотипова и субфенотипова хипертрофичне кардиомиопатије

Циљ 3. Утврђивање повезаности генотипа и исхода фенотипа хипертрофичне кардиомиопатије уз помоћ алгоритама машинског учења

Циљ 4. Израда модела машинског учења за предвиђање исхода хипертрофичне кардиомиопатије на основу генотипских и фенотипских података

Циљ 5. Идентификација генотип-специфичних налаза ехокардиограма хипертрофичне кардиомиопатије уз помоћ алгоритама машинског учења

Хипотезе истраживања су логичне, јасно и прецизно формулисане и у складу су са постављеним циљевима истраживања:

Хипотеза 1. Субфенотипови хипертрофичне кардиомиопатије могу се идентификовати кластеризацијом

Хипотеза 2. Постоји корелација између генотипова и субфенотипова хипертрофичне кардиомиопатије

Хипотеза 3. Повезаност генотипа и исхода фенотипа хипертрофичне кардиомиопатије може се утврдити уз помоћ алгоритама машинског учења

Хипотеза 4. Модели машинског учења могу да предвиде исходе хипертрофичне кардиомиопатије на основу генотипских и фенотипских података

Хипотеза 5. Генотип-специфични налази ехокардиограма хипертрофичне кардиомиопатије могу се идентификовати уз помоћ алгоритама машинског учења

Комисија сматра да су циљеви истраживања и радне хипотезе логично и прецизно формулисани, јасно изложени, научно аргументовани, као и да омогућавају доношење конкретних закључака.

Материјали и методе – у овом поглављу су јасно и прецизно описани материјали и методе који су коришћени у раду: прецизно је дефинисан начин избора и величина узорка истраживања, као и критеријуми за укључивање испитаника у студију, односно искључивање из ње. Детаљно је описан ток истраживања, а предложене методе рада су јасно формулисане. Методологија рада је у складу са принципима научноистраживачког рада у медицини и омогућава понављање истраживања. Коришћене методе су примерене врсти студије, прихваћене и коришћене у широј научној заједници. Изабране методе одговарају постављеним циљевима и омогућавају добијање конкретних и квалитетних научних резултата. Спровођење истраживања је одобрено одлукама етичких комисија сваке од институција које су учествовале у истраживању, као и Комисије за етичност клиничких испитивања Медицинског факултета Универзитета у Новом Саду. Студија је била мултицентрична и ретропроспективна, обухватила је 143 одрасла пацијента са потврђеном дијагнозом хипертрофичне кардиомиопатије, који су испунили укључујуће и искључујуће критеријуме. Прикупљени су демографски и анамнестички подаци, извршена су антропометријска мерења, узорковање крви и генетичко тестирање, хематолошке и биохемијске анализе крви, трансторакална ехокардиографија са доплером, кардиопулмонални тест оптерећењем, електрокардиографија (ЕКГ) и ЕКГ-холтер мониторинг. Затим су прикупљени подаци припремљени за анализу (највећим делом коришћењем библиотеке Pandas програмског језика Python), анализирани применом кластеризације (коришћењем

Scikit-learn модула програмског језика Python), креирањем модела са предвиђање исхода фенотипа (коришћењем Scikit-learn и SHAP модула програмског језика Python), као и креирањем модела за рачунарски вид приликом идентификације генотип-специфичних налаза ехокардиограма (коришћењем библиотека за дубоко учење и рачунарски вид програмског језика Python).

Комисија сматра да су изабране методе прикупљања података и методе статистичке обраде података адекватне, у складу су и примерене постављеним циљевима и хипотезама, те омогућавају добијање поузданих, актуелних и квалитетних научних резултата, а описане су јасно и детаљно, чиме се у потпуности омогућава поновљивост експеримената.

Резултати су уверљиво, прегледно и стручно приказани табеларно и графички. Прикази су праћени јасним текстуалним тумачењима. Резултати јасно произилазе из примењене методологије, уз коришћење одговарајућих статистичких метода. Подтипови хипертрофичне кардиомиопатије идентификовани су на основу свих доступних података о фенотипу. Пронађене су одређене значајне корелације између генотипа и субфенотипова. Предикција исхода фенотипа извршена је на основу информације о мутираним генима, као и на основу података о генотипу и фенотипу. Идентификовани су генотип-специфични налази ехокардиограма.

Комисија сматра да су резултати оригинални, свеобухватни и јасни за интерпретацију, да произилазе из примењене методологије и да су приказани систематично, прегледно, логичним редоследом и на начин који је примерен типу података, са одговарајућим статистичким методама обраде података.

Дискусија је свеобухватна и јасно, прегледно и детаљно написана, по логичном редоследу. Кандидат аргументовано и критички анализира резултате истраживања и упоређује их са наводима из релевантне литературе. Литературни подаци су адекватно одабрани, актуелни и релевантни за извођење валидних закључака из проучаване проблематике. Коментари и тумачења резултата су исцрпни, логични и научно утемељени. Стил писања је јасан, прецизан, логичан и прегледан, и недвосмислено указује на упућеност кандидата у истраживану проблематику. Кандидат је у тексту дискусије показао способност критичке анализе и интерпретације научних података, како резултата сопственог истраживања, тако и података из релевантних литературних извора.

Комисија сматра да је дискусија добијених резултата истраживања исцрпна, прецизна, научно утемељена и свеобухватна уз критички осврт који је кандидат изложио у односу на актуелне литературне податке, указује на завидно познавање проучаване проблематике од стране кандидата, те значајно доприноси научној вредности ове дисертације.

Закључци дисертације су прецизни, концизни, научно засновани, јасно формулисани и приказани прегледно. Изнети закључци су у складу са постављеним циљевима и хипотезама докторске дисертације. Темеље се на резултатима добијеним истраживањем ове докторске дисертације, заснованом на адекватно изабраним и примењеним методама рада. Добијени резултати могу имати значајну вредност у препознавању повезаности генотипа и фенотипа хипертрофичне кардиомиопатије, а идентификација подтипова хипертрофичне кардиомиопатије представља помак ка прецизној медицини хипертрофичне кардиомиопатије.

Комисија сматра да су закључци јасно формулисани и прегледно приказани, сагласни са постављеним циљевима и хипотезама и у складу са примењеном методологијом рада и добијеним резултатима истраживања.

Литература - коришћена литература је обимна, актуелна и релевантна у односу на тему која се разматра, а библиографске јединице су цитиране Ванкуверским цитатним стилем на прописан начин.

Комисија сматра да су литературни наводи актуелни и адекватно одабрани за приказ истраживане проблематике, као и за поређење са добијеним резултатима истраживања.

На основу вредновања појединачних делова докторске дисертације,
Комисија позитивно оцењује све делове докторске дисертације.

VI СПИСАК НАУЧНИХ И СТРУЧНИХ РАДОВА КОЈИ СУ ОБЈАВЉЕНИ ИЛИ ПРИХВАЋЕНИ ЗА ОБЈАВЉИВАЊЕ НА ОСНОВУ РЕЗУЛТАТА ИСТРАЖИВАЊА У ОКВИРУ РАДА НА ДОКТОРСКОЈ ДИСЕРТАЦИЈИ:

Таксативно навести називе радова, где и када су објављени. Прво навести најмање један рад објављен или прихваћен за објављивање у складу са *Правилима докторских студија Универзитета у Новом Саду* који је повезан са садржајем докторске дисертације. У случају радова прихваћених за објављивање, таксативно навести називе радова, где и када ће бити објављени и приложити потврду уредника часописа о томе.

Glavaški M, Preveden A, Jakovljević Đ, Filipović N, Velicki L. Subtypes and mechanisms of hypertrophic cardiomyopathy proposed by machine learning algorithms. *Life (Basel)*. 2022;12(10):1566. (M22)

Glavaški M, Velicki L. Shared molecular mechanisms of hypertrophic cardiomyopathy and its clinical presentations: Automated molecular mechanisms extraction approach. *Life (Basel)*. 2021;11(8):785. (M22)

Glavaški M, Velicki L. Humans and machines in biomedical knowledge curation: Hypertrophic cardiomyopathy molecular mechanisms' representation. *BioData Min*. 2021;14(1):45. (M21)

Glavaški M, Velicki L. More slices, less truth: Effects of different test-set design strategies for magnetic resonance image classification. *Croat Med J*. 2022;63(4):370-8. (M22)

Glavaški M, Stankov K. Epigenetics in disease etiopathogenesis. *Genetika*. 2019;51(3):975-94. (M23)

Preveden A, Golubovic M, Bjelobrk M, Miljkovic T, Ilic A, Stojsic S, Gajic D, Glavaski M, Maier LS, Okwose N, Popovic D, Barlocco F, Ristic A, MacGowan GA, Olivotto I, Filipovic N, Jakovljevic DG, Velicki L. Gender related differences in the clinical presentation of hypertrophic cardiomyopathy—An analysis from the SILICOFM Database. *Medicina (Kaunas)*. 2022;58(2):314. (M22)

Glavaški M, Velicki L. HCM Map: Interactive knowledge resource about molecular mechanisms of hypertrophic cardiomyopathy. 5th Disease Maps Community Meeting (DMCM20), 12-14 November 2020, Conference abstract. (M34)

Glavaški M, Velicki L. Intersection—Simple, yet effective for exploring molecular mechanisms; Lessons learned from the HCM representations. 6th Disease Maps Community Meeting (DMCM2021), 29-30 November 2021, Conference abstract. (M34)

Glavaški M, Velicki L. HCM Clinical: An interactive knowledge resource for HCM diagnosis, monitoring, and treatment. 6th Disease Maps Community Meeting (DMCM2021), 29-30 November 2021, Conference abstract. (M34)

<p>VII ЗАКЉУЧЦИ ОДНОСНО РЕЗУЛТАТИ ИСТРАЖИВАЊА:</p>
<p>На основу добијених резултата истраживања, формулисани су јасни закључци који дају одговоре на постављене циљеве и хипотезе:</p> <ol style="list-style-type: none"> 1. Четири подтипа хипертрофичне кардиомиопатије идентификована су на основу свих доступних података о фенотипу. 2. Иако нити једна од одређених поставки субфенотипова не корелира у потпуности са генотипом, неке значајне корелације су пронађене у поставкама са 3, 4, 5 и 6 субфенотипова. 3. Предикција исхода фенотипа на основу информације о мутираним генима је могућа за присуство или одсуство синусног ритма и присуство или одсуство оштећења миокарда. 4. Алгоритми машинског учења извршили су предикцију следећих исхода на основу података о генотипу и фенотипу: замор, диспнеја, бол у грудима, палпитације, синкопа, шум на срцу, претибијални едем, покретање митралног залистка унапред, абнормалност папиларних мишића, хипокинезија, атријална фибрилација, атриовентрикуларни блок првог степена, блок леве гране, блок десне гране, предњи леви хемиблок, абнормалности СТ сегмента и негативни Т талас. 5. Идентификовани су генотип-специфични налази ехокардиограма. За мутацију у <i>MYH7</i> гену, структуре које највише утичу на распознавање су септум, излазни тракт леве коморе, предњи зид, врх срца, десна комора и митрални апарат; за мутацију у <i>TNNT2</i> гену, то су септум и десна комора; док су за мутацију у <i>MYBPC3</i> гену ове структуре септум, лева комора и шупљина леве коморе. <p>Комисија сматра да су формулисани закључци логично изведени из добијених резултата и јасно одражавају значај и научни допринос спроведеног истраживања.</p>
<p>VIII ОЦЕНА НАЧИНА ПРИКАЗА И ТУМАЧЕЊА РЕЗУЛТАТА ИСТРАЖИВАЊА:</p> <p>Експлицитно навести позитивну или негативну оцену начина приказа и тумачења резултата истраживања.</p>
<p>Мишљење Комисије је да су резултати истраживања у оквиру докторске дисертације Миле Главашки систематично, прегледно, јасно, детаљно и прикладно приказани и интерпретирани, као и критички продискутовани. Табеле и графикони доприносе прегледности и јасноћи резултата. Тумачење резултата се заснива на савременим и доступним научним сазнањима, те се начин приказа и тумачења резултата истраживања оцењује позитивно. Закључци су адекватни, научно оправдани и логично произилазе из добијених резултата рада.</p> <p>Тестирањем на плагијаризам помоћу програмског пакета iThenticate (https://www.ithenticate.com), утврђен је проценат преклапања од 27% (текст докторске дисертације на енглеском језику) и 4% (текст проширеног извода на српском језику). Увидом у извештај о тестирању на плагијаризам Комисија констатује да је кандидат адекватно цитирао коришћену литературу.</p> <p>Комисија позитивно оцењује начин приказа и тумачење резултата истраживања на основу закључка да је истраживање спроведено у складу са стандардима истраживања у области медицине, на репрезентативном узорку, да су добијени резултати овог истраживања оригинални, репрезентативни, прецизно приказани и адекватно интерпретирани, те да су научно оправдани и применљиви.</p>
<p>IX КОНАЧНА ОЦЕНА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:</p> <p>Експлицитно навести да ли дисертација јесте или није написана у складу са наведеним образложењем, као и да ли она садржи или не садржи све битне елементе. Дати јасне, прецизне и концизне одговоре на 3. и 4. питање:</p>
<p>1. Да ли је дисертација написана у складу са образложењем наведеним у пријави теме? Дисертација је написана у потпуности у складу са предложеним испитивањима и образложењем наведеним у пријави теме.</p>

<p>V Да ли дисертација садржи све битне елементе? Дисертација садржи све битне елементе, укључујући сва потребна поглавља и свеобухватне оригиналне резултате истраживања. Дисертација је написана концизно, разумљиво и резултат је самосталних истраживачких напора кандидата.</p>
<p>V По чему је дисертација оригиналан допринос науци? Докторска дисертација по својој свеобухватности, добијеним резултатима и изнетим закључцима представља оригинални допринос науци. Резултати овог истраживања омогућују даља истраживања у области кардиологије. Идентификована и описана су четири подтипа хипертрофичне кардиомиопатије. Пронађене су неке значајне корелације између генотипова и субфенотипова. Алгоритми машинског учења извршили су предикцију следећих исхода на основу података о генотипу и фенотипу: замор, диспнеја, бол у грудима, палпитације, синкопа, шум на срцу, претибијални едем, покретање митралног залистка унапред, абнормалност папиларних мишића, хипокинезија, атријална фибрилација, атриовентрикуларни блок првог степена, блок леве гране, блок десне гране, предњи леви хемиблок, абнормалности СТ сегмента и негативни Т талас. За сваки од ових исхода одређени су подскупови параметара који су довољни за вршење предикције и утврђен је њихов релативни значај у свакој од предикција. Идентификовани су генотип-специфични налази ехокардиограма (структуре које највише утичу на распознавање) за мутације у генима <i>MYH7</i>, <i>TNNT2</i> и <i>MYBPC3</i>. Добијени резултати докторске дисертације су јасни и примењиви у свакодневној медицинској пракси. Поред наведеног, добијени резултати отварају могућности за спровођење додатних истраживања у погледу остваривања поменутих циљева, а пре свега представљају корак ка прецизној медицини хипертрофичне кардиомиопатије који би у будућности могао да представља основу за стварање стратегија за превентиву и третман одређених група пацијената са дијагнозом хипертрофичне кардиомиопатије.</p>
<p>V Који су недостаци дисертације и какав је њихов утицај на резултат истраживања? Докторска дисертација не садржи формалне нити суштинске недостатке који би могли утицати на резултате истраживања и/или научну вредност докторске дисертације.</p>
<p>X ПРЕДЛОГ:</p>
<p>На основу наведеног, комисија предлаже:</p> <p>а) да се докторска дисертација прихвати, а кандидату одобри одбрана; б) да се докторска дисертација врати кандидату на дораду (да се допуни односно измени); в) да се докторска дисертација одбије.</p> <p>Докторска дисертација кандидата Миле Главашки написана је јасно и прегледно. Хомогеност текста показује да кандидат изврсно познаје проблематику као и способност кандидата да на оптималан начин користи расположиве литературне ресурсе. На основу позитивно вреднованих свих појединачних делова докторске дисертације, као и укупне позитивне оцене, Комисија за оцену докторске дисертације предлаже Наставно-научном већу Медицинског факултета у Новом Саду и Сенату Универзитета у Новом Саду да се докторска дисертација Миле Главашки под називом „Утврђивање повезаности генотипа и фенотипа хипертрофичне кардиомиопатије применом машинског учења“ прихвати, и кандидату одобри спровођење поступка јавне одбране.</p>

Место и датум: **Нови Сад, 09.05.2023.**

1. _____
Проф. др Александар Рецек, редовни професор, председник

2. _____
Проф. др Голуб Самарџија, ванредни професор, члан

3. _____
Проф. др Ненад Филиповић, редовни професор, члан

НАПОМЕНА: Члан комисије који не жели да потпише извештај јер се не слаже са мишљењем већине чланова комисије, дужан је да унесе у извештај образложење односно разлоге због којих не жели да потпише извештај и да исти потпише.

DOCTORAL DISERTATION EVALUATION REPORT

I COMMITTEE INFORMATION		
1. Date and body that assigned the committee: 26.04.2023. The Academic Council of the Faculty of Medicine, University of Novi Sad.		
2. Committee members information in accordance with regulations for doctoral studies at University Novi Sad:		
1. Prof. Dr. Aleksandar Redžek	Full Professor	Surgery with Anesthesiology (Cardiac surgery), 01.08.2020.
Surname and name	Title	Scientific field and date of election
Faculty of Medicine, University Novi Sad		Chair of the committee
Institution of employment		Committee member role
2. Prof. Dr. Golub Samardžija	Associate Professor	Pathology, 13.4.2022.
Surname and name	Title	Scientific field and date of election
Faculty of Medicine, University Novi Sad		Member of the committee
Institution of employment		Committee member role
3. Prof. Dr. Nenad Filipović	Full Professor	Applied mechanics and Applied informatics and computer engineering, 27.5.2010.
Surname and name	Title	Scientific field and date of election
Faculty of Engineering Sciences, University of Kragujevac		Member of the committee
Institution of employment		Committee member role
II CANDIDATE INFORMATION		
1. Name, name of one parent, surname: Mila, Sava, Glavaški		
2. Date, municipality and country of birth: 22.03.1985., Senta, Serbia		
3. Name of the academic institution, previously completed academic program and the title acquired: Faculty of Pharmacy, University of Belgrade, graduate studies, pharmacist-medical biochemist Faculty of Medicine and Faculty of Sciences, University of Novi Sad, master's academic studies, master of embryology		
4. Year of enrollment into doctoral studies and name of the PhD study program: 2016. Doctoral academic studies – Preclinical research.		

III TITLE OF THE DOCTORAL DISSERTATION:

Deciphering Genotype-Phenotype Associations in Hypertrophic Cardiomyopathy Using Machine Learning

IV OVERVIEW OF THE DOCTORAL DISSERTATION:

Concisely summarize the dissertation content with indicated number of pages, chapters, figures, images, charts etc.

The doctoral dissertation has been written systematically, clearly, and comprehensibly in a style that includes necessary elements of scientific research. It encompasses 428 pages and is organized into 7 chapters:

1. INTRODUCTION - written in 76 pages, including a total of 2 figures and 5 tables.
2. RESEARCH AIMS AND HYPOTHESES - presents the aims based on data presented in the introduction, and working hypotheses derived from the aims.
3. MATERIALS AND METHODS - this chapter contains a detailed description of the selection and sample size, protocols and research methods, and statistical analysis of data. The chapter is written in 24 pages, including 1 figure.
4. RESULTS - the results are systematically presented with textual descriptions. The chapter is written in 249 pages, including a total of 50 figures, 149 graphs, and 126 tables.
5. DISCUSSION - a detailed discussion of the obtained results and comparison with relevant published studies. The chapter is written in 11 pages.
6. CONCLUSION - derived from the obtained results of the dissertation, presented in 2 pages.
- 1-6. EXTENDED ABSTRACT IN SERBIAN LANGUAGE - includes all previously mentioned chapters. It is written in Serbian and spans 25 pages.
7. LITERATURE - 349 bibliographic units were cited in accordance with the Vancouver citation style. A distinct part of the doctoral dissertation is a list of all abbreviations used in the manuscript, which is given on 8 pages.

The dissertation presents 131 tables, 53 figures, and 149 graphs. Key documentation information in Serbian and English with abstracts is given at the beginning of the doctoral dissertation, and a data treatment plan is located at the end of the text.

V EVALUATION OF INDIVIDUAL DISSERTATION CHAPTERS:

The title of the doctoral dissertation “Deciphering Genotype-Phenotype Associations in Hypertrophic Cardiomyopathy Using Machine Learning“ is clearly and precisely formulated and entirely in line with the research content presented.

The committee considers that the title of the doctoral dissertation is clearly and precisely formulated and adequately reflects the content and topic of the doctoral dissertation.

The introduction of the doctoral dissertation is systematically written, with a detailed overview of current knowledge on the researched problem, and clearly indicates the complexity of the identified problem and the need for research. This chapter is divided into 8 smaller sections with sub-sections. The first section provides an overview of hypertrophic cardiomyopathy, and the second discusses cardiomyopathies in general. Since hypertrophic cardiomyopathy is often described as the "disease of the sarcomere," a detailed presentation of the sarcomere and its components follows. After that, genotype of hypertrophic cardiomyopathy is presented, followed by phenotype of hypertrophic cardiomyopathy, and then, the associations between the genotype and phenotype in hypertrophic cardiomyopathy are presented. After that, the importance of hypertrophic cardiomyopathy in contemporary cardiology and cardiac surgery is described. The last section of the introduction provides an overview of machine learning and its applicability in biomedicine.

In the introductory section, the candidate clearly, comprehensively, systematically, and using current and relevant literature sources, identified the complexity of the identified problem and justified the significance of conducting research.

The research aims are clearly defined and enable the drawing of specific conclusions. The research aims outlined in the dissertation are based on the results of previous studies published in relevant scientific publications. The objectives of the conducted research were as follows:

Aim 1. Identification of HCM subphenotypes using clustering

Aim 2. Correlation determination between genotypes and subphenotypes in HCM

Aim 3. Determination of associations of genotype and phenotype outcomes in HCM using ML algorithms

Aim 4. Creation of ML models for HCM outcomes prediction based on genotypic and phenotypic data

Aim 5. Identification of genotype-specific echocardiogram findings in HCM using ML algorithms

The research hypotheses are logical, clearly and precisely formulated, and in line with the research objectives:

Hypothesis 1. HCM subphenotypes can be identified using clustering.

Hypothesis 2. There is a correlation between genotypes and subphenotypes in HCM.

Hypothesis 3. The associations of genotype and phenotype outcomes in HCM can be determined using ML algorithms.

Hypothesis 4. ML models can predict HCM outcomes based on genotypic and phenotypic data.

Hypothesis 5. Genotype-specific echocardiogram findings in HCM can be identified using ML algorithms.

The committee considers that the research objectives and working hypotheses are logically and precisely formulated, clearly presented, scientifically justified, and enable the drawing of specific conclusions.

Materials and methods - In this chapter, the materials and methods used in the study are clearly and precisely described. The selection criteria and sample size of the study are defined precisely, as well as the criteria for inclusion and exclusion of participants in the study. The course of the study is described in detail, and the proposed work methods are clearly formulated. The research methodology is in accordance with the principles of scientific research in medicine and enables the reproducibility of the study. The methods used are appropriate for the type of study, accepted and used in the wider scientific community. The chosen methods correspond to the stated aims and enable the obtaining of specific and high-quality scientific results. The conduct of the research was approved by the decisions of the ethical committees of each institution participating in the research, as well as of the ethical committee of Faculty of Medicine, University Novi Sad. The study was multi-centric and retrospective, involved 143 adult patients with a confirmed diagnosis of hypertrophic cardiomyopathy, who met the inclusion and exclusion criteria. Demographics, medical and family history, anthropometric measurements were recorded; blood sampling, genetic testing, blood markers, transthoracic echocardiography with Doppler, cardiopulmonary exercise testing, electrocardiography (ECG) and ECG-holter-monitoring were all performed. Next, the collected data was prepared for analysis, mostly using the Pandas library in the Python programming language. It was analyzed using clustering techniques, employing the Scikit-learn module in Python. Models were created for predicting phenotypic outcomes using both the Scikit-learn and SHAP modules in Python. Additionally, a computer vision model was created for identifying genotype-specific echocardiogram findings, utilizing deep learning and computer vision libraries in the Python programming language.

The committee considers that the selected methods for data collection and statistical analysis are adequate, in line with the research objectives and hypotheses, and allow obtaining reliable, up-to-date, and high-quality scientific results. They are described clearly and in detail, thus enabling experiment reproducibility.

The results are presented convincingly, clearly, and professionally in tables and graphs. The displays are accompanied by clear textual explanations. The results are clearly derived from the applied methodology, using appropriate statistical methods. Hypertrophic cardiomyopathy subtypes were identified based on the overall phenotypic appearance. Some significant correlations between genotype and phenotype were found. Phenotypic outcomes were predicted based on mutated genes and using genotypic and phenotypic data. Genotype-specific echocardiogram findings were identified.

The committee considers that the results are original, comprehensive, and clear for interpretation, stemming from the applied methodology, and presented systematically, coherently, and in a way appropriate to the data type, with appropriate statistical methods for data processing.

The discussion is comprehensive and clearly and detailed written in a logical sequence. The candidate provides well-supported and critical analysis of the research results and compares them to citations from

relevant literature. The literature data is appropriately chosen, current, and relevant for drawing valid conclusions from the studied problem. The comments and interpretations of the results are exhaustive, logical, and scientifically grounded. The writing style is clear, precise, logical, and organized, and unambiguously indicates the candidate's knowledge of the researched problem. In the discussion text, the candidate demonstrates the ability to critically analyze and interpret scientific data, including the results of their own research, as well as data from relevant literature sources.

The committee considers that the discussion of the obtained research results is exhaustive, precise, scientifically grounded, and comprehensive, with a critical review presented by the candidate regarding current literature data, indicating an admirable knowledge of the researched topic by the candidate, and significantly contributing to the scientific value of this dissertation.

The conclusions of the dissertation are precise, concise, scientifically grounded, clearly formulated, and presented in an orderly manner. The conclusions are consistent with the set objectives and hypotheses of the doctoral dissertation, and they are based on the results obtained from the research carried out in this doctoral dissertation, which was based on adequately chosen and applied research methods. The obtained results can have significant value in recognizing genotype-phenotype associations in hypertrophic cardiomyopathy. Hypertrophic cardiomyopathy subtypes identification represents a shift towards precision medicine of hypertrophic cardiomyopathy.

The committee considers that the conclusions are clearly formulated and presented in a clear manner, consistent with the stated aims and hypotheses, and in accordance with the applied research methodology and obtained results.

The literature used is extensive, current, and relevant to the topic being discussed, and the bibliographic units are cited using the Vancouver citation style in the prescribed manner.

The committee considers that the literary references are current and adequately selected to present the researched issues, as well as to compare with the obtained research results.

On the basis of the evaluation of individual parts of the doctoral dissertation,
the committee positively assesses all parts of the doctoral dissertation.

VI LIST OF SCIENTIFIC AND PROFESSIONAL MANUSCRIPTS PUBLISHED OR ACCEPTED FOR PUBLICATION, WHICH ARE BASED ON THE RESULTS OF THE DOCTORAL RESEARCH:

Glavaški M, Preveden A, Jakovljević Đ, Filipović N, Velicki L. Subtypes and mechanisms of hypertrophic cardiomyopathy proposed by machine learning algorithms. *Life (Basel)*. 2022;12(10):1566. (M22)

Glavaški M, Velicki L. Shared molecular mechanisms of hypertrophic cardiomyopathy and its clinical presentations: Automated molecular mechanisms extraction approach. *Life (Basel)*. 2021;11(8):785. (M22)

Glavaški M, Velicki L. Humans and machines in biomedical knowledge curation: Hypertrophic cardiomyopathy molecular mechanisms' representation. *BioData Min*. 2021;14(1):45. (M21)

Glavaški M, Velicki L. More slices, less truth: Effects of different test-set design strategies for magnetic resonance image classification. *Croat Med J*. 2022;63(4):370-8. (M22)

Glavaški M, Stankov K. Epigenetics in disease etiopathogenesis. *Genetika*. 2019;51(3):975-94. (M23)

Preveden A, Golubovic M, Bjelobrck M, Miljkovic T, Ilic A, Stojisic S, Gajic D, Glavaski M, Maier LS, Okwose N, Popovic D, Barlocco F, Ristic A, MacGowan GA, Olivotto I, Filipovic N, Jakovljevic DG, Velicki L. Gender related differences in the clinical presentation of hypertrophic cardiomyopathy—An analysis from the SILICOFM Database. *Medicina (Kaunas)*. 2022;58(2):314. (M22)

Glavaški M, Velicki L. HCM Map: Interactive knowledge resource about molecular mechanisms of hypertrophic cardiomyopathy. 5th Disease Maps Community Meeting (DMCM20), 12-14 November 2020, Conference abstract. (M34)

Glavaški M, Velicki L. Intersection—Simple, yet effective for exploring molecular mechanisms; Lessons learned from the HCM representations. 6th Disease Maps Community Meeting (DMCM2021), 29-30 November 2021, Conference abstract. (M34)

Glavaški M, Velicki L. HCM Clinical: An interactive knowledge resource for HCM diagnosis, monitoring, and treatment. 6th Disease Maps Community Meeting (DMCM2021), 29-30 November 2021, Conference abstract. (M34)

VII CONCLUSIONS AND RESEARCH RESULTS:

Based on the obtained research results, clear conclusions were formulated that provide answers to the set goals and hypotheses:

1. Four hypertrophic cardiomyopathy subtypes were identified based on the overall phenotypic appearance
2. Although none of the determined subtypes settings could be completely correlated with genotype, some significant correlations with genotype were found in the 3-subtypes, 4-subtypes, 5-subtypes, and 6-subtypes settings.
3. Two phenotypic outcomes that can be predicted from mutated genes are the absence or presence of sinus rhythm and the absence or presence of myocardial injury
4. The ML algorithms were able to predict phenotypic outcomes — fatigue, dyspnea, chest pain, palpitations, syncope, heart murmur, pretibial edema, systolic anterior motion, papillary muscle abnormalities, hypokinesia, atrial fibrillation, atrioventricular block I, left bundle branch block, right bundle branch block, left anterior hemiblock, ST segment abnormalities, and negative T wave — using genotypic and phenotypic data.
5. Genotype-specific echocardiogram findings were identified: for mutations in the *MYH7* gene, the most discriminative structures are the septum, left ventricular outflow tract, anterior wall, apex, right ventricle, and mitral apparatus; for mutations in the *TNNT2* gene, the most discriminative structures are septum and right ventricle; while for mutations in *MYBPC3* gene these are septum, left ventricle, and left ventricle chamber.

The committee considers that the formulated conclusions are logically derived from the obtained results and clearly reflect the significance and scientific contribution of the conducted research.

VIII EVALUATION OF STUDY RESULTS PRESENTATION AND INTERPRETATION:

Explicitly indicate positive or negative assessment of study results presentation and interpretation.

The committee considers that the research results presented in Mile Glavaški's doctoral dissertation are systematically, clearly, thoroughly, and appropriately demonstrated and interpreted, as well as critically discussed. Tables and graphs contribute to the clarity and comprehensibility of the results. The interpretation of the results is based on contemporary and available scientific knowledge, and the way the results are presented and interpreted is positively evaluated. The conclusions are adequate, scientifically justified, and logically derived from the obtained results of the study.

After testing for plagiarism using the iThenticate software package (<https://www.ithenticate.com>), the overlap percentages were determined to be 27% (for the doctoral dissertation text in English) and 4% (for the extended abstract text in Serbian). Upon review of the plagiarism testing report, the committee has determined that the candidate has adequately cited the used literature.

The committee positively assesses the presentation and interpretation of the research results based on the conclusion that the research was conducted in accordance with the standards of medical research, on a representative sample, that the results obtained from this research are original, representative, precisely presented, and adequately interpreted, and that they are scientifically justified and applicable.

IX FINAL ASSESSMENT OF DOCTORAL DISSERTATION:	
Explicitly state if the dissertation is or is not written in accordance with submitted dissertation proposal, as well as does it contain all the important elements. Provide clear, precise and concise answers to questions 3 and 4:	
1. Was the dissertation written in accordance with the previously submitted dissertation proposal? Dissertation is written entirely in accordance with previously submitted dissertation proposal.	
V	Does the dissertation contain all the important elements? The dissertation contains all essential elements, including all necessary chapters and comprehensive original research results. The dissertation is written concisely and clearly, and the results are the independent research efforts of the candidate.
V	Why does this dissertation provide original contributions to science? The doctoral dissertation, in terms of its comprehensiveness, obtained results, and presented conclusions, represents an original contribution to science. The results of this research enable further exploration in the field of cardiology. Four hypertrophic cardiomyopathy subtypes were identified and described based on the overall phenotypic appearance. Some significant correlations between genotype and phenotype in hypertrophic cardiomyopathy were found. Machine learning algorithms predicted phenotypic outcomes — fatigue, dyspnea, chest pain, palpitations, syncope, heart murmur, pretibial edema, systolic anterior motion, papillary muscle abnormalities, hypokinesia, atrial fibrillation, atrioventricular block I, left bundle branch block, right bundle branch block, left anterior hemiblock, ST segment abnormalities, and negative T wave — using genotypic and phenotypic data. Subsets of features that are sufficient for prediction of these phenotypic outcomes were identified and their relative importance was assessed. Genotype-specific echocardiogram findings were identified for mutations in the <i>MYH7</i> , <i>TNNT2</i> , and <i>MYBPC3</i> gene. The results obtained from the doctoral dissertation are clear and applicable in everyday medical practice. In addition, the obtained results open up possibilities for conducting further research in order to achieve the mentioned goals, and above all they represent shift towards precision medicine of hypertrophic cardiomyopathy that could serve as a basis for creating strategies for prevention and treatment of specific hypertrophic cardiomyopathy patient groups.
V	What are the shortcomings of the dissertation and what is their impact on the research results? The doctoral dissertation does not contain any formal or substantive shortcomings that could affect the research results and/or the scientific value of the doctoral dissertation.
X PROPOSAL:	
Based on the indicated information, committee is proposing:	
<ul style="list-style-type: none"> a) To accept the doctoral dissertation and approve the candidate's defense; b) To return the doctoral dissertation to the candidate for revisions (to supplement or modify); c) To reject the doctoral dissertation. <p>The Mila Glavaški's doctoral dissertation is written clearly and concisely. The homogeneity of the text indicates that the candidate has an excellent understanding of the issues and is able to make optimal use of available literary resources. Based on the positive evaluation of all individual parts of the doctoral dissertation, as well as the overall positive assessment, the Committee for the Evaluation of the Doctoral Dissertation recommends to the Academic Council of the Faculty of Medicine in Novi Sad and the Senate of the University in Novi Sad to accept the Mila Glavaški's doctoral dissertation titled “Deciphering Genotype-Phenotype Associations in Hypertrophic Cardiomyopathy Using Machine Learning” and approve the public defense procedure.</p>	

Novi Sad, **09.05.2023.**

1. _____
Prof. Dr. Aleksandar Redžek, Full
Professor, Chair of the committee

2. _____
Prof. Dr. Golub Samardžija, Associate
Professor, Member of the committee

3. _____
Prof. Dr. Nenad Filipović, Full Professor,
Member of the committee

NOTE: A committee member who does not want to sign the report because they disagree with the majority opinion of the committee is obliged to provide an explanation or reasons why they do not want to sign the report and to sign it accordingly.