

Plan upravljanja istraživačkim podacima projekta "Otkrivanje biomarkera za bolest Hashimotov tireoiditis koristeći multi-omics pristup"

Boraska Perica, Vesna

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2025**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:171:102623>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-04**



Repository / Repozitorij:

[MEFST Repository](#)



PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)

Opće informacije		
	Ime i prezime predlagatelja	Vesna Boraska Perica
	Matična organizacija	Medicinski fakultet Sveučilišta u Splitu
	Naziv projekta	Otkrivanje biomarkera za bolest Hashimotov tireoiditis koristeći multi-omics pristup
	Upravitelj podacima	Vesna Boraska Perica, vboraska@mefst.hr
1.	Prikupljanje podataka i dokumentacija	
	<p>Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite formate, vrste i opseg svih podataka s kojima ćete raditi, a ne samo krajnji skup podataka koji će biti rezultat istraživanja)</p>	<p>Tijekom projekta prikupit ćemo i ponovo koristiti nekoliko tipova podataka (iz postojeće CRO-HT biobanke bioloških uzoraka i fenotipova ispitanika s HT-om i kontrolnih ispitanika):</p> <ul style="list-style-type: none"> - podaci iz prikupljenih upitnika (sociodemografski podaci, podaci vezani uz reproduktivni život, podaci vezani uz rani razvoj, podaci vezani za okoliš, pušenje, fizičku aktivnost, prehranu, strane objekte u tijelu te stolicu). Podaci iz upitnika će se konvertirati u digitalni oblik u Excel tablicu i neće zauzimati više od 3MB memorijskog prostora). - medicinski podaci iz prikupljenih upitnika (antropometrijski podaci, anamnestički podaci – terapija, simptomatologija, klinički nalaz, povijest bolesti te podaci o izgledu i veličini štitnjače prikupljeni ultrazvukom). Podaci iz upitnika će se konvertirati u digitalni oblik u Excel tablicu i neće zauzimati više od 3MB memorijskog prostora). - uzorci plazme, seruma i DNA. Centrifugiranjem krvi će se odvojiti plazma i serum te pohraniti u zamrzivač - 80°C, a iz pune krvi ispitanika će se izolirati DNA. Iz izolirane DNA će se vršiti genotipizacija čitavog genoma te ćemo dobiti genotipske podatke. - genotipski podaci – Genotipovi (bialelne varijante) i obiteljski podaci će biti spremljeni u datotekama u .bim i .fam obliku. Prva tri bajta datoteke trebaju biti u redosljedu 0x6c, 0x1b i 0x01, a ostatak datoteke čine nizovi V blokova od N/4 (zaokruženih) bajtova svaki, gdje je V broj varijanti, a N broj uzoraka. Prvi blok odgovara prvom genetskom markeru u .bim datoteci itd. Dva bita prvog bajta pohranjuju genotip prvog uzorka (prvi uzorak se smatra onaj koji je prvi naveden u pratećoj .fam datoteci). Sljedeća dva bita pohranjuju genotip drugog uzorka i tako dalje za 3. i 4. uzorak. Drugi bajt pohranjuje kodove genotipa za uzorke 5.-8., treći bajt pohranjuje kodove za 9.-12., itd. Genotipski podaci će se pohraniti na računalnu platformu visokih performansi (poslužitelj) koji se planira nabaviti kroz projekt. Poslužitelj će sadržavati veći broj jezgri (32) i veći omjer radne memorije po jezgri, 16GB po logičkom procesoru tj. sveukupno 512GB RAM. Projekt će zahtijevati minimalno 7TB efektivnog kapaciteta diskovnog prostora zbog veće količine analiza. - iz plazme i seruma će se vršiti mjerenja proteina, metabolita i zonulina. Proteini će se mjeriti tzv. „proximity extension assays“ (PEA) tehnologijom. Rezultati mjerenja proteina će biti izraženi tzv. Normalized Protein

	<p>eXpression (NPX) vrijednošću koja je derivirana iz koncentracije proteina (pg/mL). NPX vrijednost je proizvoljna jedinica Olink Proteomics (Uppsala, Sweden) koja se izražava u Log2 formi. Izračunava se iz vrijednosti praga ciklusa te se vrši prethodna normalizacija podataka kako bi se smanjile varijacije unutar i između testa. NPX podaci omogućuju korisnicima da identificiraju promjene u razinama pojedinačnih proteina u skupu uzoraka, a zatim koriste te podatke za utvrđivanje karakterističnih proteinskih obrazaca. Metaboliti će se mjeriti spektroskopijom pomoću nuklearne magnetske rezonancije (NMR). Ovom tehnologijom se može kvantificirati 250 biomarkera izraženim u jedinicama apsolutne koncentracije mmol/L ili u omjerima. Koncentracija zonulina (mmol/L) će se mjeriti ELISA metodom. Rezultati svih mjerenja će se pohraniti u digitalnom obliku u Excel tablici i neće zauzimati više od 3MB memorijskog prostora.</p>
<p>Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete, načine organiziranja podataka te alate i instrumente kojima ćete se koristiti za prikupljanje i obradu)</p>	<ul style="list-style-type: none"> - koristit ćemo gore opisane podatke od već prikupljenih 700 ispitanika (500 HT i 200 kontrola) te ćemo kroz ovaj projekt prikupiti 200 novih ispitanika kojima će biti podijeljeni upitnici u papirnatost formi na Kliničkom zavodu za nuklearnu medicinu Kliničkog bolničkog centra u Splitu. Ispunjeni upitnici u papirnatost formi će se, sukladno načelima povjerljivosti, držati zaključanim na Katedri za medicinsku biologiju, Medicinskog fakulteta Sveučilišta u Splitu u sklopu CRO-HT, Biobanke ispitanika oboljelih od Hashimotovog tireoiditisa. - koristit ćemo medicinske podatke od već prikupljenih 700 ispitanika (500 HT i 200 kontrola) te ćemo prikupiti medicinske podatke 200 novih ispitanika. Upitnici vezani uz medicinske podatke će biti podijeljeni ispitanicima u papirnatost formi te ispunjeni u razgovoru s liječnikom ili izmjereni od strane liječnika na Kliničkom zavodu za nuklearnu medicinu Kliničkog bolničkog centra u Splitu. Ispunjeni upitnici s medicinskim podaci će se, sukladno načelima povjerljivosti, držati zaključanim na Katedri za medicinsku biologiju, Medicinskog fakulteta Sveučilišta u Splitu u sklopu CRO-HT Biobanke ispitanika oboljelih od Hashimotovog tireoiditisa. - koristit ćemo informacije o genotipovima u već genotipiziranih 405 HT ispitanika iz postojeće CRO-HT biobanke HT ispitanika, a iz novo prikupljenih 200 uzoraka će se izolirati DNA, na Katedri za medicinsku biologiju, Medicinskog fakulteta Sveučilišta u Splitu. Četiristo DNA uzoraka (200 novo izoliranih DNA i 200 već izoliranih DNA od kontrolnih ispitanika) će se poslati u genotipizacijsku kompaniju gdje će se napraviti genotipizacija cjelokupnog genoma koristeći Illumina genotipizacijske čipove. Genotipizacijski uzorci će proći dodatnu kontrolu kvalitete koja će se vršiti na računalnoj platformi visokih performansi koristeći slobodno dostupne programe za statističku analizu podataka (PLINK, IMPUTE, R). <p>Iz plazme i seruma će se mjeriti koncentracije serumskih proteina, metabolita i zonulina na gore opisani način, a analizirati će se također na računalnoj platformi visokih performansi.</p> <p>Dosljednost mjerljivih podataka procijenit će se usporedbom ponovljenih mjerenja. Osnovi rezultat analiza podataka, koji mogu biti od interesa za širu znanstvenu zajednicu, će se pohraniti u institucijski repozitorij. Svi podaci će biti pohranjeni u datotekama koje će se nazvati prema prije dogovorenim normama. Skup podataka pratit će i README dokumenti u kojima će se opisati hijerarhija direktorija. Svaki direktorij sadržavat će i INFO.txt datoteku u kojoj će se opisati objašnjenja i korisne informacije. Povremeno će se</p>

		<p>provести kontrole kvalitete procesa da bi se uklonili pogreške i nepotrebni podatci. Podaci će prilikom unosa u digitalni oblik biti dodatno provjereni i analizirani koristeći odgovarajuću statističku metodu.</p>
	<p>Koju ćete dokumentaciju i metapodatke izraditi osim podataka? (dokumentacija mora sadržavati informacije i standarde potrebne korisnicima kako bi mogli samostalno čitati i interpretirati podatke u budućnosti, primjerice, kodne knjige, <i>ReadMe</i> datoteke i sl.)</p>	<p>Svaki dobiveni podatak će se spremirati sa svim dostupnim metapodacima. Dokumenti i datoteke će se imenovati standardnim nazivljem s popratnom dokumentacijom s objašnjenjima, kako bi se ostalim korisnicima omogućilo samostalno pronalaženje i korištenje podataka. Preciznije, završni skup podataka tj. sumarni rezultati pohranit će se u odabranom repozitoriju, popraćen s README dokumentom sa sadržajem svih datoteka kao i koja je konvencija korištena za njihovo imenovanje.</p>
2.	Pravna i sigurnosna pitanja	
	<p>Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci obrađuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka, navesti metode anonimizacije podataka)?</p>	<p>Ovaj projektni prijedlog uključuje istraživanja koja uključuje klinički pregled ispitanika, prikupljanje njihovih osobnih i medicinskih podataka, kao i provedbu dijela istraživanja na humanim biološkim uzorcima (DNA/serum/plazma). Prije samog sudjelovanja u istraživanju, ispitanicima će se dostaviti informirana privola (engl. Informed Consent) u papirnoj formi koju potpisuju te na taj način daju dozvolu za obradu i objavu podataka prikupljenih u svrhu istraživanja. U privoli će ispitanici biti upoznati sa svrhom istraživanja, načinima prikupljanja i analize podataka, oblicima diseminacije i objave podataka te o upravljanju podacima te potencijalnoj budućoj upotrebi nakon završetka projekta. Biti će navedeno i osiguranje zaštite podataka, kao i pravo ispitanika na povlačenje iz istraživanja. Potpisane informirane privole ispitanika će se držati zaključane na Katedri za medicinsku biologiju, Medicinskog fakulteta Sveučilišta u Splitu u sklopu CRO-HT Biobanke ispitanika oboljelih od Hashimotovog tireoiditisa. Osobni i osjetljivi podaci će biti anonimizirani kako bi se zaštitio identitet ispitanika. Istraživanje je odobreno od strane Etičkog povjerenstva Medicinskog fakulteta u Splitu te Etičkog povjerenstva KBC-a Split.</p> <p>Dio istraživanja koji uključuje pacijente usklađen je s odredbama o zaštiti prava i osobnih podataka ispitanika iz Zakona o zaštiti prava pacijenta (NN169/04, 37/08) i Zakona o provedbi Opće uredbe o zaštiti podataka (NN 42/18), te odredbama Kodeksa liječničke etike i deontologije (NN55/08, 139/15) i pravilima Helsinške deklaracije WMA 1964-2013 na koje upućuje Kodeks. U projektu će se koristiti i obrađivati podaci samo onih osoba koje su dale pisanu informiranu privolu.</p>

<p>Kako će se regulirati pristup podacima i njihova sigurnost? Koji su potencijalni rizici koje treba uzeti u obzir? Kako ćete osigurati sigurnost pohrane osjetljivih podataka?</p>	<p>Osobnim podacima ispitanika imat će pristup samo voditelj projekta i liječnik koji sudjeluje u provedbi istraživanja koji će svakom ispitaniku u studiji dodijeliti jedinstveni kod. Dokument koji povezuje ime ispitanika s kodom će biti zaštićen lozinkom koja će samo biti poznata glavnom istraživaču. Svi ostali suradnici na projektu koristit će se kodovima i neće imati uvida u osobne podatke ispitanika. Glavni istraživač i svi njegovi suradnici pridržavat će se interne procedure za zaštitu osobnih podataka. Ime ispitanika nikada neće biti otkriveno. Pristup dokumentaciji mogu imati i predstavnici Etičkog povjerenstva Medicinskog fakulteta. Podaci će se pohraniti na računalnu platformu visokih performansi, kojima će pristup imati samo suradnici s projekta. Svaki suradnik na projektu će dobiti svoju korisničku oznaku i zaporku za pristup poslužitelju.</p>
<p>Kako ćete upravljati zaštitom autorskih prava i drugog intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?</p>	<p>Podaci koji su prikladni za dodatno korištenje u znanstvene svrhe (sumarni rezultati provedenih analiza u kojima neće biti osobnih podataka) će se moći dijeliti. Zbog toga, naš cilj je da omogućimo najširu moguću uporabu podataka i objavit ćemo ih pod Creative Commons CC0 prilikom objave znanstvenih radova. Osobni podaci se nikada neće dijeliti. Podaci o prikupljenim medicinskim podacima, genotipovima, mjerama proteina, metabolita i zonulina će se također moći dijeliti kroz znanstvene suradnje, ali bez mogućnosti da se identificiraju pojedinci.</p>
<p>3. Pohrana i čuvanje podataka</p>	
<p>Kako će radne verzije podataka biti pohranjene tijekom projekta? Kako će se napraviti sigurnosne kopije tih podataka (<i>backup</i>)? Koja je očekivana količina podataka koja će se prikupiti i čuvati tijekom projekta (izraženo u MB/GB/TB)?</p>	<p>Uz pohranu podataka na računalnoj platformi visokih performansi, podatke ćemo pohraniti i na drugom poslužitelju (back-up uređaju) kojeg također planiramo nabaviti kroz projekt, a koji će se koristiti za potrebe izrade sigurnosnih kopija rezultata analiza sa minimalno dva puta većim diskovnim efektivnim prostorom od 16TB, što je očekivana količina podataka i rezultata analize koji se planiraju generirati kroz ovaj projekt. U oba poslužitelja biti će ugrađene 10Gbit mrežne kartice za potrebe brzog prijenosa sigurnosnih kopija rezultata analiza. Svi osnovni sumarni rezultati pohranit će se na institucionalnom repozitoriju koji redovito izrađuje sigurnosne kopije pomoću RSYNC alata.</p>
<p>Kako će se završne verzije podataka dugotrajno pohraniti i čuvati (i nakon završetka projekta)? U kojim će se formatima čuvati podaci? Koja je očekivana količina podataka koja će se trajno pohraniti (izraženo u MB/GB/TB)?</p>	<p>Podaci će se čuvati najmanje tri godine nakon završetka projekta. Podatke ćemo čuvati na računalnoj platformi visokih performansi, a rezultate istraživanja ćemo pohraniti u institucijski repozitorij Medicinskog fakulteta, Sveučilišta u Splitu, Dabar (https://repozitorij.mefst.unist.hr/). Tablične podatke čuvat ćemo u binarnim fileovima, excel obliku, a tekstualne u DOCX, TXT ili PDF obliku. Očekivana količina podataka koja će se pohraniti zahtjeva oko 16TB diskovnog prostora.</p>

4. Dijeljenje i ponovna uporaba podataka	
Kako i gdje će se podaci dijeliti? Koji repozitorij će se koristiti za dijeljenje podataka? Kako će potencijalni korisnici doznati za podatke?	Konačnu verziju rezultata istraživanja voditelj projekta podijelit će putem objavljenih znanstvenih radova i u sustavu Dabar, a nakon publiciranja istih. Institucijski repozitorij u sustavu Dabar odabrali smo jer podržava FAIR principe: skupovima dodjeljuje trajni identifikator URN:NBN, osigurava vidljivost podataka putem OpenAIRE portala i Google Scholar te tražilice dabar.srce.hr, a ujedno doprinosi vidljivosti i transparentnosti rada (Medicinski fakultet Sveučilišta u Splitu).
Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Podaci neophodni za bilo koju publikaciju bit će dostupni u trenutku objavljivanja, budući da dijeljenje podataka na bilo koji način prije objavljivanja rada/publikacije, ugrožava naš znanstveni rad. Svi neobjavljeni podaci pohranit će se u repozitoriju na najmanje 12 mjeseci od završetka projekta. Svi osobni podaci bit će anonimni prije distribucije na temelju preporuke Zakona o zaštiti osobnih podataka.
Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR-a</i> .	Da
Potvrdite da ćete se koristiti digitalnim repozitorijem koji održava neprofitna organizacija (ako ne, objasnite zašto ne možete dijeliti podatke na digitalnom repozitoriju koji nije komercijalan).	Da