

Zemunik, Tatijana

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2022**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:171:939868>

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

Download date / Datum preuzimanja: **2024-12-26**



Repository / Repozitorij:

[MEFST Repository](#)

Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Prof. dr. sc. Tatijana Zemunik, dr. med.
	Matična organizacija	Medicinski fakultet Sveučilišta u Splitu
	Naziv projekta	Regulacija funkcije štitne i doštitne žlijezde i homeostaze kalcija u krvi (IP 2593)
	Upravitelj podacima	Nikolina Pleić, mag. math., doktorand, npleic@mefst.hr
1.	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	<p>Tijekom projekta ćemo raspolagati s četiri tipa podataka:</p> <p>Već ranije prikupljeni podaci</p> <ol style="list-style-type: none"> 1. Genotipski podaci ispitanika određeni pomoću platformi za genotipizaciju od Illumine 2. Različiti fenotipski podaci ispitanika prikupljeni pomoću anketa i drugih fenotipskih mjerenja <p>Podaci koji će se prikupljati tijekom provedbe projekta</p> <ol style="list-style-type: none"> 3. Različiti parametri kojima se određuje funkcija štitne i doštitne žlijezde te homeostaze kalcija u krvi mjereni u plazmi ispitanika (vrijednosti TSH, hormona štitnjače, tireoglobulina, antitijela na tireoglobulin (TgAt) i tiroidnu peroksidazu (TPOAt), paratiroidnog hormona, kalcitonina i osteokalcina) 4. Slike sa fluorescencijskog mikroskopa na obojanim uzorcima ljudskog tkiva štitnjače <p>Podaci u 1. kategoriji biti će pohranjeni u VCF formatu (.vcf). Za podatke iz 1. kategorije će biti potrebno 50 GB prostora.</p> <p>Podaci iz 2. kategorije će se upisivati ručno iz anketa u <i>excel</i> (.xlsx) file. Za podatke iz 2. kategorije će biti potrebno 50 MB prostora.</p> <p>Podaci u 3. kategoriji biti će pohranjeni u <i>excel</i> (.xlsx) obliku. Za podatke iz 3. kategorije će biti potrebno 50 MB prostora.</p> <p>Podaci u 4. kategoriji biti će pohranjeni u <i>TIFF</i> (.tiff) obliku. Za podatke iz 4. kategorije će biti potrebno 5-10 GB prostora.</p>
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	<p>Različiti mjereni biokemijski parametri (vrijednosti TSH, hormona štitnjače, tireoglobulina, antitijela na tireoglobulin i tiroidnu peroksidazu, paratiroidnog hormona, kalcitonina i osteokalcina mjereni u plazmi ispitanika te vrijednosti ekspresije različitih proteina na tkivu štitnjače) će se upisivati u <i>excel</i> (.xlsx) tablice (prepisivanjem iz ručno pisanih zapisa – laboratorijskih dnevnika). Podatke će u tablicu unositi jedan istraživač, dok će, nakon unosa dva neovisna istraživača provjeravati unos podataka. Tijekom mjerenja biokemijskih parametara će se provoditi kontrola kvalitete (usporedbom dobivenih rezultata ponovljenim mjerenjima). Mjerni instrumenti će se redovito kalibrirati i servisirati. Sa svakim biokemijskim mjerenjem pohranit će se nekoliko metapodataka (mjerna jedinica, esej korišten za analizu, razrjeđenje uzorka, vrijeme inkubacije, itd.). Dodatno će se i sa svakom mikroskopskom slikom pohraniti nekoliko metapodataka (povećanje, veličina polja, tip</p>

		<p>analiziranog tkiva, kontrola/patološko tkivo, itd.). Fenotipski podaci neće biti memorijski zahtjevni, stoga će se moći obrađivati na standardnom <i>desktop</i> sučelju računala. Za obradu podataka koristit ćemo programe Excel, R te ostale programe koji budu potrebni u organiziranju i obradi podataka.</p> <p>Genotipski podaci će biti pohranjeni u <i>VCF</i> formatu na serveru istraživačke skupine voditelja projekta za koji pristup ima voditelj projekta i istraživači kojima je odobren pristup. Obzirom da su genotipski podaci iznimno memorijski zahtjevni i nije ih moguće otvoriti niti obrađivati na standardnom <i>desktop</i> sučelju računala, bit će organizirani unutar direktorija na serveru istraživačke skupine te će se kroz platformu Unix obrađivati programskim alatima specijalno prilagođenima takvoj vrsti podataka.</p>
	<p>Koju ćete dokumentaciju i metapodatke ustupiti osim podataka? (navedite koje su informacije potrebne korisnicima kako bi mogli čitati i interpretirati podatke u budućnosti te koji će se standardi koristiti pri tumačenju podataka)</p>	<p>Podaci istraživačkog projekta bit će popraćeni metapodacima tj. dokumentacijom s objašnjenjima. Metapodaci će biti zapisani u tekstualnu <i>readme</i> datoteku (.txt) i pridruženi pojedinom skupu podataka u vršnom direktoriju. Sa svakim biokemijskim mjerenjem pohranit će se nekoliko metapodataka (mjerna jedinica, esej korišten za analizu, razrjeđenje uzorka, vrijeme inkubacije, itd.). Dodatno će se i sa svakom mikroskopskom slikom pohraniti nekoliko metapodataka (povećanje, veličina polja, tip analiziranog tkiva, kontrola/patološko tkivo, itd.). Svi fenotipski podaci će biti popraćeni metapodacima koji će sadržavati rječnik o istraživačkim podacima tj. objašnjenja varijabli, mjernih jedinica, raspona vrijednosti i dopuštenih vrijednosti, vrijednosti koje nedostaju, metoda i slično.</p> <p>Metapodaci će biti ažurirani tijekom trajanja istraživačkog projekta kako bi podaci i izmjene na podacima bili čitljivi, razumljivi i pogodni za daljnju upotrebu. Izmjene na podacima bit će dokumentirane opisno i datumom izmjene unutar metapodataka, a podaci će biti verzionirani s datumom izmjene u imenu datoteke.</p> <p>Genotipski podaci će biti pohranjeni u <i>VCF</i> formatu koji je poseban po tome što se zaglavlje svake <i>VCF</i> datoteke sastoji od 100-150 redaka metapodataka. Svaka promjena na radnim genotipskim podacima je automatski evidentirana unutar zaglavlja novog <i>VCF</i> filea što uvelike olakšava upravljanje genotipskim podacima. Direktoriji unutar servera istraživačke skupine će, osim radnih genotipskih podataka, sadržavati i <i>readme</i> datoteke (.txt) u kojima će biti zabilježena metodologija korištena pri dobivanju rezultata genomskih analiza.</p> <p>Završni skup podataka pohranit će se na serveru istraživačke skupine voditelja projekta, popraćen s <i>README</i> dokumentom. Metapodaci će se temeljiti na općenitoj shemi koju koristi <i>Zenodo</i> (https://zenodo.org/record/1914401#.Ykre8ChByUk):</p> <p>Primjer tema u <i>README</i> dokumentu:</p> <ul style="list-style-type: none"> • NASLOV: Ime podatka ili znanstvenog projekta. • STVARATELJ: Ime i prezime istraživača koji je kreirao podatak. • OPIS sadržaja podataka i ostale popratne dokumentacije. • METODOLOGIJA: Opis kako su prikupljeni podaci. • PRAVO PRISTUPA • INFORMACIJE O PODACIMA:

		<ul style="list-style-type: none"> - Identifikator - Lokacija - Datumi - Tema - Format
2.	Pravna i sigurnosna pitanja	
	Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci pohranjuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka)?	Svi ispitanici uključeni u ovo istraživanje potječu iz projekta „10 001 Dalmatinac“ i već su dali informirani pristanak za sudjelovanje u istraživanju. Ispitanici su popunili ankete u pisanom obliku. Prilikom upisivanja podataka iz anketa u elektronski oblik (excel format), podaci su anonimizirani (svakom ispitaniku je dodijeljena šifra). Podaci o identitetu sudionika nisu pohranjeni u elektronskom obliku, a čuva ih voditelj projekta „10 001 Dalmatinac“ u zaključanom ormaru u zaključanoj sobi. Identifikacijski podaci nisu dostupni ni voditelju ni suradnicima projekta „Regulacija funkcije štitne i doštitne žlijezde i homeostaze kalcija u krvi“. Projekt poštuje sva ograničenja i zahtjeve kako je utvrđeno Zakonom o provedbi Opće uredbe o zaštiti podataka NN 42/18. Rezultati koji će se dobiti tijekom istraživačkog projekta su sumarni, nisu vezani za pojedinačne sudionike, te se objavom rezultata (u obliku publikacija), ispitanici neće moći identificirati. Ovo istraživanje se provodi u skladu s Helšinskom deklaracijom te je odobreno od Etičkog odbora Medicinskog fakulteta Sveučilišta u Splitu (Br.: 2181-198-03-04-14-0031 i 2181-198-03-04- 19-0022).
	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?	Podaci istraživačkog projekta će se čuvati na 3 lokacije; serveru istraživačke skupine, računalu odgovorne osobe, te na prijenosnom disku. Server je smješten u Informatičkoj službi Medicinskog fakulteta Sveučilišta u Splitu, a prijenosni disk na drugoj geografskoj lokaciji u zaključanom ormaru u zaključanoj sobi. Računalo je smješteno u zaključanom uredu na Medicinskom fakultetu Sveučilišta u Splitu. Stoga je rizik od kvara, gubitka ili otuđenja navedenih uređaja sveden na minimum. Pristup računalu je zaštićen odgovarajućom lozinkom, računalo je zaštićeno antivirusnim programom i uključenim vatrozidom te se operativni sustav redovito ažurira i nadograđuje. Pristup serveru se odvija kroz enkriptiranu SSH (<i>Secure Shell</i>) konekciju, a mogućnost pristupa uz opsežne sigurnosne lozinke imaju isključivo odobreni istraživači. Podaci će se među istraživačkim timom razmjenjivati isključivo putem enkriptiranog SSH kanala tj. zajedničkih direktorija na serveru istraživačke skupine.
	Kako ćete upravljati zaštitom autorskih prava i intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?	S obzirom da ni etičkim odobrenjem ni informiranim pristankom nije obuhvaćena dozvola za slobodno dijeljenje individualnih podataka ispitanika, nismo u mogućnosti dijeliti niti jednu vrstu sirovih podataka ispitanika prikupljenih u ovom istraživanju, niti će podaci moći biti javno dostupni. No sumarne statistike cjelogenomskih studija i rezultati statističkih usporedbi različitih fenotipskih podataka će biti pohranjeni na serveru kupljenom iz sredstava HRZZ projekta 2593 kojim upravlja voditelj projekta, a koji je smješten u Informatičkoj službi Medicinskog fakulteta Sveučilišta u Splitu. Rezultirajući sumarni podaci će također biti objavljeni uz pripadajuće publikacije pod licencijom CC BY-SA. Ne očekujemo da će naše istraživanje dovesti

		do patenta. No, u slučaju patentiranja dobivenih rezultata, problemi intelektualnog vlasništva će se riješavati prema preporukama Medicinskog fakulteta Sveučilišta u Splitu.
3.	Pohrana i čuvanje podataka	
	Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna kopija podataka (<i>backup</i>) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolazete? Kojim se procedurama koristite za sigurnosnu kopiju (<i>backup</i>)?	Laboratorijski dnevnicu će biti pohranjeni u laboratoriju voditelja projekta. Sigurnosne kopije (<i>backup</i>) podataka istraživačkog projekta će biti pohranjene na 3 lokacije; serveru istraživačke skupine, računalu odgovorne osobe, te na prijenosnom disku. Na podacima sa servera se jednom mjesečno rade sigurnosne kopije (kopije radi doktorandica na projektu koja je magistra matematike). Server ima na raspolaganju 7,6 TB prostora za pohranu. Sve informacije vezane za stvaranje sigurnosnih kopija bit će pohranjene u <i>readme</i> tekstualnim datotekama na računalu odgovorne osobe.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Elektroničke podatke ćemo čuvati 10 godina na serveru istraživačke skupine voditelja projekta za koji pristup ima voditelj projekta i istraživači kojima je odobren pristup. Sumarne statistike cjelogenomskih studija i rezultati statističkih usporedbi različitih fenotipskih podataka će se čuvati 10 godina na serveru projekta. Tekstualne datoteke će biti pohranjene u PDF obliku, a <i>excel</i> datoteke ćemo pretvoriti u CSV oblik. Genotipski podaci će biti pohranjeni u <i>VCF</i> formatu, a slike dobivene na fluorescencijskom mikroskopu će biti pohranjene u <i>TIFF</i> formatu.
4.	Dijeljenje i ponovna uporaba podataka	
	Kako i gdje će se podaci dijeliti? Na kojem repozitoriju planirate dijeliti podatke? Kako će potencijalni korisnici doznati za podatke?	Objavljene publikacije iz našeg projekta će biti pohranjene na institucijskom repozitoriju Medicinskog fakulteta Sveučilišta u Splitu, unutar nacionalnog sustava Dabar (https://dabar.srce.hr/repozitoriji). Sumarne statistike cjelogenomskih studija i rezultati statističkih usporedbi različitih fenotipskih podataka bit će pohranjeni na serveru projekta smještenom u Informatičkoj službi Medicinskog fakulteta Sveučilišta u Splitu u trenutku objavljivanja znanstvenih radova. Sumarni rezultati će također biti dostupni javno uz odgovarajuće publikacije ili na zahtjev drugih istraživača, uz licenciju CC BY-SA.
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	S obzirom da ni etičkim odobrenjem ni informiranim pristankom nije obuhvaćena dozvola za slobodno dijeljenje individualnih podataka ispitanika, nismo u mogućnosti dijeliti niti jednu vrstu sirovih podataka ispitanika prikupljenih u ovom istraživanju, niti će podaci moći biti javno dostupni. Sirovi podaci prikupljeni u ovom istraživanju će stoga biti dostupni isključivo na opravdani upit, u slučaju da se zahtjevi iz upita podudaraju s dozvolama iz informiranog pristanka.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR</i> -a.	Objavljene publikacije iz našeg projekta će biti pohranjene na institucijskom repozitoriju Medicinskog fakulteta Sveučilišta u Splitu, unutar nacionalnog sustava Dabar (https://dabar.srce.hr/repozitoriji).

<p>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).</p>	<p>Objavljene publikacije iz našeg projekta će biti pohranjene na institucijskom repozitoriju Medicinskog fakulteta Sveučilišta u Splitu, unutar nacionalnog sustava Dabar (https://dabar.srce.hr/repozitoriji). S obzirom da ni etičkim odobrenjem ni informiranim pristankom nije obuhvaćena dozvola za slobodno dijeljenje individualnih podataka ispitanika, nismo u mogućnosti dijeliti niti jednu vrstu sirovih podataka ispitanika prikupljenih u ovom istraživanju, niti će podaci moći biti javno dostupni. Sirovi podaci prikupljeni u ovom istraživanju će stoga biti dostupni isključivo na opravdani upit, u slučaju da se zahtjevi iz upita podudaraju s dozvolama iz informiranog pristanka.</p>
---	--

Ref:

[1] Celjak, D., Dorotić Malič, I., Matijević, M., Poljak, Lj., Posavec K. i Turk, I.: „Istraživački podaci - što s njima?“ [Istraživački podaci - što s njima? : priručnik o upravljanju istraživačkim podacima | Digitalni repozitorij Srca \(unizg.hr\)](#)