**PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)**

|  |  |  |
| --- | --- | --- |
| Opće informacije | | |
|  | Ime i prezime predlagatelja | Olga Gornik Kljaić |
|  | Matična organizacija | Farmaceutsko-biokemijski fakultet, Sveučilište u Zagrebu |
|  | Naziv projekta | Glikozilacija komponente komplementa C3 u razvoju šećerne bolesti tipa 1 i autoimunosti |
|  | Upravitelj podacima | Dr Domagoj Kifer |
| 1. | Prikupljanje podataka i dokumentacija | |
|  | 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) | Antropometrijski i biokemijski podaci ispitanika – dob (godine), spol, srodnici, trajanje bolesti (godine), titar autoantitijela, koncentracija C-peptida, HbA1c (postotak), koncecentracija glukoze postprandijalno i natašte (mmol/L), status dijabetičkih komplikacija (identifikacijski podaci poznati će biti samo kliničarima, ostali istraživači dobit će šifrirane uzorke i podatke)  Svi prikupljeni podaci su u obliku tablice u xlsx formatu. |
|  | 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) | ENDIA – već prikupljeni podaci bit će dostavljeni bez imena ispitanika već uz šifre (šifriranje u Australiji) elektroničkom poštom voditeljici projekta  KB Merkur – iz registra bolnice će se probrati ispitanici koji zadovoljavaju uključne kriterije te će oni biti pozvani (pozivno pismo) na uključivanje u studiju. Njihov identitet će biti poznat kliničaru (dr Pape-Medvidović), ali će ostalim istraživačima uključenima u realizaciju ovog projektnog zadatka biti dostavljeni šifrirani. Svi podaci (šifra, dob, spol, status, biokemijska mjerenja…) bit će u bolnici upisani u excel tablicu i voditeljici projekta poslani elektroničkom poštom.  Stvaranje podataka. U analizi uzorka uređajima HPLC i MS uključen je odgovarajući standard, čime se osigurava valjanost podataka. Dosljednost podatka procijenit će se usporedbom ponovljenih mjerenja.  Podaci će biti obrađeni primjenom programskog jezika R. |
|  | 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, *ReadMe* datoteke i sl.) | Dokumentacija (objašnjena pojedinih varijabli) će se navesti unutar istog excela u kojem se nalaze i podaci, u posljednjem listu dokumenta. |
| 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 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)? | Svi ispitanici prvo će biti temeljito informirani te zatim potpisuju informirane pristanke kojima daju suglasnost za uzimanje njihovih bioloških uzoraka, sakupljanje podataka i njihovu obradu uz šifrirane identitete. Koji identitet pripada kojoj šifri bit će poznato samo dr Penno (ENDIA) i dr Pape-Medvidović (KB Merkur)  U svrhu zaštite osjetljivih podataka koristit će se GDPR. |
|  | 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? | Tijekom istraživanja ispitanicima će se određivati glikozilacija C3 proteina i IgG-a. Mjerenja će biti upisana u posebnu tablicu uz pripadajuću šifru te pohranjena na server koji je zaštićen šifrom te dostupan samo zaposlenicima FBF-a koji su suradnici na ovom projektu. Statistička obrada podataka odvijat će se na osobnim računalima dr Kifera i dr Rudman koji su zaštićeni lozinkom i antivirusnim programom. Tablica sa šifriranim podacima zajedno sa izmjerenim glikoformama C3-a i IgG-a bit će uploadana na OneDrive koji je također zaštićen upotrebom određenih AAI identiteta. |
|  | 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? | Intelektualno vlasništvo i autorska prava dijele se između:   1. suradnika iz Australije i FBF-a (potpisan sporazum o dijeljenju uzoraka za studiju ENDIA) 2. suradnika s KB Merkur i FBF-a (za prikupljene uzorke autoimunog dijabetesa) 3. FBF-a (za novorazvijenu metodu analize C3 N-glikozilacije na LC-MS QQQ sustavu) |
| 3. | Pohrana i čuvanje podataka | |
|  | Kako će radne verzije podataka biti pohranjene tijekom projekta?  Kako će se napraviti sigurnosne kopije tih podataka (*backup*)?  Koja je očekivana količina podataka koja će se prikupiti i čuvati tijekom projekta (izraženo u MB/GB/TB)? | Sve radne verzije podataka bit će pohranjene na server koji je zaštićen šifrom te dostupan samo zaposlenicima FBF-a koji su suradnici na ovom projektu. Backup podataka radi se na dnevnoj bazi pomoću NAS sustava, a dodatan backup bit će odrađen na cloud serveru. Procijenjena veličina jednog raw file-a LC-MS analize je oko 300 MB, a s obzirom da će biti analizirano otprilike 1300 uzoraka procijenjena ukupna veličina raw podataka iznosi oko 400 GB. S obzirom na kapacitet servera od 2 TB to je izvediva količina podataka za pohranu i obradu. |
|  | 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)? | Završne verzije podataka ostaju na hard disku kompjutera, serveru te cloudu.  *Raw data* datoteke bit će pohranjene u njihovom originalnom formatu, dok će se podaci analize, infomacije o sudionicima te rezultati statističkih analiza pohraniti u tabličnom formatu excel.  Procijenjena finalna veličina od 400GB nije prevelika s obzirom na Zavodske kapacitete pohrane. |
| 4. | Dijeljenje i ponovna uporaba podataka | |
|  | Kako i gdje će se podaci dijeliti? Koji repozitorij će se koristit za dijeljenje podataka? Kako će potencijalni korisnici doznati za podatke? | Raw data fileovi bit će uploadani na PRIDE (Proteomics Identifications Database) digitalni repozitorij (<https://www.ebi.ac.uk/pride/>). Završne verzije podataka bit će statistički obrađene te će rezultati biti objavljenu u obliku barem dva znanstvena rada. |
|  | Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja. | - |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima *FAIR-a*. | Koristit će se repozitorij koji je u skladu s načelima FAIR-a. |
|  | 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). | PRIDE digitalni repozitorij održava neprofitna organizacija. |