Plan upravljanja istraživačkim podacima

|  |  |  |
| --- | --- | --- |
| Opće informacije | | |
|  | Ime i prezime predlagatelja | prof. dr.sc. Josip Begovac |
|  | Matična organizacija | Medicinski fakultet Sveučilišta u Zagrebu |
|  | Naziv projekta | Molekularna epidemiologija, kliničke osobitosti i skrb za osobe zaražene HIV-om |
|  | Upravitelj podacima | Josip Begovac, josip.begovac@gmail.com |
| 1. | Prikupljanje podataka i dokumentacija | |
|  | Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka) | Prikupljani podatci mogu se klasificirati u dvije kategorije:  -Socio-demografski  1.datumu ulaska u kliničku skrb  2. datum rođenja  3. dob kod ulaska u skrb  4. rizična skupina  5. zemlja porijekla  -laboratorijski podaci (imunološki i molekularni )  1. viremija virusa u krvi  2. broj CD4 limfocita T  Laboratorijski i klinički podatci koji su odabrani za istraživanje prikupit će se iz medicinske dokumentacije bolesnika koja je, sukladno internim standardnim operativnim postupcima Klinike, pohranjena u bolničkom informacijskom sustavu (BIS). Osim BIS-a kliničk-demografski podatci će se ekstrahirati i iz elektroničke baze podataka Klinike za infektivne bolesti. Pristup bazi je ograničen na 5 osoba.  Za analizu molekularne epidemiologije HIV infekcije u Hrvatskoj te utjecaja transmisijskih klastera na prevalenciju i obrazac primarne rezistencije HIV-a na antiretrovirusne lijekove u novodijagnosticiranih zaraženih osoba planira se uključiti oko 100 pacijenata godišnje.  U eksperimentanom dijelu istraživanja stvarat će se podaci populacijskog sekvenciranja u obliku nukleinskih kiselina za određivanje genotipova, podtipova i varijanti HIV-a  Kreirat će se bazu podataka bolesnika u formatu Excell gdje će biti pohranjeni rezultati populacijskog sekvenciranja. Za pohranu svih podataka bit će potrebno do 2 MB. |
|  | Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka) | Za određivanje rezistencija HIV-a na antiretrovirusne lijekove koristit će se slijed nukleinskih kiselina koji se generiraraju populacijskim sekvenciranjem na ABIprism 3500 sekvenceru. Podatci se dobiju u ABI formatu koji će se analizirati u SNAP Gene 5.3.2 software radi poravnavanja sekvenci, a potom unositi u Stanford HIV datebase, besplatne online platforme, s kojom se utvrđuju mutacije povezane sa rezistencijom. Slijed nukleinskih kiselina korist će se i za subtipizaciju virusa HIV-a koji se u fasta formatu unosi u REGA algoritam za određivanje subtipa. Za filogenetiku i filodinamiku koristit će se podaci koji se generiraju sekvenciranjem, tj, slijed nukleinskih kiselina u fasta formatu.  ZA NGS sekvenciranje podatci se generiraju u obliku fastq\_gz file-ova generiranih sa Ilumina (Miniseq) tehnologijom te unose u HyDRA Web aplikaciju (besplatno na internetu) za detekciju niskofrekventnih mutacija.  Eksperimentalne metode (sekvenciranje nukleinskih kiselina) bit će provedene prema ranije optimiziranim laboratorijskim postupcima koji sadrže potrebne analitičke kontrole. U sklopu sustava osiguranja kvalitete, u Klinici za infektivne bolesti instituciji se provodi validacija i redovito servisiranje instrumenata. U svrhu dodatnog osiguranja kvalitete, povodi se i vanjska kontrola kvalitete za laboratorijske postupke. Izvorni rezultati eksperimentalnih metoda pohranjuju se u instrumentu. |
|  | 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) | Svaki sekvencirani uzorak u obliku slijeda nukleinskih kiselina koji će se koristiti za filodinamiku i filogenetiku unijet će se u GeneBank, bazu sekvenci NIH te dobiti jedinstveni broj s kojim će biti naveden u radu koji se planira objaviti. Tako da će svaka sekvenca pod odgovarajućim brojem biti dostupna online u GeneBank-u.  Eksperimentalni postupci bit će opisani i dostupni u dijelu Materijals and methods znanstvenog rada koji će se objavljen u sklopu projekta te biti dostupni znanstvenoj javnosti bez ograničenja. |
| 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)? | Etičko povjerenstvo Medicinskog fakulteta odobrilo je ovaj projekt. Sukladno uobičajenim procedurama za biomedicinska istraživanja, svi ispitanici potpisali su informirani pristanak za sudjelovanje u istraživanju nakon što pročitali su obavijest o projektu. Zaštita osjetljivih podataka provodi se sukladno standardnim operativnim postupcima Klinike za infektivne bolesti. Projekt poštuje sva pravila i propise utvrđene Zakonom o podacima i informacijama u zdravstvu (NN 14/19). |
|  | 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? | Dio podataka pohranjen je u bolničkom informacijskom sustavu (BIS) te zaštićen u skladu sa standardnim operativnim postupcima Klinike za infektivne bolesti uz ograničenje pristupa (šifrirano). Podatci generirani u analitičkim instrumentima obrađuju se u zaštićenom nemrežnom okruženu korištenjem softvera instrumenta. Obrađeni podatci se dodatno obrađuju u anonimiziranom obliku. Pristup izvornim podatcima iz instrumenta kao i obrađenim podatcima je ograničen na tri člana istraživačke skupine te zaštićen. Svi podatci o identitetu ispitanika fizički su pohranjeni i nisu dostupni u elektronskom obliku. Svim pacijentima je pridružena jedinstvena šifra (CRO\_šifra) koju dobivaju pri ulasku u skrb i ona je oznaka i njihove sekvence u daljem radu. |
|  | 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? | Rezultati istraživanja se ne namjeravaju patentirati i ne očekuje se problem oko pitanja intelektualnog vlasništva. Rezultati istraživanja bit će objavljeni znanstvenoj javnosti bez ograničenja. Podatci koji će biti prikupljeni tijekom ovog istraživanja neće biti korišteni u bilo koju drugu svrhu (npr. neka druga istraživanja). |
| 3. | Pohrana i čuvanje podataka | |
|  | Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna kopija podataka (*backup*) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolažete? Kojim se procedurama koristite za sigurnosnu kopiju (*backup*)? | Laboratorijski zapisi i rezultati obrade podataka bit će pohranjeni u instrumentima u laboratoriju suradnika koji izvodi eksperimenatalnio dio istraživanja. Sigurnosne kopije se generiraju svakih nekoliko dana. Kapacitet pohrane svih podataka projekta bi bio DO 3MB. |
|  | Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati? | Rezultate istraživanja planiramo sačuvati barem 10 godina. Podatci će se čuvati u elektroničkom obliku na računalu voditelja projekta i suradnika u laboratoriju te u fizičkom obliku unutar institucije. |
| 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? | Kliničko-demografski podatci ispitanika se neće dijeliti i nisu javno dostupni zbog povjerljivosti podataka o pacijentu. Osjetljiva priroda ovih informacija kao i relativno mali broj pacijenata uključen u ovaj skup podataka može dovesti do identificiranja i stoga ne dopušta da ovaj skup podataka bude javan.  Sekvence HIV gena bit će unešene u bazu GeneBank (https://www.ncbi.nlm.nih.gov/genbank/) prilikom slanja izvornih znanstvenih radova na recenziju. Nakon što se objavi znanstveni rad proistekao iz istraživanja , svi zainteresirani mogu pristupiti izvornim rezultatima temeljem pristupnog broja u GeneBank koji će biti naveden u objavljenom radu. |
|  | Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja. | Ne postoje. |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima *FAIR-a*. | 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 |

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)](https://repozitorij.srce.unizg.hr/islandora/object/srce:327)