

Universitetet i Oslo - Det matematisk-naturvitenskapelige
fakultet
Eivind Hovig
Postboks 1032 Blidern
0315 OSLO

Deres ref.:
Vår ref.: 17/34742-7
Saksbehandler: Kari Steig
Dato: 28.11.2018

Hjemmelsgrunnlag for anonymiseringsprosess i forbindelse med etablering av nasjonal, anonym sekvensvariantdatabase

Innledning

Som et ledd i etableringen av en nasjonal database over genetiske varianter, har det blitt reist flere spørsmål om utlevering av anonyme opplysninger fra et forskningsregister til denne databasen. Det er også stilt spørsmål til hjemmelsgrunnlag for anonymiseringsprosessen av genetiske varianter som skal utleveres anonymt fra interne kvalitetsregistre til nasjonal anonym variantdatabase. Spørsmålene er knyttet til anonymisering og anonymitet. Hva som anses som en anonym opplysning er ferdig vurdert. Man kan opprette nasjonal anonym frekvensvariantdatabase med varianter som er observert 4-5 ganger eller flere.¹ Spørsmålet som gjenstår, er hjemmelsgrunnlaget for selve anonymiseringsprosessen.

Anonymisering av helseopplysninger fra forskningsprosjekt

Spørsmålet gjelder hjemmelsgrunnlaget for anonymiseringen av helseopplysninger fra et eksisterende godkjent forskningsprosjekt der formålet med anonymiseringen er å utlevere anonyme data til den nasjonale anonyme variantdatabasen².

Vi legger følgende faktum til grunn

1. Det finnes en database som inneholder DNA-varianter som kommer fra genomsekvensering av et antall norske individer utført i forskningssammenheng. Disse variantene er frikoplet fra hverandre og fra andre DNA-varianter fra samme kromosomstreng, slik at en ikke kan sammenkople variasjon fra flere posisjoner i genomet til et individ basert på DNA.
2. Fra denne databasen skal det hentes ut et anonymt datasett: bare varianter der det er 4 eller flere like observasjoner i hver registrert posisjon i genomet – eksempelvis minst 4xT og/eller 4xC.

¹ Vurdering av anonymitet ref. 17/34742-2 og Referat fra møte om nasjonal anonym sekvensvariantdatabase, ref. 17/34742-3

² 2015/2382/REK sør-øst

Hesledirektoratets vurdering

Det fremkommer av helseforskningsloven § 2 annet ledd, at personvernforordningen (GDPR) kommer til anvendelse i medisinsk og helsefaglig forskning på helseopplysninger med mindre noe annet følger av helseforskningsloven.

Det er klart at verken helseforskningsloven eller personvernforordningen (GDPR) kommer til anvendelse på anonyme opplysninger. Dette følger av bl.a. helseforskningslovens § 20, og fortalen til forordningen, punkt 26. Dette fordi anonyme opplysninger verken er person- eller helseopplysninger og følgelig faller utenfor virkeområdet til lov og forordning. Helseforskningslovens § 20 omtaler anonymiserte opplysninger, men ikke hva hjemmelsgrunnlaget for selve anonymiseringen er. Det aktuelle spørsmålet er om det er behov for et selvstendig/eget behandlingsgrunnlag for behandlingen av helseopplysninger i anonymiseringsprosessen, eller om denne behandlingen dekkes av det allerede eksisterende behandlingsgrunnlaget i forskningsprosjektet. Behandlingsgrunnlaget for dette forskningsprosjektet er samtykke.

Opplysningene som i dette tilfellet skal anonymiseres hentes fra et samtykkebasert forskningsprosjekt, der formålet er å identifisere gener og genvarianter som disponerer for arvelig bryst- og tarmkreft. Deltakerne har avgitt et bredt samtykke til forskning "med formål å finne underliggende arvelige årsaker til bryst- eller tarmkreft". Det fremkommer ikke eksplisitt av samtykkeskrivet at samtykket også gjelder anonymisering av helseopplysningene til annen bruk.

Formålet med anonymiseringen (viderebehandlingen) av enkelte opplysninger fra forskningsdatabasen, er å utlevere anonyme data til en variantdatabase over genetiske varianter. Etter vår vurdering, vil dette anses som statistikk (det telles hvor mange ganger man har sett varianter der det er 4 eller flere like observasjoner i hver registrert posisjon – eksempelvis minst 4xT og/ eller 4 xC).

Viderebehandlingen av helseopplysningene, dvs. anonymiseringen, er et annet formål enn det opprinnelige formålet i forskningsprosjektet, og trenger derfor i utgangspunktet et nytt behandlingsgrunnlag.

De overordnede prinsippene for all behandling av personopplysninger fremkommer av forordningens artikkel 5. Av bestemmelsens bokstav b fremkommer det at en viderebehandling av opplysningene vil være begrenset av det opprinnelige formålet. Dette utdypes i forarbeidene kapittel 6.6.³ "Dersom formålet *ikke* er forenelig med innsamlingsformålet, må viderebehandlingen ha grunnlag i lov eller samtykke." Det fremkommer videre av artikkel 5 nr. 1 bokstav b, at viderebehandling for statistiske formål ikke skal anses som uforenelige med det opprinnelige formålet. Dette presiseres også i fortalen, punkt 50. Viderebehandling for nye forenlige formål, er tillatt uten eget rettsgrunnlag for viderebehandlingen, jf. forordningen artikkel 6 nr. 4.

³ Prop. 56 LS (2017-2018) Lov om behandling av personopplysninger (personopplysningsloven m.m.)

Det opprinnelige formålet i forskningsprosjektet er slik vi ser det dermed ikke i strid med det nye formålet, som er anonymisering av helseopplysninger til bruk i statistikk. Det vises også til fortalen i forordningen pkt. 162 sin omtale av behandling av personopplysninger for statistiske formål.

Det følger også av helseforskningsloven at helseopplysninger ikke kan anvendes til formål som er uforenlig med det opprinnelige formålet uten at forskningsdeltakeren samtykker, med mindre annet følger av lov. jf. helseforskningsloven § 32 2. ledd.

Artikkel 5 nr. 1 bokstav b, forutsetter at vilkårene i forordningens art 89 nr. 1 følges. Dette innebærer at den registrertes rettigheter skal garanteres. Dette kan gjøres ved bl.a. dataminimering. Pseudonymisering er nevnt som et eksempel. I dette tilfellet skal det gjøres ytterligere dataminimeringstiltak ved at opplysningene skal anonymiseres. Dataansvarlig har lovlig tilgang til helseopplysningene i forskningsprosjektet. Det er opplyst at man vil gå inn i et eksisterende datasett internt i virksomheten og foreta en teknisk anonymisering av deler av datasettet. Denne anonymiseringen vil foregå der hvor dataene oppbevares og analyseres lovlig i dag – i Tjenester for sensitive data (TSD), som er godkjent for forskningsformål ved OUS, og der det foreligger databehandleravtale mellom OUS og UiO. Anonymiseringen vil foregå ved at kopling mellom personidentitet og variant fjernes i maskinens hukommelse og skrives til en ny fil. Denne filen med den anonymiserte informasjonen er utledet via et program som er startet av en autorisert forsker i det aktuelle TSD-prosjektet. Noen nøkkel i datasettet i resultatfilen finnes ikke. Kun varianter med 4-5 eller flere vil bli utlevert, ved at resultatfilen filtreres for slik at alle posisjoner med varianter med færre enn 4 vil bli fjernet før eksport ut av TSD-prosjektet. Dataansvarlig vil ikke kunne tilegne seg noen ny informasjon om den enkelte forskningsdeltaker.

Avslutningsvis bemerkes det at det er datasansvarlig(OUS) som er ansvarlig for å sikre at forordningens krav til behandling av personopplysninger er oppfylt. Det vises i den forbindelse særlig til retten til informasjon i artiklene 13 og 14. For dette prosjektet er det rett til informasjon etter artikkel 14 som vil være bestemmende da opplysningene samles inn fra andre enn den registrerte. I bestemmelsens punkt 4, har den registrerte rett til informasjon når opplysningene viderebehandles. Dette utdypes også i fortalens punkt 50 og Article 29 Working Party Guidelines on "Transparency under regulation 2016/679 punkt 45"⁴. Det følger dog av artikkel 14 nr. 5 at det er flere unntak fra informasjonsplikten. Det fremkommer bl.a. av bestemmelsens bokstav b), at dersom det er umulig eller vil innebære en uforholdsmessig stor innsats, kan informasjonskravet under visse vilkår fravikes. Helsedirektoratet bemerker at informasjon kan gis på ulike måter, som f.eks. på nettstedet til selve prosjektet eller tilsvarende. Det er dataansvarlig som skal foreta en konkret vurdering av dette.

Anonymisering av helseopplysninger i interne kvalitetsregistre

Du har også stilt spørsmål til hjemmelsgrunnlag for anonymisering av helseopplysninger til nasjonal anonym variantdatabase når disse er samlet inn fra behandlingsrettede helseregistre.

⁴ http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227

Den anonyme variantdatabasen skal anrikes med anonyme genetiske varianter fra helseforetakenes kvalitetsregistre. Se vedlagt beskrivelse i notat.⁵

Vi legger til grunn at de genetiske avdelingene i foretakene har etablert interne kvalitetsregistre over genetiske varianter for tidligere pasienter som benyttes for å kvalitetssikre diagnostikk i avdelingen, med grunnlag i helsepersonelloven § 26 og pasientjournalloven § 6.

Helsedirektoratets vurdering

Anonymisering av disse dataene for utlevering til nasjonal anonym variantdatabase er et annet formål enn det opprinnelige formålet i kvalitetsregisteret, og trenger derfor i utgangspunktet et nytt behandlingsgrunnlag. Etter vår vurdering vil dette anses som statistikk (det telles hvor mange ganger man har sett varianter der det er 4 eller flere like på hvert punkt – altså minst 4xT, 4xG, 4xA eller 4xC). Etter artikkel 5 nr. 1 bokstav b, står det at viderebehandling for statistiske formål ikke skal anses som uforenlige med det opprinnelige formålet. Viderebehandling for nye forenlige formål er tillatt uten eget rettsgrunnlag for viderebehandlingen, jf. forordningen artikkel 6 nr. 4. Også her er det datasansvarlig ved foretakene som må vurdere om forordningens krav til behandling av personopplysninger er oppfylt.

Utlevering av anonyme data fra kvalitetsregistrene til nasjonal anonym variantdatabase kan skje med hjemmel i helsepersonellovens § 23 nr.3. Det vises også til at personvernforordningen (GDPR) ikke kommer til anvendelse på anonyme opplysninger.

Spørsmål til anonymiseringsprosess ved anriking av data fra helseforetakenes kvalitetsregistre til nasjonal anonymt kvalitetsregister.

I ditt notat beskriver du hvordan anonyme data skal anrikes fra foretakenes kvalitetsregistre til nasjonal anonymt variantregister. Flere interne kvalitetsregistre vil motta den anonyme datafilen fra nasjonalt variantdataregister. Data fra interne kvalitetsregistre ved flere ulike virksomheter (leverandørfiler) hentes inn parallelt. Varianter som det finnes 4-5 av lokalt kan anrikes til basisfila. Lokale varianter som det er mindre enn 4 av kan anrikes til basisfila hvis denne allerede inneholder 4 eller flere av den aktuelle varianten. Deretter slås alle data som er innhentet i parallell sammen til en ny resultatfil og samtlige data utlevert fra kvalitetsregistrene (leverandørfiler) slettes. Det opprinnelige nasjonale variantdatasettet som er utlevert fra ditt forskningsprosjekt (basisfil) slettes også. Det eneste gjenværende aggregerte fila er den nye, som inneholder variantinformasjon fra den opprinnelige basisfila og leverandørfilene. Resultatfila er en sammenslått informasjonsfil med anonyme varianter og vil være et anonymt datasett.

Vi legger til grunn at dette opplytter krav til utlevering av anonyme data. Vi forutsetter da at det ikke vil være mulig å identifisere hvilke kilder hver enkelt variant har, og at ingen vil ha oversikt over andre enn sin egen kilde og resultatfila, der den siste er sammensatt av flere kilder.

⁵ Prosedyre for anonymisering av variantdata

Konklusjon

- Behandling av anonyme opplysninger trenger ikke et selvstendig behandlingsgrunnlag da anonyme opplysninger faller utenfor GDPR sitt virkeområde
- Anonymiseringsprosessen trenger ikke et selvstendig hjemmelsgrunnlag da anonymisering for statistikk er forenelig med det opprinnelige forskningsformålet og formålet med kvalitetsregisteret. Dette forutsetter blant annet at artikkel 89 nr. 1 om nødvendige garantier er oppfylt. Vi antar at når opplysningene minimeres, anonymiseres og håndteres som beskrevet ovenfor, er dette tilstrekkelig for at nødvendige garantier oppfylt. Det er dataansvarlig som er ansvarlig for å sikre at forordningens krav til behandling av personopplysninger er oppfylt.
- Forskningsdeltakerene har rett til informasjon om viderebehandlingen, med mindre unntaket i artikkel 14 nr. 5 bokstav b) kommer til anvendelse.

Vennlig hilsen

Anne Louise Valle e.f.
Avdelingsdirektør

Kari Steig
Seniorrådgiver

Dokumentet er godkjent elektronisk

Kopi:
Helse- og omsorgsdepartementet

Møte om: Etablering av nasjonal anonym sekvens-variantdatabase

Referat

Møteleder: Anne Louise Valle

Dato: 02.05.2018

Referent: Ingunn Myklebust

Saksnr: 17/34742-3

Til stede: HSØ: Torunn Berge, Heidi Thorstensen, Øystein Krüger, Randi Lilletvedt,
Hdir: Hege Wang, Anne Loise Valle, Ola Johnsborg, Anne Fous og Ingunn Myklebust

Referat fra møte om nasjonal anonym sekvensvariantdatabase

Bakgrunn:

Helse- og omsorgsdepartementet (HOD) har gitt Helse Sør-Øst (HSØ) i oppdrag å etablere en anonym frekvensdatabase over sekvensvarianter. I brev av 21. november 2017 ba HOD Helsedirektoratet svare på HSØs henvendelse om juridisk avklaring knyttet til etablering av anonym sekvensvariantdatabase. I brev av 28. november 2017 ga Helsedirektoratet en juridisk vurdering av hvilke opplysninger som kan være med slik at databasen kan anses å være anonym, dvs. at en sekvensvariant ikke vil anses å være anonym dersom den er observert svært få ganger (< 4-5).

Tema for møtet:

- 1) Formål med databasen er å filtrere vekk normal genetisk variasjon i pasientprøver

I arbeidet med å bestemme hvorvidt en genetisk variant er sykdomsgivende eller ikke er det viktig å kunne sammenligne funn hos pasient med genetisk normalvariasjon i den øvrige befolkningen. Ved å sammenligne varianter i pasientprøven mot normalvariasjon vil mengden av varianter som må analyseres videre reduseres og dermed blir det lettere å identifisere sjeldne og potensielt sykdomsfremkallende varianter hos pasienten.

Det ble avklart i møtet at man vil oppnå dette formålet med opprettelse av nasjonal anonym frekvensvariantdatabase med varianter som er observert 4-5 ganger eller flere.

Det er derfor ikke nødvendig å gå videre med en rettslig vurdering av om man også kan lagre varianter anonymt hvis cut-off er observasjoner færre enn 4-5 ganger.

Vurdering av om genetiske opplysninger er anonyme eller indirekte identifiserbare har vist seg vanskelig å gjennomføre i praksis. Lagring av genetiske varianter som kan knyttes til helseopplysninger om enkeltpersoner vil være en alvorlig personvernkrænelse og brudd på taushetsplikten.

Det var enighet om at Helsedirektoratet i brev av 28. november gir en tilstrekkelig redegjørelse av anonyme opplysninger for å starte arbeidet med å etablere en nasjonal anonym sekvensvariantdatabase.

- 2) Etablere en nasjonal sekvensvariantdatabase som inkluderer sjeldne genetiske varianter og eventuelt også tilhørende kliniske metadata. Formålet med databasen vil være å kvalitetssikre diagnostikk.

I møte ble det pekt på at dette vil kreve behandlingsgrunnlag da slike opplysninger vil være identifiserbare. Man kan vurdere å opprette et slikt kvalitetsregister med grunnlag i samtykke og/eller eventuelt dispensasjon fra taushetsplikt etter helseregisterloven § 29 b. Dette må utredes nærmere juridisk. Før Helsedirektoratet kan utrede dette må faktum beskrives mer detaljert. Viktige moment som vil ligge til grunn for juridisk vurdering vil bl.a. være; 1) Hvilke data er tenkt inkludert i databasen 2) Hvordan skal data innhentes (f. eks. samtykke) 3) Nøyaktig beskrivelse av formålet med bruken 4) Hvem skal ha tilgang 5) Hvordan blir tilgang styrt 6) Hvem skal være databehandlingsansvarlig.

Helse Sør-Øst RHF
Torunn Berge
Postboks 404
2303 HAMAR

Deres ref.: 16/724
Vår ref.: 17/34742-2
Saksbehandler: Tonje Borch
Dato: 28.11.2017

Etablering av nasjonal anonym sekvensvariantdatabase - vurdering av anonymitet

HelseDirektoratet viser til oppdrag fra Helse- og omsorgsdepartementet av 21. november 2017, der departementet ber oss om å svare på Helse Sør-Østs (HSØ) henvendelse om avklaringer knyttet til etablering av anonym sekvensdatabase. Videre viser direktoratet til notat som ble forelagt arbeidsgruppen for etablering av nasjonal anonym sekvensdatabase 31. januar 2018, samt mail av 6. februar fra Torunn Berge med vedlegg til henvendelsen av 21. november.

HSØ oppstiller i sin henvendelse tre problemstillinger de ønsker en juridisk avklaring knyttet til:

1. Vil en sekvensvariant fortsatt anses anonym dersom den kun er observert svært få ganger (<5)?
2. Kan kobling til tekniske metadata som tidsperiode og laboratorium understøtte kravet om anonymitet?
3. Er databasen å anse som anonym dersom sekvensvarianter kobles til en klinisk fortolkning/indikasjon?

Sakens faktum

Utgangspunktet for henvendelsen er HODs oppdrag til HSØ om å etablere en anonym frekvensdatabase over sekvensvarianter. I brevet fra HSØ omtales sekvensvariasjon primært som de endringer i DNA-sekvensen som omfatter ett til noen få nukleotider.

En norsk frekvensdatabase skal være et verktøy til bruk i kvalitetssikring av genetisk diagnostikk. I henvendelsen fra HSØ er innholdet i frekvensdatabasen beskrevet på følgende måte: «*Datasettet i den norske variantdatabasen vil være aggregert fra sekvensinformasjon innhentet fra norske pasienter og pårørende (>5000 individer og stadig voksende) med allel-frekvenser estimert fra dette aggregatet. Dette innebærer at innholdet i databasen vil være av dynamisk karakter så lenge ny informasjon blir lagt til.*

Den enkleste versjonen av databasen vil bygge på såkalte frekvensstudier der man enkelt forklart teller antall forekomster av en bestemt sekvensvariant i et innsamlet gendatamateriale. En populasjonsspesifikk variantdatabase slik den er beskrevet i oppdraget vil derfor kunne si noe om

HelseDirektoratet

Avdeling bioteknologi og helserett
Tonje Borch, tlf.: +4724163946

Postboks 7000 St. Olavs plass, 0130 Oslo • Besøksadresse: Universitetsgata 2, Oslo • Tlf.: 810 20 050
Faks: 24 16 30 01 • Org.nr.: 983 544 622 • postmottak@helsedir.no • www.helseDirektoratet.no

utbredelsen av bestemte genvarianter i den norske befolkningen».

Det framgår av henvendelsen at det i tillegg er ønskelig og nødvendig å koble tekniske data til sekvensvariantene for å understøtte kvalitetssikring og vedlikehold av databasen. Dette vil f.eks. være informasjon om type sekvenseringsplattform, dato for analyse og lokalisering av sekvenseringslaboratorium.

Innhold og bruk av databasen er nærmere beskrevet i notat til møte RHFenes arbeidsgruppe 31.01.18., samt i vedlegg til henvendelse mottatt i mail 6. februar. Disse dokumentene bekrefter at den anonyme frekvensdatabasen skal være en referanse med norske variantfrekvenser. Ved trio-sekvensering (foreldre+ barn) er det bare data fra foreldre som vil inngå.

Rettslig grunnlag

Pasientjournalloven, helseregisterloven og personopplysningsloven oppstiller ingen krav til behandling av anonyme opplysninger, og en korrekt identifisering av hva som kan defineres som anonyme opplysninger blir derfor viktig for å ivareta den registrertes personvern. Å anonymisere data er utfordrende, og det er mer utfordrende i dag enn det var tidligere. Det enorme tilfanget av offentlig tilgjengelige data, kombinert med tilgang til stadig billigere og mer kraftfull analyseteknologi, har bidratt til å øke faren for reidentifisering.¹

Ved behandling av anonyme opplysninger skal det ikke, direkte eller indirekte, være mulig å spore opplysningene tilbake til de enkeltpersoner opplysningene knytter seg til. Det fremgår av Helseregisterlovens forarbeider og av fortalen til ny personvernforordning (avsnitt 26) at det bør tas hensyn til alle midler som det med rimelighet kan tenkes at den behandlingsansvarlige eller en annen person kan ta i bruk for å identifisere vedkommende direkte eller indirekte. For å fastslå om midler med rimelighet kan tenkes å bli tatt bruk for å identifisere den fysiske personen bør det tas hensyn til alle objektive faktorer, f.eks. kostnadene for og tiden som er nødvendig for å foreta identifikasjonen. Det må tas hensyn til teknologien som er tilgjengelig på behandlingstidspunktet, samt den teknologiske utvikling.

I praksis er det i Norge lagt til grunn at opplysningene først kan anses som anonyme dersom de kan knyttes til en av fire eller fem individer. Dette er en streng fortolkning og samsvarer med de anonymiseringsprinsippene SSB anvender i sin statistikkfremstilling.

Det er ifølge Datatilsynets rapport: "Personvernutfordringer ved genetiske undersøkelser" gjort rede for genetiske undersøkelser status i nasjonalt og internasjonalt regelverk. Med utgangspunkt i et sammensatt rettskildebilde har Datatilsynet konkludert med at genetiske undersøkelser har en særskilt status som spesielt sensitive opplysninger.²

¹ <https://www.datatilsynet.no/globalassets/global/regelverk-skiema/veiledere/anonymisering-veileder-041115.pdf>

² https://www.datatilsynet.no/globalassets/global/om-personvern/rapporter/genetiske_undersokelser_rapport_v2.pdf

Helsedirektoratets vurdering

Til spm 1

Vil en sekvensvariant fortsatt anses anonym dersom den kun er observert svært få ganger (<5)?

Som nevnt ovenfor tilsier dagens praksis at opplysningene som et utgangspunkt først kan anses som anonyme dersom de kan knyttes til en av fire eller fem individer.

Teoretisk kan det være mulig å identifisere et individ ut fra beskrivelse av en sjelden variant. Det kan for eksempel være mulig hvis den samme varianten finnes i en annen database med *personidentifiserbare* opplysninger, eller man har tilgang til personidentifiserbare kliniske data som utvetydig kan knyttes til den aktuelle varianten. Dette medfører at en sekvensvariant ikke vil anses for å være anonym dersom den kun er observert svært få ganger (<5).

Det fremgår imidlertid av notatet til møtet i arbeidsgruppa 31.01.18 at det er mulig å innrette databasen slik at den ikke oppgir eksakt frekvens for eller inneholder de sjeldne variantene. Cut-off for hva som anses som sjeldent kan for eksempel settes til 4/5 alleler eller færre i den aktuelle sykdomsgruppen.

Slik vi forstår dette, vil en slik database ikke ha informasjon om eksakt hvor mange ganger en sjelden variant er observert og heller ikke nærmere beskrivelse av den sjeldne varianten (jf. begrepet «non disclosed, rare variant»). Et søk i, eller filtrering mot, databasen vil kunne oppgi at det finnes en sjelden variant i den aktuelle posisjonen, og at det er observert 4/5 alleler eller færre. Vi kan ikke se at det vil være mulig å identifisere enkeltindivider på bakgrunn av dette, og mener at et register som er innrettet på denne måten må kunne sies å være anonymt. Det er vår oppfatning at en slik innretning vil være i tråd med dagens prinsipper om anonymisering.

Til spm 2

Kan kobling til tekniske metadata som tidsperiode og laboratorium understøtte kravet om anonymitet?

Det er i tillegg stilt spørsmål om inkludering av tekniske data. Vi oppfatter at dette er viktig og nødvendig for å sikre god kvalitet i datasettet som til enhver tid ligger i databasen. Notatet til arbeidsgruppa gir flere eksempler på tekniske data som er ønskelig å inkludere i databasen. Så lenge de sjeldne variantene håndteres slik det er beskrevet i notatet til arbeidsgruppa ser vi ikke at rent tekniske opplysninger om sekvenseringsplattform, protokoll for prøvepreparering, dekningsgrad, dato, laboratorium etc. vil bidra til at enkeltindivider kan identifiseres. Det vil imidlertid være avgjørende at inkludering av slike data ikke gjør det mulig å re-identifisere enkeltindivider i datasettet. For å sikre at dette ikke skjer vil direktoratet foreslå at tekniske data knyttes til databasene som den enkelte avdelingen har over sine varianter, og at det settes tekniske krav til datasett som inkluderes i den anonyme frekvensdatabasen som må oppfylles.

Til spm 3

Er databasen å anse som anonym dersom sekvensvarianter kobles til en klinisk fortolkning/indikasjon?

Som ovenfor vil utgangspunktet være at det er faren for reidentifikasjon av enkeltindivider som avgjør om databasen vil være anonym eller ikke. Om man kobler variantene til en klinisk

fortolkning/indikasjon vil dette fortsatt kunne være anonymt dersom samme fortolkning knyttes til varianter som er observert 4-5 ganger eller flere.

Helsedirektoratets videre oppfølging av de nevnte problemstillinger

HSØ sier i vedlegg til henvendelsen; "*for at en nasjonal database skal ha størst mulig klinisk verdi, er det viktig at det kan gjennomføres aggregering av variantfrekvenser slik at databasen får bredest mulig nasjonal representasjon. Dette vil innebære at det må legges til rette for innhenting av data, med godkjent prosedyre, fra alle nasjonale leverandører av genetisk data. Skulle denne leveransen blir begrenset på lokalt nivå, vil også databasen kliniske betydning bli betydelig redusert. Det følger av dette at dersom hver enkelt lokale kilde pålegges en nedre terskelverdi eller rapporteringsgrense, vil det samlet sett være et stort antall varianter som ikke vil fanges opp, fordi de ikke kan summeres nasjonalt. I dette tilfelle vil databasen miste mye av sin nytteverdi for klinisk anvendelse.*" Dette betyr, slik Helsedirektoratet oppfatter det, at en anonymitetsvurdering som baseres på at variantdatabasen først anses som anonym dersom variantene kan knyttes til en av fire eller fem individer, vil påvirke databasens nytteverdi. Etter vår oppfatning vil dette kunne avhenge av hvordan data samles inn. Hvis alt samles inn nasjonalt uten tekniske data og varianter færre enn 4/5 vil databasen kunne være nyttig som verktøy til bruk i kvalitetssikring av diagnostikk.

Vi ser imidlertid at en frekvensdatabase vil kunne ha større potensiale dersom man også kan lagre varianter hvor cut off ikke er 4/5, og hvor man kan knytte flere variabler til dataene. Spørsmålet blir imidlertid da om databasen fortsatt kan sies å være anonym. Det er Helsedirektoratets oppfatning at man bør se nærmere på hvilke krav til anonymisering man skal legge til grunn når det gjelder genetiske varianter. Særlig når det, som her, er snakk om genetiske varianter som i seg selv ikke er personidentifiserbare med mindre de har andre variabler knyttet til seg. Den rettslige vurderingen er, som nevnt ovenfor, knyttet til hva som skal til for å identifisere et enkelt individ ut i fra de opplysninger man har tilgjengelig i registeret og tilfanget av offentlig tilgjengelige data, kombinert med tilgang til stadig billigere og mer kraftfull analyseteknologi. Vurderingen vil videre være knyttet til en vurdering av alle midler som det med rimelighet kan tenkes at den behandlingsansvarlige eller en annen person kan ta i bruk for å identifisere vedkommende direkte eller indirekte. Konklusjonen vil ha betydning for hvordan vi i fremtiden behandler genetiske opplysninger i Norge, og bør etter Helsedirektoratets oppfatning bygges på en grundig utredning. Direktoratet vil prioritere arbeidet med en slik utredning våren 2018.

Vennlig hilsen

Anne Louise Valle e.f.
avdelingsdirektør

Tonje Borch
seniorrådgiver

Dokumentet er godkjent elektronisk

Kopi:
Helse- og omsorgsdepartementet