

Januar 2020

Hjemsendelse af analyseresultater

Retningslinjer for Sundhedsdatastyrelsens forskermaskine



**SUNDHEDSDATA-
STYRELSEN**

Hjemsendelse af analyseresultater

Retningslinjer for Sundhedsdatastyrelsens forskermaskine

Sundhedsdatastyrelsens retningslinjer for hjemsendelse af analyseresultater har som udgangspunkt samme indhold som Danmarks Statistiks retningslinjer. Dette er gjort for at ensarte retningslinjerne for de to forskermaskine-miljøer. Der er dog to væsentlige forskelle: Sundhedsdatastyrelsen håndhæver, at hjemsendte data som minimum skal indeholde 5 observationer pr. celle mod Danmarks Statistiks minimum på 3 observationer, og det vil altid være den samme autoriserede institution, der er ansvarlig for samtlige brugere på et projekt.

Den enkelte borger har krav på, at fortrolige data behandles med den største omhyggelighed. Formålet med disse retningslinjer er at beskytte mod misbrug af borgernes data. Bruges data korrekt kan analyserne være til gavn for det danske samfund. Derfor er det vigtigt, at anvendelsen sker på en måde, hvor både datasikkerhed og datafortrolighed er i top.

Reglerne for hvordan fortrolige data skal behandles er fastlagt i databeskyttelsesloven og i de aftaler som brugere af Forskermaskinen skal indgå for at få adgang til data. Data på Forskermaskinen er fortrolige og brugere af Forskermaskinen er forpligtet til kun at hjemsende generelle resultater og aldrig individdata fra Forskermaskinen.

Nedenfor gives en beskrivelse af hovedreglerne for hjemsendelse af filer, herunder hvad der menes med hhv. mikrodata og detaljerede tabeller og resultater; proceduren ved et brud på datafortroligheden samt nogle gode råd til, hvordan man undgår at hjemsende data man ikke må.

Hvis Sundhedsdatastyrelsen konstaterer, at en bruger har hjemsendt mikrodata (se mere herom i afsnittet Hjemsendelse af mikrodata), lukkes der omgående for adgangen til Forskermaskinen. Det gælder både for den bruger, som har hjemsendt mikrodata og for alle andre brugere fra den autoriserede institution.

Hvis en forsker, der er bruger under forskellige autorisationer, hjemsender mikrodata fra et projekt under den ene autorisation, foretages der intet i forhold til de øvrige autoriserede institutioner som personen er bruger under. For den person, der har hjemsendt mikrodata, lukkes adgangen til alle projekter, uanset hvilken institution de hører under.

Gode råd

Nedenfor følger en liste med gode råd om, hvordan man undgår brud på reglerne om datafortrolighed og datasikkerhed:

For brugeren, der har adgang til Forskermaskinen:

1. Hjemsendte data skal som hovedregel indeholde minimum 5 observationer pr. celle
2. Skriv aldrig programmer der skriver direkte til mailmappen
3. Kontroller altid samtlige filer, der skal hjemsendes
 - a. Kontakt altid din autorisationsansvarlige i tvivlstilfælde

4. Begræns hjemsendelsen af filer til det mest nødvendige. Spørg fx dig selv:
 - a. Er det nødvendigt at hjemsende programmer og logfiler? – de opbevares sikkert på Forskermaskinen
 - b. Er det nødvendigt at hjemsende foreløbige resultater?

For den ansvarlige for autorisationen:

1. Lav en plan der sikrer, at alle forskere er orienteret om reglerne vedrørende datafortrolighed og datasikkerhed under forskerordningen, herunder også:
 - a. At både den sidst tilkomne og den udenlandske forsker har tilstrækkelig viden om reglerne
 - b. At der i miljøet findes en kontaktperson ved tvivlstilfælde
2. At alle autoriserede forskere/analytikere, du har ansvar for og har underskrevet brugeraftaler for, kender proceduren ved et regelbrud
3. At alle forskere, der arbejder under autorisationen er bekendt med de "fælder" der kan være med at arbejde med mikrodata – fx at logfiler kan indeholde mikrodata
4. Lav evt. kontrolforanstaltninger ved hjemsendelse af resultater
5. Begræns evt. antallet af forskere pr. projekt, der må sende output hjem
6. Kontroller jævnligt, at kun de relevante forskere har adgang til Forskermaskinen og orienter Forskerservice, hvis der er brugere, der skal have deaktiveret deres adgang, jvf. Autorisationsaftalen.

Hvad må hjemsendes?

Som hovedregel må der kun hjemsendes analyseresultater, aggregerede tabeller eller figurer, hvor det ikke er muligt at identificere enkelte enheder, dvs. hverken personer, familier eller andre enheder. Data på enkelt individer må ikke hjemsendes. Dette gælder også selvom det pseudonymiserede CPR-nummer eller løbenummer er fjernet. **Det må med andre ord ikke være muligt at genkende nogen enkeltpersoner i det hjemsendte materiale.**

Generelt gælder nedenstående tommelfingerregler for aggregerede data:

Hjemsendte resultater bør som udgangspunkt have sådan en form, at de umiddelbart kan anvendes i en publikation.

Det er tilladt at arbejde videre med det hjemsendte aggregerede materiale, fx til at dannelse af figurer eller videre analyser i statistikprogrammer.

For tabeller gælder nedenstående:

Tabeller bør indeholde mindst 5 observationer pr. celle
For tabeller er tommelfingerreglen, at tabeller skal indeholde mindst 5 observationer, men aggregeringsniveauet skal altid afhænge af en konkret vurdering af, hvad tabellen indeholder.
Max. og min. og percentiler, som fx median, må kun hjemsendes, hvis der ikke er fare for identifikation af enkeltindivider. Ingen personer må kunne genkendes i det hjemsendte output.

Funktioner som min. og max. henviser ofte til enkeltindivider, men de kan hjemsendes, hvis det er sikret, at ingen personer kan genkendes i materialet.

Figurer må kun hjemsendes, hvis de ikke indeholder information, som kan identificere enkeltindivider. Man skal ved figurer være særligt opmærksomme på ekstreme værdier såkaldte outliers. Det skal derudover sikres, at figurerne ikke indeholder indlejrede mikrodata. Sidstnævnte kan fx sikres ved at hjemsende figuren i PDF-format.

Hvad må ikke hjemsendes?

Ingen form for mikrodata må hjemsendes

Mikrodata defineres som data, der fx indeholder nedenstående:

- Datasæt – eller dele af datasæt - med oplysninger på individniveau, dvs. hvor der i datasættet findes en række for hvert individ (enkeltekords), også i de tilfælde, hvor man har fjernet det pseudonymiserede løbenummer fx det pseudonymiserede personnummer. Dvs. datasæt bestående af observationer på enkeltindivider med fx fødselsdato, kommune, køn, diagnose eller ATC kode.
- Pseudonymiserede nøglevariabler såsom personnumre og ydernumre betragtes altid som mikrodata, da de angiver et unikt nummer, der henviser direkte til en enkelt person eller praktiserende læge. Derfor er det heller ikke tilladt at hjemsende pseudonymiserede personnumre alene. Dette gælder også egne data, dvs. data som man selv har indsendt til Sundhedsdatastyrelsen og som skal indgå i et konkret projekt.

Pas på med programfiler og logfiler

Programmer og logfiler kan også indeholde mikrodata fx:

- Programkode der indeholder mikrodata skrevet ind i en betingelse
- Logfiler der lister observationer. Nogle procedurer lister fx et lille udsnit af det datasæt, der er anvendt til analysen

Hvis man sender en logfil hjem, der lister pseudonymiserede personnumre er det et brud på reglerne – også selv om disse pseudonymiserede personnumre ikke indeholder anden information.

Al arbejde med mikrodata skal foregå på Forskermaskinen

Det er tilladt at danne logfiler, der lister pseudonymiserede personnumre – fx som kontrol af et program, men en sådan fil SKAL, i lighed med alle andre filer der indeholder mikrodata, blive på forskermaskinen og må IKKE hjemsendes.

Detaljerede tabeller og resultater bør ikke hjemsendes

Tabeller og resultater der indeholder information, der gør det muligt at identificere enkelt individer, må ikke hjemsendes. Som udgangspunkt er det derfor ikke tilladt at hjemsende tabeller og resultater, der indeholder celler med mindre end fem observationer.

Indeholder et eller flere felter i en tabel færre end fem observationer skal du overveje at aggregere disse tabeller yderligere, alternativt blænde celler med få observationer før tabellen

hjemsendes. Afblanding kan fx ske ved at slette cellens indhold eller erstatte det med en markering såsom "n<5".

Vær opmærksom på, at man heller ikke af en tabels marginale totaler, procenter eller rater må kunne udlede om en given information stammer fra for få observationer. Fx må der ikke af en tabels summer kunne udledes hvor mange personer, der er blevet diskretioneret. Man må med andre ord ikke kunne "regne tilbage".

Modsat mikrodata er det, når det gælder tabeller og resultater, ikke muligt at lave helt faste regler for, hvad der må hjemsendes og hvad der ikke må. Det er med andre ord tilladt at hjemsende tabeller, figurer mv., der har en større detaljeringsgrad end de tommelfingerregler som vi anbefaler, hvis det er sikret, at det ikke er muligt at identificere enkeltpersoner i det hjemsendte output. Fx er det OK at hjemsende medianen, selv om det reelt er mikrodata, hvis det er sikret, at ingen personer kan genkendes i materialet. Det er dog vigtigt at understrege, at al hjemsendelse sker på eget ansvar.

Procedurer og sanktioner ved et brud på datafortroligheden

Hvis vi konstaterer et brud på datafortroligheden, skelner vi mellem hjemsendelse af detaljerede tabeller og resultater og hjemsendelse af mikrodata.

Hvis vi konstaterer, at der er hjemsendt mikrodata, lukker vi for institutionens adgang og følger proceduren beskrevet nedenfor.

Hvis vi konstaterer, at der er hjemsendt detaljerede tabeller eller resultater, lukker vi IKKE for adgangen, men kontakter forskeren for en redegørelse, vejleder forskeren i reglerne for hjemsendelse af data og lukker herefter sagen. Kun hvis der er tale om meget detaljerede tabeller, som har karakter af mikrodata, dvs. langt de fleste celler indeholder 1-taller, vil vi overveje at behandle sagen som hjemsendelse af mikrodata.

Sanktioner rammer både den enkelte forsker og den autoriserede institution

Sundhedsdatastyrelsen autoriserer institutioner til at kunne få adgang til Forskermaskinen i forbindelse med konkrete nærmere afgrænsede analyseprojekter. De autoriserede institutioner er ansvarlige for, at de brugere, der arbejder under autorisationen, har de nødvendige kompetencer, og at de er godt bekendt med reglerne om datafortrolighed mv.

I hver autoriseret institution er der derfor en autorisationsansvarlig, der har ansvaret for - og fører tilsyn med - at brugere, der er tilknyttet deres autorisationen, er bekendt med Sundhedsdatastyrelsens datafortrolighedsregler, som de er beskrevet i autorisationsaftale, bruger-aftaler samt generelle vilkår for Forskerservice.

Af de ovennævnte aftaler fremgår det bl.a., at en alvorlig overtrædelse af datafortrolighedsreglerne vil medføre, at brugeren, der overtræder reglerne kan udelukkes fra at anvende Sundhedsdatastyrelsens forskermaskine permanent eller i en periode. Ligeledes gælder det, at den institution, som brugeren arbejder for, kan blive udelukket fra at bruge Forskermaskinen.

En overtrædelse af datafortrolighedsreglerne kan altså få konsekvenser både for den person, der overtræder datafortrolighedsreglerne og for de projekter og brugere, der er knyttet til den autoriserede institution, som brugeren arbejder under.

Begrundelsen for dette er, at en alvorlig overtrædelse af datafortrolighedsreglerne tyder på, at den involverede autoriserede institution ikke har taget sit ansvar tilstrækkeligt alvorligt. Derfor er sanktionen ikke kun rettet mod den enkelte bruger, men mod den autoriserede institution, der har ansvaret for at pågældende bruger kender reglerne og efterlever reglerne.

Hjemsendelse af detaljerede tabeller

Ved hjemsendelse af tabeller, hvor vi vurderer, at der er fare for identifikation af enkeltindivider, bliver brugeren, der har hjemsendt filen, samt den autorisationsansvarlige for den institution som brugeren er tilknyttet, kontaktet og bedt om en redegørelse for hjemsendelsen samt bekræfte, at filen er slettet. Ligeledes vil der blive bedt om en redegørelse for, hvordan man fremadrettet vil undgå sådanne hjemsendelser.

Ved enkeltstående overtrædelser vil det for bruger og institutionen normalt ikke få konsekvenser for adgangen til Forskermaskinen, men ved gentagne regelbrud, hvor der er sendt tabeller hjem, hvor der er fare for identifikation af enkeltpersoner, kan denne praksis føre til lukning for institutionens adgang til Forskermaskinen og en videre behandling af sagen svarende til hjemsendelse af mikrodata.

Hjemsendelse af mikrodata

Hvis Sundhedsdatastyrelsen konstaterer, at en bruger har hjemsendt mikrodata, lukkes der omgående for adgangen til Forskermaskinen. Det gælder både for den bruger, som har hjemsendt mikrodata og for alle andre brugere fra den autoriserede institution.

Hvis en forsker, der er bruger under forskellige autorisationer hjemsender mikrodata fra et projekt under den ene autorisation, foretages der intet i forhold til de øvrige autoriserede institutioner som personen også er bruger under. For den person, der har hjemsendt mikrodata, lukkes adgangen til alle projekter, uanset hvilken institution de hører under.

Derefter bliver forskeren, der har hjemsendt filen, samt den ansvarlige for institutionens autorisationsaftale kontaktet. De bliver bedt om øjeblikkeligt at slette samtlige filer, der overtræder reglerne for hjemsendelse af mikrodata samt sende en bekræftelse på, at alle filer der overtræder reglerne er slettet. Dette gælder data beliggende på harddiske, mailkonti og hvor disse data ellers måtte være opbevaret.

Sundhedsdatastyrelsen vil derudover udbede sig en redegørelse for regelbruddet samt omfanget af hjemsendte filer, der bryder reglerne. Det er den autorisationsansvarlige for institutionen, der har ansvaret for, at redegørelsen bliver udarbejdet. Efterfølgende vil Sundhedsdatastyrelsen bede institutionen om snarest muligt at sende en plan for, hvordan den fremadrettet vil undgå yderligere regelbrud. Alle adgange forbliver lukket indtil både redegørelsen og planen er modtaget.

Herefter beslutter Sundhedsdatastyrelsen, hvilken sanktion institutionen skal have. Sagsbehandlingen kan tage op til 8 arbejdsdage.

Som udgangspunkt vil sanktionen være, at der lukkes for adgangen til Forskermaskinen for alle brugere tilknyttet institutionen i minimum en måned samt al adgang til Forskermaskinen for den berørte person i samme periode. Ved gentagne brud vil adgangen blive lukket i en længere periode og i særlig alvorlige tilfælde permanent.

Hvis brugeren selv gør Sundhedsdatastyrelsen opmærksom på, at vedkommende har hjemsendt mikrodata, vil det blive betragtet som en formildende omstændighed, og det vil muligvis gøre, at adgangen til Forskermaskinen vil kunne åbnes igen, så snart Sundhedsdatastyrelsen har fået klarhed over overtrædelsens omfang og karakter.

En vejledende oversigt over sanktioner fremgår af oversigten i tabel 1.

Tabel 1: Vejledende oversigt over sanktioner ved regelbrud.

Hjemsendelse af mikrodata	Sanktion mod institution			Sanktion mod forsker	
	1. gang for institution	2. gang for institution	Gentagne gange for institution	1. gang	Flere gange
a. Teknisk uheld, uden forsæt til at gøre dette	1 måneds karantæne for institutionen og for alle tilførende projekter og brugere *	2 måneders karantæne for institution	3 måneders karantæne for institution	1 måneds karantæne fra alle forskningsprojekter*	3 måneders karantæne fra alle forskningsprojekter*
b. Bevidst handling, fx for at fejlsøge data	2 måneders karantæne for institution	3 måneders karantæne for institution	3 måneders karantæne for institution	2 måneders karantæne fra alle forskningsprojekter	Permanent udelukkelse
Bevidst forsøg på at identificere individer på Forskermaskinen	3 måneders karantæne for institution	3 måneders karantæne for institution	Permanent udelukkelse	Permanent udelukkelse	Kan ikke forekomme
Password eller adgang videregivet af autoriseret person	1. gang for institution	2. gang for institution	Gentagne gange for institution	1. gang	Flere gange
Med vilje	1 måneds karantæne for institution	2 måneders karantæne for institution	Permanent udelukkelse	3 måneders karantæne for institution	Permanent udelukkelse

* Det betragtes som en formildende omstændighed, hvis institutionen selv underretter om fejlen

Hvis der sker brud på datafortroligheden

Hvis der ved en fejl sendes mikrodata hjem så kontakt Sundhedsdatastyrelsen hurtigst muligt ved at skrive til Forskerservice gennem vores henvendelsesformular, som du finder på vores hjemmeside under [lgangværende aftale](#).

Orienter hurtigst muligt Sundhedsdatastyrelsen ved et regelbrud

Det betragtes som en formildende omstændighed, at institutionen selv rapporterer regelbruddet.

Beskriv i henvendelsen, hvornår data er sendt hjem samt omfanget af hjemsendelsen. Adgangen til Forskermaskinen vil fortsat blive lukket, mens vi afventer en redegørelse og plan for forebyggelse.

Orienteres der hurtigt om et regelbrud, betragtes det som en formildende omstændighed

Ved hjemsendelse af tabeller eller resultater hvor vi vurderer, at der er fare for identifikation af enkeltindivider, bliver brugeren, der har hjemsendt filen, samt den autorisationsansvarlige for den institution som brugeren er tilknyttet, kontaktet og bedt om en redegørelse for hjemsendelsen samt bekræfte, at filen er slettet. Derudover skal der redegøres for, hvordan man fremadrettet vil undgå at hjemsende detaljerede tabeller eller resultater. Ved enkeltstående overtrædelser vil det for bruger og institutionen normalt ikke få konsekvenser for adgangen til Forskermaskinen.