

## Sammenstilling av data med Reseptregisteret

### Innholdsfortegnelse

1	Formål	1
2	Målgruppe	1
3	Forutsetninger	1
4	Prosedyre	1
4.1	Preprosessering og tilrettelegging av datafiler hos dataleverandør	1
4.2	Tildeling av prosjektspesifikke løpenummer	3
4.3	Kvalitetskontroll av filer	3
4.4	Tilgjengeliggjøring til søker	3
5	Relevante dokumenter	3
6	Flytskjema for sammenstilling av data med Reseptregisteret	4

## 1 Formål

Sikre at sammenstillinger med Reseptregisteret gjennomføres på en trygg måte.

## 2 Målgruppe

Saksbehandlere og teknisk personell ved Reseptregisteret. Prosjektledere som skal søke om data fra Reseptregisteret sammenstilt med andre datakilder og dataleverandører av data som skal inngå i sammenstillingen.

## 3 Forutsetninger

Søknaden skal være behandlet og godkjent i henhold til RF-RD-AR-002 Saksbehandling av søknader om tilgang til data og biologisk materiale.

Tilgjengeliggjøring av de ferdig sammenstilte filene, fakturering, oppfølging og avslutning av saken skjer også i henhold til RF-RD-AR-002 Saksbehandling av søknader om tilgang til data og biologisk materiale.

## 4 Prosedyre

Data som skal sammenstilles med reseptregisterdata kan være fra ulike dataleverandører f.eks. lovbestemte helseregistre, helseundersøkelser eller fra forskers egeninnsamlede data.

Koblingsnøkkel er fødselsnummer, evt. helsepersonellnummer. Denne må pseudonymiseres med samme pseudonymiseringsnøkkel som i Reseptregisteret før sammenstilling kan skje. Se nærmere beskrivelse i pkt. 4.1 og flytskjema i pkt. 6.

### **4.1 Preprosessering og tilrettelegging av datafiler hos dataleverandør**

Datauttrekk forberedes for sammenstilling hos dataleverandøren. Fødselsnummer/helsepersonellnummer (nøkkelfil) og opplysningene fra dataleverandøren som skal inngå i sammenstillingen (datafil) må splittes og sendes som separate filer med felles løpenummer per individ i alle filer. Dataleverandøren som definerer utvalget i sammenstillingen distribuerer utvalget

## Sammenstilling av data med Reseptregisteret

til alle datakilder i sammenstillingen. De andre dataleverandørene trenger ikke å sende nøkkelfil til Reseptregisteret dersom dataleverandøren har benyttet samme løpenummer i alle de utsendte nøkkelfilene, og datakildene som inngår benytter det tilsendte løpenummeret.

### 4.1.1 Nøkkelfilen

Nøkkelfilen skal inneholde gyldig fødselsnummer (11 siffer) eller helsepersonellnummer og løpenummer, én linje per individ. Nøkkelfilen må ha format semikolon-, komma- eller tabulatorseparert og må ha kolonneoverskrifter. Fast bredde kan ikke brukes.

Reseptregisteret skal aldri ha tilgang til informasjon om fødselsnummer/ helsepersonellnummer. Nøkkelfilen må derfor krypteres ved hjelp av et preprosesseringsprogram før oversendelse til Reseptregisteret. Ved bruk av programmet formateres nøkkelfilen til XML-format og krypteres. Preprosesseringsprogrammet fås sammen med et hjelpeprogram (veiviser) for tilrettelegging nøkkelfil ved henvendelse til [reseptregister.data@fhi.no](mailto:reseptregister.data@fhi.no).

### 4.1.2 Datafilen

Datafilen(e) skal inneholde løpenummer og opplysningene fra dataleverandør som skal inngå i sammenstillingen, denne kan inneholde flere linjer per individ. Dataleverandør må benytte samme løpenummer per individ i nøkkelfil og i datafilen(e).

Det godtas oversendelse av følgende filformater: sav (SPSS), dta (Stata) og separert tekstfil (csv). Dersom det skal sendes filer i annet format må dette avtales med Reseptregisteret.

### 4.1.3 Oversendelse av nøkkelfil, datafil(er) og informasjon om filene

XML-filer fra preprosesseringsprogrammet (nøkkelfilen) samt datafil(er) med samme løpenummer per individ sendes som passordbeskyttet ZIP-fil via fildelingstjenester, på DVD/CD, rekommandert per post eller som vedlegg til e-post til [reseptregister.data@fhi.no](mailto:reseptregister.data@fhi.no). Ved forsendelse per post skal datafil(ene) og nøkkelfilen sendes i to separate forsendelser.

Dataleverandør(ene) må sende Reseptregisteret en oversikt over hvilke variabler som inngår i datafilen(e) samt en beskrivelse av disse variablene. Det skal også følge med dokumentasjon for kolonnerekkefølge og antall rader på datafilen(e).

Ved oversendelse av separerte tekstfiler (csv) skal det i tillegg sendes informasjon om hvilken separator som er benyttet, f.eks. komma, semikolon eller tab. Merk at separatortegnet utelukkende skal benyttes som separator i datasettet. Det skal ikke inngå i dataverdier. Det må også gis informasjon dersom det er benyttet apostrofer ( ' ' eller " "). Det må i tillegg gis informasjon om hver enkelt variabel for å sikre at filen leses korrekt. Følgende datatyper (formater) er vanligvis aktuelle: Heltall, desimaltall, streng og dato.

## Sammenstilling av data med Reseptregisteret

Eksempel på informasjon som skal sendes om hver enkelt variabel:

Variabelnavn	Datatype (format)	Beskrivelse	Variabellengde
Alder	Heltall (integer)	Alder på diagnosetidspunktet	3
Status	Streng	B = I live og bosatt i Norge, D = Død, blank = «lost to follow-up»	1
Diagnosedato	Dato	Dato for diagnose. Angitt som 15.mm.åååå	10

Datafilinnholdsoversikten kan sendes sammen med datafilene eller med e-post til [reseptregister.data@fhi.no](mailto:reseptregister.data@fhi.no).

### 4.2 Tildeling av prosjektspesifikke løpenummer

All pseudonymisering foregår hos SSB som er tiltrudd pseudonymforvalter for Reseptregisteret. Ved mottak av nøkkelfilen ved FHI sender IT-personell ved Reseptregisteret de krypterte XML-filene (nøkkelfilen) som er behandlet med Reseptregisterets preprosesseringsprogram til SSB på lukket nett mellom Folkehelseinstituttet og SSB. Via dette nettet går filene direkte til SSBs pseudonymiseringsløsning. Pseudonymiseringen skjer ved helautomatiske prosesser og ihht. skriftlig avtale mellom Folkehelseinstituttet og SSB. Pseudonymiserte filer sendes automatisk tilbake til Reseptregisteret på lukket nett, transportkryptert for lesing kun av Folkehelseinstituttet. Hvert fødselsnummer og evt. helsepersonellnummer blir deretter tildelt et prosjektspesifikt løpenummer ved FHI.

IT-personell ved Reseptregisteret gjør uttrekk av opplysninger og variabler fra Reseptregisteret ihht. godkjent søknad, vilkår og tillatelser. I tillegg blir alle løpenummer i alle tilsendte datafiler fra alle dataleverandører som skal inngå i sammenstilling erstattet med de prosjektspesifikke løpenumrene. De samme løpenumrene blir benyttet i datafilen(e) fra Reseptregisteret.

### 4.3 Kvalitetskontroll av filer

Før tilgjengeliggjøring kontrolleres det at alle filer er lesbare og inneholder alle variabler. Filer som ble sendt inn til sammenstilling, skal fortsatt ha samme antall rader, og variablene skal ligge i samme kolonnerekkefølge, som oppgitt fra dataleverandøren.

### 4.4 Tilgjengeliggjøring til søker

Filene til sammenstilling sendes som kryptert ZIP-fil med passord på e-post, DVD/CD eller annen sikker dataoverføring til prosjektleder. Passordet overleveres separat, for eksempel per telefon eller SMS. Beskrivelse av uttrekket fra Reseptregisteret vedlegges. FHI dokumenterer hvilke prosjekter som har fått tilgang til opplysninger fra Reseptregisteret både i interne systemer, og i en allment tilgjengelig oversikt på Reseptregisterets nettsider.

## 5 Relevante dokumenter

RF-RD-AR-002 Saksbehandling av søknader om tilgang til data og biologisk materiale

## Sammenstilling av data med Reseptregisteret

### 6 Flytskjema for sammenstilling av data med Reseptregisteret

