|  |
| --- |
| 10. februar 2023  Forskerservice  Sundhedsdatastyrelsen  Ørestads Boulevard 5  2300 København S  FSEID-00006014 |

## Bilag B til aftale om dataadgang

FSEID-00006014

Dataspecifikation

Nærværende bilag er en del af aftale om dataadgang for FSEID-00006014. Bilaget beskriver følgende:

1. Type af dataadgang
2. Datatype
3. Registre
4. Egne data
5. Afgrænsning af population(er)
6. Afgrænsning af dataadgang
7. Bemærkninger til dataadgangen
8. Adgang til applikationsserver og analyseværktøjer
9. Betaling

Hvis der sker ændringer i ovenstående punkter opdateres nærværende bilag uden, at der skal indgås fornyet aftale om dataadgang.

# Type af dataadgang

**Views:**

* CAR
* CPR
* DAR
* LAB
* LPR
* LPR-PSYK
* LPR3\_F
* LSR
* DDV\_SSI
* OVD\_SSI
* MFR
* SSR
* PCR

# Datatype

Individdata med krypterede cpr-numre

# Registre

* Cancerregisteret (CAR)
* CPR-registeret (CPR)
* Det Psykiatriske Centrale Forskningsregister (PCR)
* Det Danske Vaccinationsregister - COVID-19-vaccinationsdata fra SSI (DDV\_SSI)
* Dødsårsagsregisteret (DAR)
* Laboratoriedatabasens Forskertabel (LAB)
* Landspatientregisteret (LPR)
* Landspatientregisteret - Psykiatri (LPR-PSYK)
* Lægemiddelstatistikregisteret (LSR)
* Overvågningsdata - COVID-19-testsvar fra SSI (OVD\_SSI)
* Sygesikringsregisteret (SSR)
* Landspatientregisteret - forskerrettet datamodel (LPR3\_F)
* Medicinsk Fødselsregister (MFR)

# Egne data

Ja.

Forskerservice afventer, at du leverer data. Dataskal uploades til Sundhedsdatastyrelsen via vores upload-løsning. Når du er klar til at uploade dit datasæt, skal du give Forskerservice besked om dette, hvorefter du vil modtage en mail med et unikt upload-link.

Ved at følge linket vil du komme til Sundhedsdatastyrelsens upload-side, hvor du bliver guidet til, hvordan du uploader dit datasæt til Forskerservice. Når du uploader datasættet, bliver det automatisk krypteret. Det er derfor ikke nødvendigt, at du selv krypterer datasættet.

Formatet på dit datasæt skal være et SAS-datasæt (.sas7bdat) eller en tekstfil (.txt eller .csv). Uploader du mere end ét datasæt, må du meget gerne pakke datasættene i én zip-komprimeret mappe.

Bemærk, at det indsendte datasæt skal stemme overens med de angivne oplysninger i bilag C. Det betyder bl.a. at filformat, delimiter, navngivning af tabeller og angivelse af særlige variable skal svare til oplysningerne om dine egne data i bilag C.

Såfremt dit datasæt ikke er i overensstemmelse med bilag C vil vi, for at undgå misforståelser, bede dig om at indsende et revideret datasæt.

Hvis dit datasæt indeholder CPR-numre, vil kun CPR-numre der kan genfindes i CPR-registeret blive krypteret i forbindelse med indlæsning på Forskermaskinen. Rækker med ugyldige CPR-numre vil stadig være i dit datasæt, men det ugyldige CPR-nummer fjernes som udgangspunkt.

# Afgrænsning af population

Forsker leverer populationen.

Det fremgår af ansøgningen, at populationen vil bestå af personer med nydiagnosticeret type 2 diabetes, og som er rekrutteret til at deltage i projektet Dansk Center for Strategisk Forskning i Type 2 Diabetes (DD2).

Ved oprettelse af dataadgangen er der tale om ca. 10.000 individer.

Forsker vil efterfølgende opdatere populationen i takt med den løbende rekruttering til projektet. Populationen forventes i alt at omfatte ca. 50.000 individer

Vi forudsætter en population i denne størrelsesorden. CPR-numrene vil blive kontrolleret for dubletter og holdt op imod CPR-registerets persontabel, når de modtages. Vi giver kun adgang til data for CPR-numre, der kan genfindes i CPR-registeret.

# Afgrænsning af dataadgang

**Cancerregisteret (CAR)**

For populationen gives adgang til oplysninger om alle incidente cancerdiagnoser, uden afgrænsning.

**CPR-registeret (CPR)**

For populationen gives adgang til følgende oplysninger fra CPR:

* Aktuel status, køn og fødselsdato
* Aktuel og historisk civilstand
  + *Der leveres kun rækker hvor c\_annkor = ’ ’ for t\_civil\_hist*
* Perioder med bopæl i en dansk kommune (ekskl. Grønland), afgrænset ved dansk\_ophold\_periode

**Det Danske Vaccinationsregister - COVID-19-vaccinationsdata fra SSI (DDV\_SSI)**

For populationen gives adgang til alle covid-19-vaccineoplysninger i DDV\_SSI.

**Dødsårsagsregisteret (DAR)**

For populationen gives adgang til alle registrerede dødsfald i DAR, uden afgrænsning.

Der gives desuden adgang til oplysninger om midlertidige dødsårsager i perioden fra året efter seneste opdateringsår i tabellen t\_dodaarsag\_2, afgrænset ved:

* year(statdate) i schemamortality > max(year(d\_statdato)) i t\_dodsaarsag\_2

*Der medgives en oversigtstabel over antal personer i seneste validerede årgang samt næst følgende årgange.*

**Laboratoriedatabasens Forskertabel (LAB)**

For populationen gives adgang til alle laboratoriesvar i LAB, uden afgrænsning.

**Landspatientregisteret (LPR), - Psykiatri (LPR-PSYK) og - forskerrettet datamodel (LPR3\_F)**

For populationen gives adgang til alle kontakter i LPR og LPR-PSYK, inklusive uafsluttede kontakter, fra offentlige såvel som private sygehuse, samt alle kontakter i LPR3\_F, uden periodeafgrænsning.

*Kontakter fra LPR og LPR-PSYK, som er påbegyndt eller afsluttet efter 2018, kan ikke forventes at være komplette, på grund af løbende overgang til indberetning i LPR3-format*

*Vær opmærksom på, at kontakter der var igangværende ved overgangen til indberetning i LPR3-format, som udgangspunkt er registreret som afsluttet, med udskrivningsmåde (c\_udm) = 'M' (teknisk afslutning). Der findes dog fortsat uafsluttede kontakter i LPR2-format.*

*Det er i de tilgængelige data ikke muligt at afgøre hvorvidt kontakter med udskrivningsmåde ’M’ samt de uafsluttede kontakter reelt er afsluttet eller igangværende efter overgangen.*

*For tabellen nyt\_helbredsforloeb i LPR3\_F gives adgang til alle rækker hvor enten dw\_ek\_forloeb\_nyt eller dw\_ek\_forloeb\_forrige indgår i tabellen forloeb.*

**Det Psykiatriske Centrale Forskningsregister (PCR)**

For populationen gives adgang til alle psykiatriske kontakter i PCR, uden periodeafgrænsning.

**Lægemiddelstatistikregisteret (LSR)**

For populationen gives adgang til alle indløste recepter, hvor følgende atc-koder er registreret i variablen atc:

* A01\* - A16\*
* B01\* - B06\*
* C01\* - C10\*
* D01\* - D11\*
* G01\* - G04\*
* H01\* - H05\*
* J01\* - J07\*
* L01\* - L04\*
* M01\* - M05\*
* M09\*
* N01\* - N07\*
* R01\*
* R03\*
* R05\* - R07\*
* S01\* - S03\*

**Overvågningsdata - COVID-19-testsvar fra SSI (OVD\_SSI)**

For populationen gives adgang til alle registreringer i OVD\_SSI.

**Sygesikringsregisteret (SSR)**

For populationen gives adgang til alle registrerede ydelser i SSR, der opfylder følgende kriterier:

* Nedenstående specialekoder er registreret i tabellen t\_ssik (1990-2013) i variablen *c\_speciale* eller i tabellen t\_ssr (2014-) i variablen *yderens\_speciale*:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Specialekoder** | | | | |
| 02  04  06  07  08  09  11  15  16  17  18  19 | 20  21  24  25  26  28  35  39  41  42  43  45  46  47  49 | 50  51  52  53  54  55  56  58  59  61  62  63  64  68 | 71  75  76  79  80  81  82  83  84  85  86  87  88  89 | 90  91  92  93  94  95  96  97  98  99 |

*Data i sygesikringsregisteret ligger i årstabeller som er afgrænset ved afregningsår. Skellet mellem SSR og SSR2, går ved afslutningen af afregningsåret 2013.*

**Medicinsk Fødselsregister (MFR)**

For populationen gives adgang til alle oplysninger fra tabellen t\_lfoed i MFR, hvor personen er registreret som barn, afgrænset ved:

* + v\_bcpr\_encrypted = population

# Bemærkninger til dataadgangen

**Kobling mellem tabellerne i LPR3\_F**

Der kan generelt kobles mellem forskellige tabeller i LPR3\_F via de enslydende dw\_ek-nøglevariable.

* Særligt for tabellerne procedurer\_andre og procedurer\_kirurgi:
  + Enkelte procedurer er ikke knyttet til en kontakt men er i stedet knyttet direkte til patientens forløb. Disse procedurer kaldes forløbsprocedurer, og er patientrelaterede procedurer, der foregår uden samtidig patientkontakt. Forløbsprocedurer kan kobles til tabellen forloeb eller tabellen kontakter via nøglevariablen dw\_ek\_forloeb.
* Særligt for tabellen betaler
  + Betalingsoplysninger kan være knyttet til enten en kontakt, en kontaktprocedure med selvstændig betaler eller en forløbsprocedure. Betalingsoplysninger for kontakter kobles til tabellen kontakter via nøgelvariablen dw\_ek\_kontakt. Betalingsoplysninger for kontaktprocedurer med selvstændig betaler og forløbsprocedurer kobles til tabellerne procedurer\_kirurgi og procedurer\_andre via nøgelvariablen dw\_ek\_procedureregistrering.
* Særligt for tabellen henvisning\_tillaeg
  + Der kan principielt indberettes henvisningsårsager for både kontakter og forløb til LPR3. Derfor kan der være tillægskoder til henvisningsårsager som er knyttet til enten en kontakt eller et forløb i tabellen henvisning\_tillaeg. Tillægskoder til henvisningsårsager kan kobles til tabellen forloeb eller tabellen kontakter via nøglevariablene dw\_ek\_forloeb og dw\_ek\_kontakt.
* Særligt for tabellen resultater
  + Resultatindberetninger kan principielt være knyttet til enten en kontakt eller et forløb. Resultatindberetninger kan kobles til tabellen forloeb eller tabellen kontakter via nøglevariablene dw\_ek\_forloeb og dw\_ek\_kontakt.

**Kobling mellem LPR2-kontakter i LPR og LPR3-forløb i LPR3\_F**

Kontakter der var igangværende på tidspunktet for overgang til indberetning i LPR3-format bør være overført til et LPR3-forløb på en måde der gør kobling af det nye LPR3-forløb og den gamle LPR2-kontakt mulig.

Entydig kobling mellem LPR3-forløb og den modsvarende LPR2-kontakt, defineres ved følgende:

|  |  |
| --- | --- |
| **Variable i tabellen t\_adm i LPR2** | **Variable i tabellen forloeb i LPR3\_F** |
| v\_cpr\_encrypted | cpr\_encrypted |
| c\_udm (M – ’afsluttet til lpr3’) | henvisningsmaade (ALDA99 – ’overført fra lpr2’) |
| c\_sgh og c\_afd | sorenhed\_hen  *(se tabellen organisationer for oversættelse af sor-kode til shak-kode)* |
| d\_uddto og v\_udtime | dato\_henvisning  tidspunkt\_henvisning |

Ca. 58% af LPR3-forløb med henvisningsmåden 'overført fra LPR2' kan kobles til en LPR2-kontakt ud fra ovenstående krav.

Fjernes kravet om enslydende timeangivelse mellem v\_udtime og tidspunkt\_henvisning, kan ca. 81% af LPR3-forløbene kobles til en LPR2-kontakt.

Du kan læse meget mere om overgangen fra LPR2 til LPR3 i 'Datakvalitetsrapport om LPR 2018 – overgangen fra LPR2 til LPR3' her: https://sundhedsdatastyrelsen.dk/da/registre-og-services/om-de-nationale-sundhedsregistre/sygedomme-laegemidler-og-behandlinger/landspatientregisteret/landspatientregisteret-moderniseres

**Om MiniPas i Landspatientregisteret og Landspatientregisteret – Psykiatri**

MiniPas er oprettet i 2002 og er et nationalt register over aktiviteten ved danske private sygehuse og klinikker (og få offentlige institutioner). Registeret indeholder en beskrivelse over hvilke sygdomme, som patienterne har haft ved kontakten til sygehuse, og hvordan disse sygdomme er blevet behandlet. Hertil kommer en række oplysninger om patienternes bopæl, køn, alder, indlæggelses- og udskrivningstidspunkt mv.

MiniPas er formelt en del af Landspatientregisteret (LPR) og Landspatientregisteret – Psykiatri (LPR-PSYK), men da de private sygehuse indberetter til den selvstændige databaseenhed MiniPas, er denne del af Landspatientregisteret dokumenteret som et særskilt register. Der er - med få undtagelser - overensstemmelse mellem dataindhold og struktur i MiniPas og i Landspatientregisteret. De private sygehuse og klinikker skal indberette al den aktivitet, der ikke er omfattet af overenskomster på praksisområdet, til Sundhedsdatastyrelsen.

**Om kontaktnøglen i Landspatientregisteret**

Variablene c\_sgh, c\_afd, v\_cpr, c\_pattype, d\_inddto, v\_indtime, v\_indminut danner ’kontaktnøglen’ i LPR. Variablene er tilføjet dataudtrækket, således det evt. på et senere tidpunkt, er muligt at udtrække de samme kontakter fra LPR.

**Om datakompletheden og validiteten i midlertidige dødsårsager**

Data i SCHEMAMORTALITY er midlertidige og til forskel fra de publicerede årgange i Dødsårsagsregisteret, mindre komplette, ligesom de ikke har ikke undergået fuld validering.

Komplethed

Når Dødsårsagsregisteret opdateres med en ny årgang, vil denne indeholde indberetninger for alle dødsfald på nær et par procent. Når der videregives nyere data end i de publicerede årgange fra Dødsårsagsregisteret, vil der mangle betydeligt flere, da mange indberetninger af særligt dødsattestens kliniske oplysninger, kan være forsinkede.

Validitet

Til forskel fra data i de afsluttede årgange, har data i SCHEMAMORTALITY ikke undergået helt samme validering. Forskellen beskrives herunder.

Data i de afsluttede årgange af Dødsårsagsregisteret har inden offentliggørelse gennemgået stor validering. Vigtigst er, at den tilgrundliggende dødsårsag bestemmes efter internationale regler og kan ende med at afvige fra den, der oprindeligt er indberettet. Til dette benyttes dels en automatisk validering (algoritmen IRIS), som i første omgang bestemmer en tilgrundliggende dødsårsag alene ud fra de indberettede dødsårsager. Men ud over dette er i hver publiceret årgang ca. 10.000 indberetninger (ud af 50-60.000) valideret manuelt. Dette sker via løbende monitorering og bestemmes ud fra evt. afvigelser fra det normale årsagsbillede.

**Vigtig information om opdatering af registre:**

Indberetningen til Landspatientregisteret er overgået til nyt dataformat (LPR3) mellem 1. januar 2019 og 3. marts 2019. Derudover er Sundhedsdatastyrelsens Elektroniske Indberetningssystem (SEI) blevet moderniseret. Dette har følgende betydning for indeværende aftale om dataadgang:

* **Overgang til LPR3**

Efter overgangen til LPR3 følger nu en fase hvor de registre der bygger på oplysninger fra LPR3 skal tilrettelægges og gøres tilgængelige for Forskerservice på en ny måde. I denne fase vil der ikke kunne leveres opdaterede oplysninger fra de registre, der modtager data fra Landspatientregisteret. Det drejer sig i indeværende aftale om følgende registre:

* Medicinsk Fødselsregister (MFR)
* Forskerservice vil melde ud via vores nyhedsbrev, når data klar til brug i det nye dataformat. [Tilmeld dig ForskerserviceNyt her](http://sundhedsdatastyrelsen.us15.list-manage.com/subscribe?u=a2666d6e57e0a06dfeb9beb46&id=c03aaf8d29).
* Hvis du ønsker at få adgang til opdaterede data fra disse registre, skal den nuværende aftale om dataadgang opdateres. Aftalen skal opdateres, da dataformatet og selve oprettelsen af dataadgang/views er væsentlig anderledes.

Du vil ved opdatering af aftalen blive faktureret for det antal arbejdstimer, Forskerservice skal bruge til at behandle opdateringen, hvilket inkluderer opdatering af aftalen om dataadgang, programmering af dataadgang/views samt eventuel kontrol af egne udtræksprogrammer.

**Om views**

Views bliver automatisk opdateret når registeret bliver opdateret på Forskermaskinen, så du løbende får adgang til de nyeste registreringer (alt afhængig af den øvrige afgrænsning af din dataadgang selvfølgelig). Delvist undtaget her fra er data, der ligger i årstabeller, som f.eks. Sygesikringsregisteret. Her skal du kontakte Forskerservice for at få opdateret data efter et årsskifte.

# Adgang til applikationsserver og analyseværktøjer

* FM-SAS-02 – Analyseværktøjet er SAS
* FM-R-01 – Analyseværktøjet er R
* FM-Stata-02 – Analyseværktøjet er STATA

Brugerne på projektet skal have følgende individuelle adgange:

|  |  |  |  |
| --- | --- | --- | --- |
| **Email eller brugernavn** | **SAS** | **R** | **Stata** |
| **fskDiaHed** | **X** | **X** |  |
| **fskFrePag** | **X** | **X** | **X** |
| **fskHelMat** | **X** | **X** |  |
| **fskLarPed** | **X** | **X** |  |
| **fskReiNie** | **X** | **X** |  |
| **fskSiaKro** | **X** | **X** |  |

# Betaling

Al fakturering i forbindelse med projektet sker til det autoriserede forskningsmiljø:

Klinisk Epidemiologisk Afdeling

Aarhus Universitetshospital

Betalingsansvarlig: Wencke Bentsen

Betalings-ID hos Forskerservice: BET-00002994

Projektet vil blive faktureret løbende ved udgangen af hver måned. Fakturaen vil indeholde betaling for den tid, som Forskerservice har brugt på projektet den pågældende måned. Dette omfatter den tid, som Forskerservice bruger på at behandle ansøgningen og programmere dataudtrækket. Det omfatter også eventuel rådgivning, som forekommer efter levering af data.

Alle henvendelser efter indgåelse af aftale afregnes til timepris.

Du kan orientere dig om den aktuelle timepris på vores hjemmeside.