

Krav til udtræksbeskrivelse

Statistikudtræk

For at sikre, at du får de data, du har brug for og for at lette sagsgangen og arbejdet med udtrækket, skal du ved alle ansøgninger til Forskerservice vedhæfte en særskilt udtræksbeskrivelse.

Når du søger om statistikudtræk fra Forskerservice, leveres data så det på ingen måde er personhenførbart. Dette betyder, at Forskerservice som udgangspunkt slører felter med 0-4 observationer, samt observationer der på anden måde vurderes personhenførbart.

Du kan nedenfor se en række punkter, som udtræksbeskrivelsen skal indeholde. Er du i tvivl, er du velkommen til at kontakte Forskerservice.

Til sidst i dokumentet findes et eksempel på en udtræksbeskrivelse til et simpelt statistikudtræk.

Bemærk, at vores behandling af ansøgningen vil tage udgangspunkt i din udtræksbeskrivelse. Der kan dog i hvert enkelt tilfælde være en grænse for, hvor omfattende et produkt vi kan tilbyde. Herunder hvor mange tabeller vi kan tilbyde at danne inden for hver ansøgning.

1. Oplysninger om eventuel population/kohorte

Hvis der skal foretages dataudtræk ud fra en afgrænset population, skal du oplyse de nøjagtige kriterier for afgrænsningen. Det kan fx være:

- ▶ Registre
- ▶ Diagnoser
- ▶ Procedurer (operationer, undersøgelser og behandlinger)
- ▶ Dødsårsager
- ▶ Aldersintervaller
- ▶ Køn
- ▶ Tidsintervaller (datoer)

2. Tidsafgrænsning

Du skal oplyse inden for hvilken tidsperiode udtrækket skal strække sig. Søger du om data fra Landspatientregisteret vil vi på forhånd opfordre dig til at overveje om tidsperioden skal afgrænses på indskrivnings- eller udskrivningsdatoen for den enkelte hospitalskontakt.

Ønsker du fx alle kontakter i perioden 2000-2012, skal periodeafgrænsningen se således ud:

d_inddto <= 2012 OG (d_uddto >= 2000 ELLER d_uddto = .)

På denne måde inkluderes alle kontakter i perioden, også de kontakter som er startet før 2000 eller afsluttet efter 2012.

3. De ønskede tabeller og særlige afgrænsninger for hver kategori

Du skal oplyse, hvilke tabeller du ønsker, at Forskerservice danner og meget gerne vise et eksempel. Særlige afgrænsninger for hver kategori skal angives.

4. Beskrivelse af hvor de ønskede oplysninger skal findes

Du skal beskrive fra hvilket register, og hvilke variable Forskerservice skal finde de ønskede oplysninger.

Du kan med fordel udfylde vores bestillingsark med de variable, du ønsker, at vi anvender. Hvis du er i tvivl, er du altid velkommen til at kontakte Forskerservice.

Eksempel på udtræksbeskrivelse

1. Oplysninger om population/kohorte

Alle kvinder, der har modtaget diagnosen DC50* (brystkræft) i Landspatientregisteret.

Diagnosen skal være givet i perioden 01.01.2000 – 31.12.2010 afgrænset på indskrivningsdato (d_inddto).

Diagnosen skal være noteret i variablen c_diag i tabellen t_diag i Landspatientregisteret, og diagnosen skal være noteret som enten aktions- eller bidiagnose.

Personerne må ikke have modtaget diagnosen DC50* (brystkræft) før 01.01.2000.

2. Tidsafgrænsning

Der ønskes oplysninger for perioden 01.01.2000 – 31.12.2010 afgrænset på indskrivningsdato (d_inddto).

3. De ønskede tabeller og særlige afgrænsninger for hver kategori

Tabel 1:

Tabellen skal vise antallet af patienter, der har modtaget diagnosekoden DC50* som enten aktions- eller bidiagnose, inddelt i periode og på sygehus for kontakten.

Periode ¹	Sygehus ²	Antal patienter ³

1. Periode, inddelt på indskrivningsår
2. Sygehuset for kontakten (c_sgh)
3. Antal patienter.

Afgrænsninger:

- En person, kan kun indgå én gang under hver periode og hvert sygehus.
- Hver person kan indgå flere gange i tabellen, hvis personen har haft en kontakt på flere sygehuse.
- Hver person kan indgå flere gange i tabellen, hvis personen har haft kontakter med diagnosekoden DC50* i flere perioder.

Tablet 2:

Tabellen skal vise alderen på patienter, der får diagnosen DC50* som aktions- eller bidiagnose for første gang.

Periode ¹	Alder ²	Antal patienter ³

1. Periode, inddelt på indskrivningsår
2. Alder på indskrivningstidspunktet. Inddelt i følgende aldersgrupper:
 - a) <30 år
 - b) 30 – 39 år
 - c) 40 – 49 år
 - d) 50 – 59 år
 - e) 60 – 69 år
 - f) 70 – 79 år
 - g) >79 år
3. Antal patienter.

Afgrænsninger:

- En person kan kun indgå én gang i tabellen. Der kigges på den første kontakt afgrænset på indskrivningsdatoen.

4. Beskrivelse af hvor de ønskede oplysninger skal findes

Der ønskes registreringer fra Landspatientregisteret.

Diagnosekoden DC50* findes i variabelen c_diag i tabellen t_diag. Diagnosen skal være registreret som enten aktions- eller bidiagnose, dvs. c_diagtype = A eller B.

Periode afgrænses på indskrivningsdatoen (d_inddto) i t_adm.

Sygehus findes fra c_sgh i t_adm

Alder findes fra CPR registeret, baseret på variabelen d_foddato i tabellen t_person.

(*=inklusive eventuelle underkoder)