

Att tänka på i registerstudier

Christel Häggström

Registercentrum Norr, Region Västerbotten &
Inst för folkhälsa och klinisk medicin, Umeå Universitet



Definition

Registerforskning, eller mer korrekt registerbaserad forskning, använder sig av uppgifter från *ett eller flera register som finns hos myndigheter eller andra organisationer*. Uppgifterna rör oftast individer, men kan även handla om exempelvis kommuner, fastigheter eller företag.

<https://www.registerforskning.se/>

Den särskilda satsningen syftar till att stärka och utveckla tvärvetenskaplig och tvärsektoriell *forskning baserad helt eller delvis på mikrodata från exempelvis myndighets- eller kvalitetsregister (data insamlad primärt för andra syften än forskning) och som kan länkas till data från andra källor*, till exempel forskargenererade data, biobanksdata eller surveydata.

www.vr.se

Vad menar vi med register?

- Nationella register med rapporteringsplikt
 - T.ex., Socialstyrelsens databaser, SCB register
- Nationella kvalitetsregister (och andra kvalitetsregister)
 - Ej obligatorisk rapportering → varierande kvalité
 - Primärt syfte med dessa register: vårduppföljning och förbättring
- Register hos andra nationella/regionala myndigheter ej insamlat i forskningssyfte
 - Försäkringskassan
 - Värnpliktsregistret
 - Folkhälsomyndigheten (smittsamma sjukdomar, alkohol/narkotika/tobak, folkhälsa)
 - Stadsarkivet
 - digitaliserade kyrkböcker
 - data insamlade regionalt/lokalt för hälsouppföljning
 - kostnad per patient databasen (SKR)
 - Väntetidsdatabasen (SKR)
 - GIS information
 - med mera..

← Vad vi fokuserar på idag

Dessa registerkällor kan i sin tur länkas till data/biobankat material insamlat i forskningssyfte (tex tvillingregistret, U-CAN, blodprov i VIP/VHU m fl)

Agenda

- Att tänka på inför studier med:
 - Kvalitetsregister
 - Socialstyrelsens hälsodataregister
 - SCBs register
- Att tänka på inför etikansökan:
 - Kodnyckel
 - Risker/nytta (datasäkerhet, MONA systemet)
- Att tänka på inför studier med data från flera myndigheter
 - Kodnyckelhantering
 - Länkningsprotokoll
 - RUT
- Att tänka på innan/när data kommer
 - 3 månader "hitta fel"
 - Datat == registerutdrag!
 - Statistisk mjukvara?
 - Dokumentation runt datahantering
- Sammanfattning
- Framtida tips



Register inom presentationen

- Nationella Kvalitetsregister

- Vilket data finns?
 - Diagnoser, behandling, patientinfo, nyckelindikatorer, patientrapporterade mått
- Hur ansöker man?
 - Kontakt med registerhållare/styrgrupp för hjälp med ansökan. Ansökan till CUPA (central personuppgiftsansvarig) myndighet (=registercentrum på någon sjukvårdsregion)

- Socialstyrelsens register

- Vilket data finns?
 - Hälsodata (sluten /öppenvård, läkemedel, cancer, tandhälsa, osv.), socialtjänster, dödsorsaker
- Hur ansöker man?
 - <https://bestalladata.socialstyrelsen.se/data-for-forskning/>

- SCB register

- Vilket data finns?
 - Individer (ekonomi, levnadsförhållande, utbildning, befolkning, sysselsättning, osv.), företag, fastigheter
- Hur ansöker man?
 - <https://www.scb.se/vara-tjanster/bestall-data-och-statistik/bestalla-mikrodata/>



Nationella Kvalitetsregister

Hur ansöka om data?

1. Ta kontakt med styrgrupp

Ta kontakt med registerhållaren eller någon annan i kvalitetsregistrets styrgrupp för att diskutera om det går att använda registrets data för det du har tänkt.
Ladda ner variabellistor

2. Ansök om etikprövning till Etikprövningsmyndigheten

Ange register, selektion av studiepopulation och variabelgrupper i ansökan

3. Skicka in ansökan till CPUA för registret via RCO (registercentrum/regionalt cancercentrum)

Att tänka på

1. Hur samlas data in i kvalitetsregistret:

Kan individer finnas med flera gånger?

Höger/vänster?

Upprepade procedurer?

Inklusionskriterier?

2. Kvalitet i kvalitetsregistret (certifieringsgrad, bla täckningsgrad)


Beroende på design av studie: fångar kvalitetsregistret vad du vill studera?

3. Samarbeta gärna med någon representant från kvalitetsregistret som kan svara på frågor runt variablerna!








Socialstyrelsens register

Socialstyrelsens register

- Hälsodataregister 
- Socialtjänstregister
- Övriga register

Socialstyrelsens hälsodataregister

- Cancerregistret 
- Registret för insatser i kommunal hälso- och sjukvård
- Läkemedelsregistret 
- Medicinska födelseregistret 
- Patientregistret 
- Tandhälsoregistret 

Socialstyrelsen: Registerservice

En servicefunktion som hjälper till med databeställningar från forskare, kvalitetsregister, kommuner, företag, allmänhet med flera

- Hjälp för forskare:
 - Statistikdatabasen
 - Mycket bra för att få översikt
 - Lätt att navigera
 - Bra visualisering
 - Beställning av aggregerad statistik
 - Snabbare än beställning av mikrodata
 - Beställning av data
 - Mycket användbar variabelväljare
 - Tar 6-12 mån

Prislista (alla priser är exklusive moms):

Grundavgift, standard 2 200 kr

Grundavgift, om samtliga i studiepopulationen är avlidna 1 650 kr

Timpris handläggning, 1 100 kr

Reklamation som inkommer inom reklamationstiden 0 kr

Ändra förbehållsbeslut 1 100 kr

Förlänga en sparad nyckel 0 kr

Register		Årtal		
		Från	Till	
Hälsodataregister				
<input type="checkbox"/>	Patientregistret Öppenvård			
<input type="checkbox"/>	Patientregistret Slutenvård			
<input type="checkbox"/>	Läkemedelsregistret			
<input type="checkbox"/>	Dödsorsaksregistret (inkl. avi)			
<input type="checkbox"/>	Cancerregistret			
<input type="checkbox"/>	Medicinska Födelseregistret			Mamma ▼
Socialtjänstregister				
<input type="checkbox"/>	Ekonomiskt bistånd EKB, År			Hushåll ▼
<input type="checkbox"/>	Ekonomiskt bistånd EKB, Månad			Individ ▼
<input type="checkbox"/>	Registret över insatser för barn och unga BU			
<input type="checkbox"/>	Registret över socialtjänstinsatser till äldre och personer med funktionsnedsättning SOL			
<input type="checkbox"/>	Registret över insatser enligt lagen om stöd och service till vissa funktionshindrade LSS			
<input type="checkbox"/>	Registret över tvångsvård enligt lagen om vård av missbrukare LVM			
Övriga hälsodataregister				
<input type="checkbox"/>	Tandhälsoregistret THR			
<input type="checkbox"/>	Olycksfallsregistret IDB			
<input type="checkbox"/>	Basalcellscancerregistret			
<input type="checkbox"/>	Medicinska Födelseregistret Läkemedel			

Välkommen till SIV, Socialstyrelsens Interaktiva Variabelväljare!

1

Välj dina register och ange tidsperiod
Klicka i registrets kryssruta och ange tidsperiod i till och från kolumnerna genom att skriva in år.
Glöm inte att trycka Enter efteråt.

2

Välj variabler för respektive register
Klicka på knappen "Välj variabler". När du valt dina variabler för ett register klicka på "Nästa register" för att komma till nästa register.

3

Se över din beställning, spara och skicka in
Klicka på "Se din variabellista" för att få en översikt. Är du nöjd sparar du excel-filen som en xls fil. Bifoga filen tillsammans med din beställning av registeruppgifter.

Selektion & Trunkering av koder

- När register och variabler är bestämt...
- Dags då att fundera på selektion och trunkering av intressanta koder..
- Dvs bestämma på vilken detaljnivå du behöver informationen
- Tex ICD-10 koden F22.9 :
 - Vill du veta? F22.9, F22, F20-25 Kroniskt vanföreställningssyndrom, ospecificerat
- Tex ATC koden A07AA02 (nystatin) mot tuberkolos (?):
 - Vill du veta A07AA02, A07AA, A07?

En rad i patientregistret motsvarar ett vårdtillfälle. Upp till 30 diagnoskoder kan finnas registrerade på ett och samma vårdtillfälle. Utöver diagnoskoder kan yttre orsakskoder och åtgärds-koder finnas registrerade. För att begränsa ett uttag av registeruppgifter från patientregistret kan Socialstyrelsen:

1. selektera ut de vårdtillfällen där en specifik diagnoskod, yttre orsakskod eller åtgärds-kod registrerats
2. trunkera, det vill säga begränsa detaljnivån på samtliga eller vissa av de diagnoskoder, yttre orsakskoder eller åtgärds-koder som förekommer på det valda vårdtillfället

Vad som ej finns på Sos registren

- Patientregistret innehåller EJ primärvård
- Läkemedelsregistret innehåller EJ läkemedel som ges via sjukhus
- Tandhälsoregistret innehåller EJ tandvård som ges gratis till barn och unga mellan 0 och 23 år, käkkirurgi eller tandvård som ett led i kortvarig sjukdomsbehandling.
- Med mera?



Statistiska centralbyråns (SCB) register

SCB's mikrodata

- Individregister ← Befolkningsregister: Register över totalbefolkningen (RTB) + flergenerationsregistret
- Fastighetsregister
- Företagsregister och företagsundersökningar
- Longitudinella register ← Longitudinell Integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier – LISA

RTB i forskning

- Skapa populationer =ofta kontrollpopulation i fall-kontrollstudier
- Folkbokföringsuppgifter – långa tidsserier
- Kopplingar till make/maka
- Migrationer/dödsdatum

LISA

Om en forskare begär data från LISA – bra att specificera från vilket årtal data behövs för projektet!

Variabelinnehållet i LISA kan huvudsakligen grupperas efter:

- Demografi
- Utbildning
- Sysselsättning och arbetslöshet
- Inkomst och socialförsäkring
- Familj
- Arbetsställe och företag.

Datakällor

- Registret över totalbefolkningen (RTB)
- Folk- och bostadsräkningen 1990
- Registerbaserad arbetsmarknadsstatistik (RAMS)
- Befolkningens utbildning
- Registret över personer i utbildning
- Inkomst- och Taxeringsregistret (IoT)
- Datalagret (Arbetsförmedlingens sökanderegister)
- STORE-databasen (Försäkringskassans datalager)
- Företagens ekonomi
- Ekonomisk data för kommuner
- Ekonomisk data för landsting.

Mer info om användbara variabler i LISA

https://www.scb.se/contentassets/f0bc88c852364b6ea5c1654a0cc90234/anvandbara_variabler_i_lisa_20210326.pdf

Vad kostar det och hur lång tid tar det?

Just nu, 1 nov 2022, hanteras beställningar som inkom i juni

Populationens storlek och antal registeruppgifter som efterfrågas påverkar priset marginellt så länge uppgifterna kan hämtas från ett och samma register. Normalt tar det 1–6 månader från det att du har tilldelats en handläggare till det att vi levererar mikrodata till dig.

Det är först efter en sekretessprövning som SCB har möjlighet att lämna en fastställd prisuppgift. SCB tillämpar principen om full kostnadstäckning, vilket innebär att SCB tar betalt för den faktiska kostnaden för att handlägga och producera det aktuella uppdraget. Merparten av uppdragen hamnar inom kostnadsintervallet 30 till 60 tkr. Mer information hittar du under [prissättning](#).

<https://www.scb.se/vara-tjanster/bestall-data-och-statistik/bestalla-mikrodata/>

Efter inskick till Sos och/eller SCB (separata processer)

- Kösystem för handläggare (~4-6mån)
- Kontakt med handläggare
 - Stämmer forskningsfrågan med etikillståndet?
 - Är registren/variablerna täckta av etiken?
 - Finns etik för känsliga personuppgifter inom projektet?
 - Kodnyckelhantering?
- Offert och beskrivning av uttag
 - Enbart nödvändiga variabler för att utföra forskningen på grövsta nivå
- Kontakt med handläggare vid andra myndigheter
 - Matchning med register vid andra myndigheter
- Sekretessprövning
- Beslut
- Datauttag (ca ~ 6mån från kontakt med handläggare)
- Leverans (till Mona-systemet eller krypterad digital eller fysisk fil)



Tänka på inför
EPM ansökan

Kodnyckelhantering!!

Pnr	Lopnr
195104011453	1
198008140231	2
199912130543	3
200509103487	4
	5
	6
	7

Om kodnyckeln ska sparas (tex för framtida uppdateringar)
på någon myndighet (SCB ofta) måste det finnas med i etikansökan!

Hur detaljerat data behövs på födelsedatum och ursprungsland?

7.1 Vilka risker kan ett deltagande medföra för de forskningspersoner som ingår i forskningsprojektet?

Risker kan vara av olika karaktär. Det kan exempelvis vara fysisk eller psykisk skada, smärta, obehag, integritetsintrång på kort eller lång sikt. Ange vilka risker det kan finnas för de forskningspersoner som deltar i detta forskningsprojekt.

INTEGRITETSINTRÅNG!!
Obehag?

7.2 Vilken nytta kan ett deltagande medföra för de forskningspersoner som ingår i forskningsprojektet?

Ange vilken hjälp de forskningspersoner som deltar i detta forskningsprojekt kan få som ett resultat av forskningsprojektet.

7.3 Gör en värdering av förhållandet mellan riskerna och nyttan av projektet.

Forskning får enligt etikprövningslagen bara godkännas om riskerna som den kan medföra för forskningspersonernas hälsa, säkerhet och personliga integritet uppvägs av dess vetenskapliga värde.

7.4 Beskriv hur projektet har utformats för att minimera riskerna för forskningspersonerna.

Forskning får enligt etikprövningslagen bara godkännas om det förväntade resultatet inte kan nås på ett annat sätt som innebär mindre risker för forskningspersoners hälsa, säkerhet och personliga integritet.

Behandling av personuppgifter som avses i 3 § får godkännas bara om den är nödvändig för att forskningen ska kunna utföras.

Observera att om en behandling ska prövas för första gången på människa måste detta tydligt framgå och relevanta säkerhetsrutiner tydligt beskrivas.

Ange vilka risker som ett deltagande i forskningsprojektet kan innebära. Finns det risker t.ex. med att ingå i en kontrollgrupp?

Datasäkerhet extremt viktigt!!!
Vart ska datat förvaras?

Kolla med din inst vad som gäller!

Tex:
Skyddade dokument
Krypterade servrar
Cryptshare för att skicka data

Eventuellt
Risk & sårbarhetsanalys
Säkerhetsklassificera data

ABSOLUT INTE MOLNTJÄNSTER

7.5 Identifiera och precisera om eventuella etiska problem (nackdelar/fördelar) kan uppstå i ett vidare perspektiv genom forskningsprojektet.

Redovisa t.ex. om vissa grupper (andra än de forskningspersoner som ingår i forskningsprojektet) kan komma att utpekats respektive få hjälp som ett resultat av projektet. Frågan avser även de indirekta riskerna eller den indirekta nyttan. T.ex. genetisk påverkan på kommande generationer eller om resultaten på annat sätt kan tänkas skada vissa grupper.

Säker plattform för registerforskning

Utlämning av data: MONA systemet

Vad är MONA

MONA (Microdata Online Access) är SCB:s plattform för tillgängliggörande av mikrodata. I MONA kan användare göra bearbetningar via internet utan att mikrodata lämnar SCB.

Innehåller bla: R, R studio, Stata, SPSS, SAS, ...



Att tänka på med
data från flera
myndigheter

Länkningsprotokoll

- Gör ett förslag/idé för vilken ordning register ska länkas
 - Ta med vilken myndighet som hanterar resp register
 - Bifoga gärna till EPM ansökan!
-
- OBS! handläggarna på resp myndighet bestämmer, men ni kan göra ett förslag

Exempel på registerlänkningsprotokoll

Frågeställning:

- **Har barn som vuxit upp i med föräldrar med låg inkomst en ökad risk för allergi och astma (jämfört med hög inkomst)?**

Design 1 + ett förslag på länkning (fall-kontroll)

- Väljer ut alla individer med astma (ICD10 J45.9) eller allergidiagnos (T78.4) ur patientregistret samt att dessa personer ska vara >5 års ålder vid diagnos (fallgrupp) och att diagnosen sker efter 1 juli 2005
- Skapar 4 kontroller (matchad för kön, ålder och län) ur totalbefolkningsregistret (kontrollgrupp)
- Länkar fall och kontrollgrupp mot flergenerationsregistret för att hitta deras föräldrar
- Länkar föräldrarna mot LISA för att hitta deras inkomst
- Väljer den exponering för fall och kontrollgrupp enl beslut tidigare
- Använder inkomstkategorierna och beräknar oddsen för astma och allergidiagnos

Länkningsprotokoll, förslag design 1

Sändare	Mottagare		Beskrivning
		Soc	Väljer ut alla individer med astma (ICD10 J45.9) eller allergidiagnos (T78.4) ur patientregistret samt att dessa personer ska vara >5 års ålder vid diagnos (fallgrupp) och att diagnosen sker efter 1 juli 2005
Soc	SCB		Skickar dessa fallen till SCB
		SCB	Skapar 4 kontroller (matchad för kön, ålder och län) ur totalbefolkningsregistret (kontrollgrupp)
		SCB	Länkar fall och kontrollgrupp mot flergenerationsregistret för att hitta deras föräldrar Länkar föräldrarna mot LISA för att hitta deras inkomst Länkar mot confounders som kan tänkas finnas i SCB register
SCB	Soc		Skickar tbks data på fall och kontroller till Soc
		Soc	Länkar fall- och kontrolldata med data på confounders från patientregistret
		Soc	Skapar internt ID nummer för alla personnummer i fall- och kontrollfilerna
		Soc	Avkodar fall- och kontrollfilerna från personnummer och arkiverar kodnyckeln
Soc	Forskaren		Sänder registerdata för de önskade registren för fall- och kontrollpersoner till forskaren

Design 2 + ett förslag på länkning (kohort)

- Väljer ut alla nyfödda barn under 2006 från medicinska födelseregistret
- Länkar alla barn mot flergenerationsregistret för att hitta deras föräldrar
- Länkar föräldrarna mot LISA för att hitta deras inkomst
- Väljer den exponering för fall och kontrollgrupp enl beslut tidigare
- Länkar alla barn mot läkemedelsregistret mellan 2006-2011 för att hitta poster där de blivit utskrivna läkemedel för astma (R03CC - bricanyl) eller allergi (R06 – antihistaminer).
- Använder inkomstkategorierna och beräknar relativa risken för utskrivna läkemedel för astma och allergi (ev med ålder som tidskala mha Cox regression)

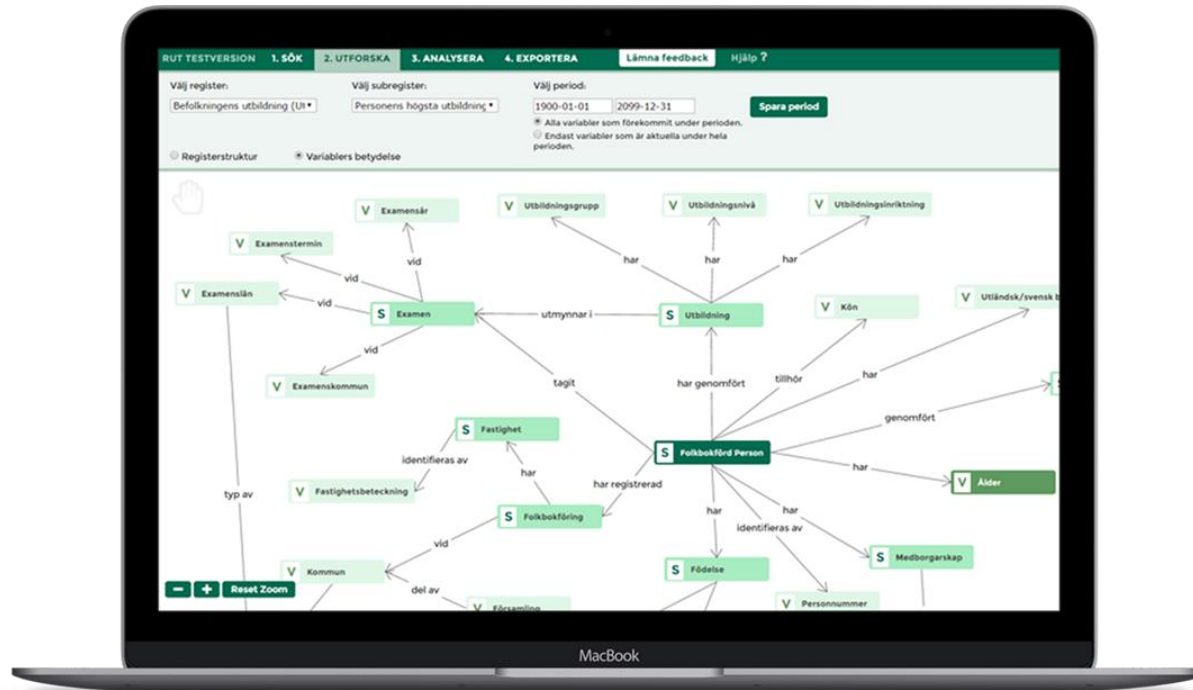
Länkningsprotokoll, förslag design 2

Sändare	Mottagare		Beskrivning
		Soc	Väljer ut alla barn födda 2006 ur det medicinska födelseregistret
		Soc	Länkar barnen mot läkemedelsregister för att se detaljerad info om mediciner, avsöker specifika ATC koder
		Soc	Skapar internt ID nummer för alla barnen
Soc	forskaren		Avkodar filerna från medicinska födelseregistret och läkemedelsregistret från personnummer och skickar till forskaren
Soc	SCB		Skickar personnummer och internt ID nummer för barnen till SCB
		SCB	Länkar barnen mot flergenerationsregistret för att hitta deras föräldrar Länkar föräldrarna mot LISA för att hitta deras inkomst
SCB	Forskaren		Sänder data barnens föräldrar och deras inkomst till forskaren

Register Utiliser Tool (RUT)

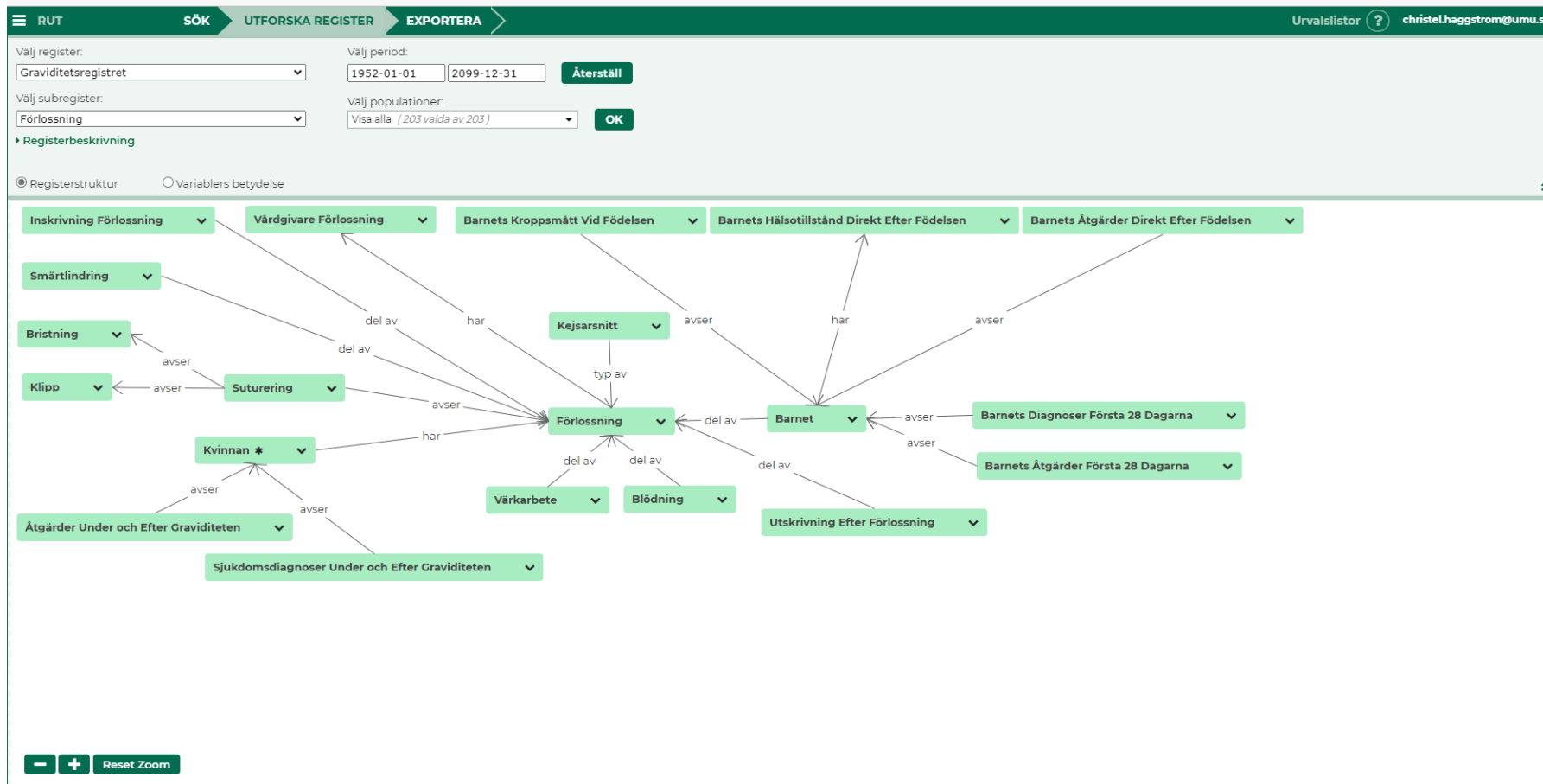
Ett verktyg under utveckling för att utforska register och biobanker på metadata-nivå

Visar registers struktur och innehåll på variabelnivå samt variablers betydelse och representation



RUT: <https://rut.registerforskning.se/>

Graviditetsregistret, förlossning, exempel registerstruktur





Att tänka på när
data kommer
(helst innan..)

När datat kommer

- Oftast (alltid?) en fil från varje register
- Filerna = råa registerutdrag (inget bearbetat data)
- Filerna kan innehålla flera rader per individ etc
- 3 månaders ”reklamations tid” från myndigheter
- Datahantering kräver tid!!

Statistisk mjukvara?

- Olika fördelar/nackdelar
- Bra kunna dela skript mellan kollegor
- Datumhantering
- "Missing" av variabler

If statistics programs/languages were cars...



Datahantering på registerdata

Ex från datahanteringsuppgift

- Fall från kvalitetsregistret Qreg länkat mot kontroller från RTB
- → Kontrollkopplingsfil
- Data från patientregistret

Du vill

- Göra en fil med en rad för varje fall och kontroller
- Veta vilken kontroll som tillhör vilket fall
- Lägga på en variabel för varje fall och kontroller för totalt antal inlagda dagar året efter diagnosår/indexår
- Hur gör du detta?

Datat från Qreg har variablerna

<u>Variabelnr</u>	Variabelnamn
1	<u>LopNr</u>
2	<u>FoddAr</u>
3	Kon
4	Dia1
5	Dia1datum

Datat från kontrollkopplingsfilen har variablerna:

<u>Variabelnr</u>	Variabelnamn
1	<u>LopNr</u>
2	<u>FoddAr</u>
3	Kon
4	<u>LopNr Kontroll</u>
6	Dia1År

Filen från patientregistret innehåller variablerna:

<u>Variabelnr</u>	Variabelnamn	Klartext
1	<u>hdia</u>	Huvuddiagnos
2	<u>LopNr</u>	Löpnummer
3	ALDER	Ålder vid besökstillfället
4	INDATUMA	Inskrivningsdatum slutenvård, besöksdatum öppenvård, alfanumerisk
5	UTDATUMA	Utskrivningsdatum, alfanumerisk
6	VTID	Vårdtid (dagar)



Sammanfattning

Do and don't i registerforskning

- Skapa kontakt med handläggare på register/myndigheter för att se om din frågeställning går att få svar på med datat på registret.
- Var noga med EPM ansökan
 - Nämn alla register (+ gärna alla variabler + intressanta koder och deras selektion redan här!)
 - Skriv hur kodnyckeln ska hanteras + eventuell tid som kodnyckeln ska sparas
 - Risker vs nytta perspektivet, finns det några ytterligare risker i "långa loppet"?
 - Gör ett förslag till länkningsordning
 - Ha en grov analysplan
 - Tänk redan här på hur det färdiga datat ska förvaras/bearbetas
- I ansökan till register
 - Kolla om viktiga variablers definition eller kodning förändrats över tid
 - Fundera på: Har registerna validerats? Går det att lita på data? Kommer datat "fånga" det vill undersöka?
 - Fundera på: Hur "noggranna" variabler behövs map ICD och ATC koder
 - Lyssna på tips från SCB/Soc/kvalitetsregister
 - Läs noga offerterna!!

Do and don't i registerforskning

När data kommer

- Kolla det noga! 3 mån reklamations tid
- Gör en kopia av rådatat som du inte "rör"
- Spara all cleaning i skript och helst i ett "DM dokument"
- Om ngn annan ska använda samma cleaning: skriv "DM dokument" i tex word med stolpar/punkter på vad skripten gör
- Det kommer vara "fel" i datat – bestäm hur det ska hanteras kollektivt
 - Bla kan det finnas "återanvända personnummer"
 - Tex mismatch mellan datum – vilket datum ska du lita på (tex olika definition i kval reg/cancer reg)
- En rad per individ gör det lätt att samköra!
- Mina datahanteringstips:
 - Börja med en datafil med enbart det viktigaste (id, kön, fall/kontroll status etc), länka sen på fler och fler variabler eftersom du hunnit cleana de filerna.
 - Skapa cleanade hjälpfiler för specifika variabler

Om ni vill veta mera..

Vecka 45, HT 2023

Introduktion till registerforskning, 1,5 hp Introduction to register-based research, 1,5 ECTS

Kursansvarig	Christel Häggström Telefon: +46 90 785 72 80 Epost: christel.haggstrom@umu.se	
Kursadministratör	Ulrika Järholm Telefon: +46 90 786 71 43 Epost: ulrika.jarholm@umu.se	
Institution	Institutionen för epidemiologi och global hälsa	
Datum	7 – 11 november	
Språk	Svenska	
Antal deltagare	20	
Undervisningsform	Föreläsningar	18 timmar
	Seminarier	12 timmar
	Praktiska exempel	8 timmar
Examinationsform	Att medverka i och klara av praktiska övningar på seminarier	

Kursens innehåll

Detta är en introduktionskurs i registerforskning. Kursen avser att ge generell teoretisk kunskap om och grundläggande praktiska färdigheter för forskning på kvalitetsregister.

Vårens seminarier

22/3 kl 12-13 Poweranalyser och stickprovdimension

Simon Vallin, Region Västerbotten / institutionen för folkhälsa och klinisk medicin, Umeå universitet

18/4 kl 12-13 En introduktion till ordinal regression

När vår utfallsvariabel mäts på en ordinal skala finns det särskilda regressionsmodeller i vår statistiska verktygslåda för att hantera detta. Här ges en introduktion till dessa modeller. Det kommer visas exempel på hur ordinala regressionsmodeller kan vara mycket användbara även när vår utfallsvariabel mäts på intervall- eller kvotskala men då normalfördelningsantagande inte uppfylls, eftersom de till skillnad från de vanliga icke-parametriska metoderna även kan justeras för kovariat.

Per Liv, Umeå universitet, institutionen för folkhälsa och klinisk medicin, Umeå universitet

9/5 kl 12-13 Att strukturera och organisera forskningsdata

Data kommer i många former, men i regel måste data lagras i rätt bestämda strukturer för att vara bearbetningsbart. Det kan också vara bra att ha en plan för hur datafiler och tillhörande material ska organiseras för att materialet ska vara överblickbart. Det senare kan vara särskilt viktigt i projekt med flera inblandade, men även för att det ska finnas en begriplig struktur när material ska arkiveras. Målet med passet är att peka på en del aspekter av detta med struktur som kan vara viktiga att tänka på och även komma med konkreta förslag.

Anna Lindam, Region Jämtland Härjedalen / Institutionen för folkhälsa och klinisk medicin, Umeå universitet,

Robert Lundqvist, Region Norrbotten / Institutionen för folkhälsa och klinisk medicin, Umeå universitet

1/6 kl 12-13 Presentatör och ämne bestäms senare

<https://www.canvas.umu.se/courses/2600>

FRÅGOR?

Tack!



Patientregistret

- Slutenvårdsregistret sen 1987
- Öppenvårdsregistret sen 2001
- Diagnoser kodade enl ICD-10 sen 1997 (innan ICD-9 sen 1987)
- Bortfallet 1% för huvuddiagnos, personnummer och yttre orsak.
- **Intern kvalitetscheck:** <https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/ovrigt/information-for-patientregistret-registerar-2020.pdf>

En rad i registret är ett vårdtillfälle!

Dvs en individ har (oftast) flera rader

Vid ansökan behövs de intressanta diagnoserna eller åtgärderna specificeras och trunkeras (med ICD eller åtgärds-koder)

Cancerregistret

- Startade 1958 med syfte kartlägga cancersjukdomar
 - Hur utvecklingen ser ut när det gäller nya antal fall.
 - Om det finns några regionala eller demografiska skillnader.
 - Hur stor överlevnadssannolikheten i olika cancersjukdomar är.
- Sos ansvarar för registret men registreringen sker regionalt (RCC)
- Hög kvalité, 99% av diagnoserna fastställda med morfologi
- Dokumentet: Kodning i cancerregistret - hjälp för ansökan

En rad i registret = en tumör (primärtumör)
Dvs en individ kan ha flera rader.

Läkemedelsregistret

- Start juli 2005
- Alla läkemedel som hämtas ut mot recept på apotek
- Många poster! Drygt 100 miljoner nya poster per år
- Information om patienten (kön, ålder, ort), varan (ATC kod), förskrivningen (mängd, datum för förskrivning och uthämtning), kostnader, detaljer om arbetsplatsen vart förskrivningen skedde
- Uppdateras 1 gång/månad

En rad i registret = ett uthämtat recept
Dvs en individ har (oftast) väldigt många rader!
Vid ansökan måste intressanta läkemedel
specificeras och trunkeras mha ATC koder!

Medicinska födelseregistret

- Startade 1973
- Innehåller data för alla graviditeter som lett till förlossning
- Syftet med registret är att kunna beskriva händelser och utfall för kvinnor och barn under graviditet, förlossning och nyföddhetsperioden.
- Sedan 1995: Två journalblad från mödrahälsovården, en från förlossningsvården, och en från det nyfödda barnet.
- Kopplat till medicinska födelseregistret finns [registret övervakning av fosterskador och kromosomavvikelser](#).

En rad i registret = en förlossning

Tandhälsoregistret

- Startades juli 2008
- Innehåller information om tandvård inom det statliga tandvårdsstödet, tandvård till personer med vissa långvariga sjukdomar och funktionsnedsättningar samt nödvändig tandvård.
- Syftet är att kunna följa tandvårdens och tandhälsans utveckling över tid
- Kvalitén uppskattas vara god men ej validerat på vissa variabler/åtgärder
- Rapporteras via Försäkringskassan till Socialstyrelsen

En rad i registret = en diagnos eller ett ärende hos tandvården

Dödsorsaksregistret

- Från 1961 (finns även ett historiskt dödsorsaksregister tidigare)
- Från 2012 alla dödsfall i Sverige oavsett folkbokföringsland
- Innehåller dödsdatum, underliggande och bidragande dödsorsak (tas från dödsorsaksintyget)
- Kvalitet: Den största källan till osäkerhet är läkarens fastställande och rapportering av dödsorsak. Dödsorsaksuppgifterna är säkrare för yngre personer än för äldre. Gamla människor lider ofta av många sjukdomar och det kan vara svårt för läkaren att avgöra vilka tillstånd som var avgörande för dödsfallet.
- Uppgifter om avlidna räknas ej som känsliga uppgifter

En rad i registret = ett dödsfall