

Data handling and register data - pitfalls and challenges, consequences after the GDPR legislation

John Kling
Henrik Nordin

Registerservice

- Databeställningar
(ca 1000 per år)
 - Forskning
 - Statistik
 - Kvalitetsregister
- Stöd till externa dataanvändare
- Stöd till kvalitetsregister
- Ingång för externa dataanvändare

Lag (1998:543) om hälsodataregister

- Regeringen bestämmer vilka hälsodataregister som ska finnas
 - en förordning per register
- Uppgiftsskyldighet för ”den som bedriver verksamhet inom hälso- och sjukvården”
- Inget samtycke krävs, ingen opt-out
- Tillåtna ändamål:
 - framställning av statistik
 - uppföljning, utvärdering och kvalitetssäkring av hälso- och sjukvård
 - forskning och epidemiologiska undersökningar

Statistiksekretess

24 kap 8 § offentlighets- och sekretesslagen (2009:400)

Huvudregel: Absolut sekretess

Kan brytas för:

- Uppgifter som behövs för forskningsändamål
- Uppgifter som behövs för statistikändamål
– främst till andra statistikansvariga myndigheter
- Uppgifter som *inte* är *direkt* hänförliga till enskild
- Dödsorsak och dödsdatum till kvalitetsregister

om det står klart att den enskilde eller närstående inte lider skada eller men

Utlämnande av *individuppgifter* kan inte göras för

- Kvalitetssäkring
- Uppföljning på individnivå
- Utvärderingar
- Kliniskt förbättringsarbete
- Examensarbeten (utbildning)
- Tillsyn

Sekretessprövning

(ej uttömmande lista)

Rätten att behandla uppgifterna

Forskningsändamål

Stöd för behandling av känsliga personuppgifter? (etiskt godkännande)

Forskningshuvudman: samma i begäran som i etikansökan

Menprövning

Kan någon lida skada eller men om uppgifter lämnas ut?

Sekretesskyddet hos mottagaren

Behovsprövning

Endast nödvändiga registeruppgifter

Sekretesskydd hos mottagaren

24 kap 8 § OSL (statistiksekretess)

25 kap 1 § OSL (hälso- och sjukvårdssekretess)

11 kap 3 § OSL (sekretess överförs från utlämnande myndighet)

7 § OSF (vissa däri angivna undersökningar)

Inget sekretesskydd (gäller tex företag) →
beslut om utlämnande förenas med förbehåll gentemot
de som ska behandla uppgifterna vid
forskningshuvudmannen

Vad vi tittar på i en etikansökan

- Att beskrivningen av **populationen** överensstämmer mellan begäran till Socialstyrelsen.
- Att **register** finns namngivna i etikansökan och dess syfte.
- Att **syftet** (behovet) med begärda registeruppgifter framgår i etikansökan.
- Att det framgår att projektet avser att följa upp individer över tid om en **nyckel** behöver sparas.
- Att det framgår varför **personnummer** krävs om personnummerutlämnande beställts (vanligen lämnas registeruppgifter ut med löpnummer).
- Att det framgår att uppgifter från **andra registerhållare** ska hanteras inom projektet om Socialstyrelsen ska samköra med data från andra registerhållare, till exempel kvalitetsregister.
- Att **forskningshuvudmannen** stämmer överens med den som angetts i beställningen. Leverans görs till den av Etikprövningsmyndigheten godkända forskningshuvudmannen.

**Mer information finns på:
www.socialstyrelsen.se**

Mikrodatauppdrag vid SCB

Freja Werke

Jonas Färnstrand

2022-09-09



”Mikrodata är data om enskilda personer, företag eller arbetsställen. Mikrodata kan också vara en rad i en tabell som summeras till en person, ett företag eller ett arbetsställe.”

Att beställa data från flera registerhållare

- Separata beställningar görs till respektive registerhållare.
- Samtliga registerhållare gör en sekretessprövning.
- Registerhållarna diskuterar:
 - ✓ Population och Nycklar
 - ✓ Registeruppgifter
 - ✓ Avslagsbeslut
 - ✓ Teknik och juridik

Beställningsformulär

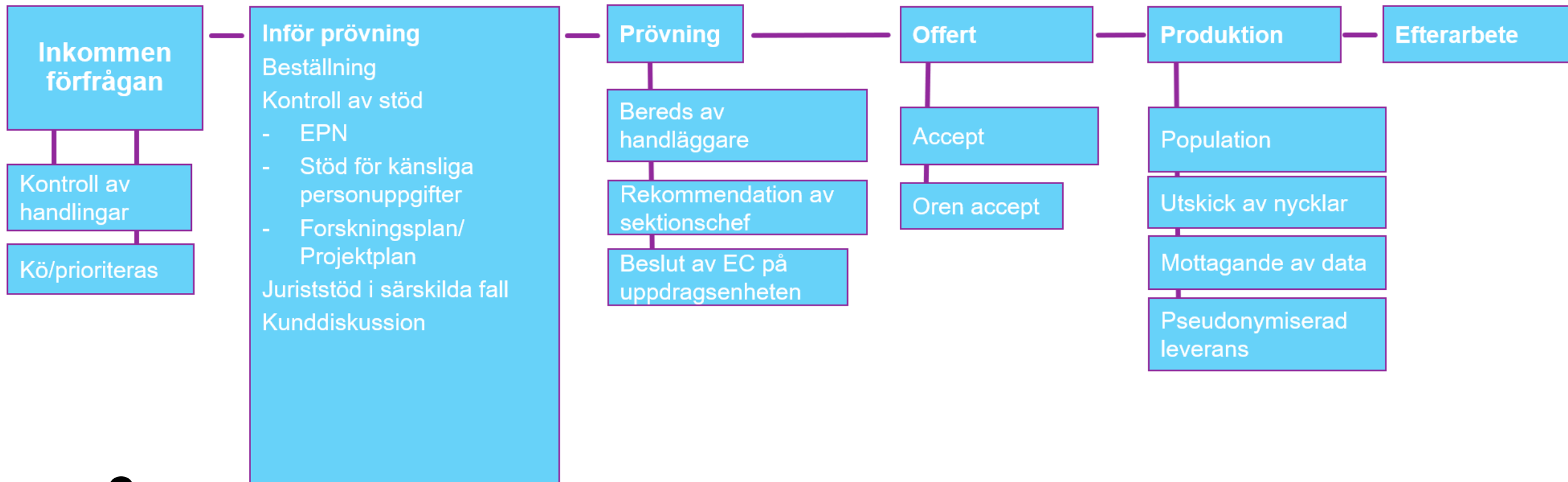
- <http://www.scb.se/vara-tjanster/bestalla-mikrodata/sa-har-bestaller-du/>

Att fylla i:

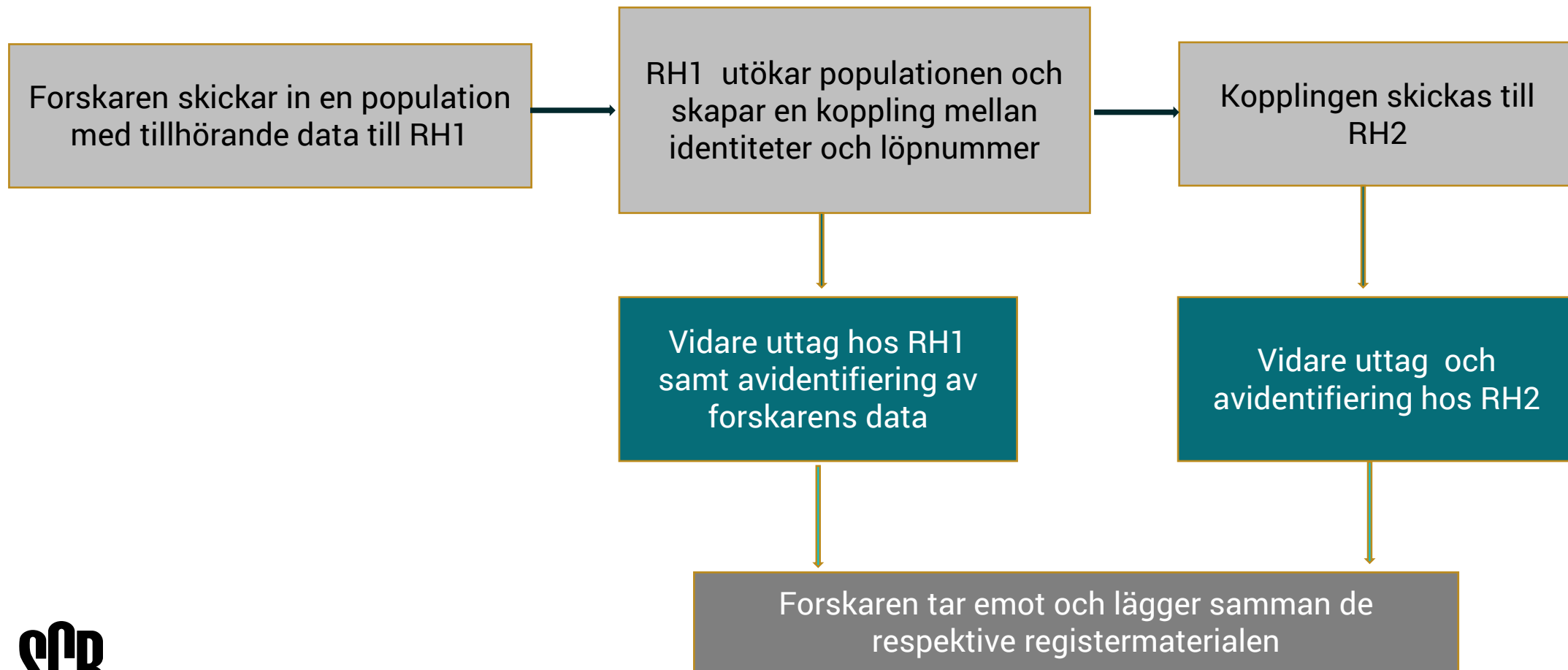
- Forskningshuvudman/Datamottagare
- Kortfattad beskrivning av uppdraget samt med vilket stöd uppgifterna ska behandlas.
- Kontaktperson, leveransadress, faktureringsuppgifter
- Övriga inblandade myndigheter
- Att bifoga: variabellista, forskningsplan, etikansökan, samt godkännande och eventuella kompletteringar



Arbetets gång



Forskningsprojekt med registerdata från flera registerhållare



Tid i de olika delmomenten under 2021



Registret över totalbefolkningen (RTB)

- Alla folkbokförda personer vid varje årsskifte med början 1968-12-31
- Ett utdrag ur Skatteverkets folkbokföringsregister:
 - Personnummer (kön, ålder)
 - Folkbokföring (namn, adress)
 - Familjer (individer folkbokförda i samma lägenhet)
 - Civilstånd (kopplingar till make/maka/partner/sambo)
 - Migrationer (datum, invandrings-/utvandringsland)
- Årsversionen tillgänglig i mitten av februari året efter

RTB i forskning

- Skapa populationer
- Folkbokföringsuppgifter – långa tidsserier
- Kopplingar till make/maka
- Migrationer
- Dödsdatum
- Lägenhetsregistret (från 2011)

Flergenerationsregistret

Individer folkbokförda någon gång sedan 1961, födda 1932 och senare (inkl. avlidna och emigrerade)

Innehåll:

- Eget personnummer, biologisk mors och fars personnummer samt adoptivmors och adoptivfars personnummer

Kontroller:

- Borttag av mammor och pappor med t.ex. orimlig ålder vid barnets födelse

Historiska Befolkningsregistret

- Baseras på Registret över totalbefolkningen
- Uppbyggt kring händelser istället för årsskiften
- Alla folkbokförda i Sverige sedan 1968-12-31
- Födelse, immigration, emigration, inrikes flytt, civilståndsändring, medborgarskapsändring, död
- 16,5 miljoner individer, 129 miljoner händelser

Historiska befolkningsregistret

Individ	Datfran	Dattill	Fodd	Immigration	Civilandring	Medborgar- byte	Inrikes flytt	Emigration	Dod
1	19721008	19740122	0	1	0	0	0	0	0
1	19740122	19750325	0	0	0	0	1	0	0
1	19750325	19760618	0	0	0	0	1	0	0
1	19760618	19760619	0	0	0	1	0	0	0
1	19760619	19801120	0	0	1	0	0	0	0
1	19801120	20160101	0	0	0	0	1	0	0
1	20160101	20210101	0	0	0	0	0	0	0
2	20180301	20180804	0	1	0	0	0	1	0

Händelser

- Flest händelser för en individ: 127
- Flest inrikes flyttar för en individ: 112
- Flest antal boendekommuner för en individ: 47
- Flest immigrationer för en individ: 21
- Flest emigrationer för en individ: 20
- Flest civilståndsändringar för en individ: 22
- Flest antal olika civilstånd för en individ: 6 (av 7 möjliga)
- Flest födslar för en individ: 1
- Flest dödsfall för en individ: 1

Historiska Befolkningsregistret i forskning

- Fall- och kontrollpopulationer
- Följa personer under en tidsperiod
- Händelser ett visst datum

Fall- och kontrolldragning

- Standardiserat script
- Kontrollerna dras på exakt datum
- Möjlighet att dra kontroller på:
 - Födelseår/månad
 - Kön
 - Boenderegion (riket, län, kommun, församling/distrikt)
- Med återläggning/utan återläggning
- Fall kan dras som kontroller innan sitt eget indexdatum/fall kan aldrig dras som kontroller

Fall- och kontrolldragning

Lopnr_fall	Lopnr_kontroll	Fodlsear	Kon	Region	Datum	Ordning
1	234	1978	1	0180	20050618	1
1	12788	1978	1	0180	20050618	2
1	43999	1978	1	0180	20050618	3
1	6	1978	1	0180	20050618	4
1	11884	1978	1	0180	20050618	5
2	799	1911	2	2403	20120830	1
2	67434	1911	2	2403	20120830	2
2	42	1911	2	2403	20120830	3

Fall- och kontrolldragning

- Följande information skickas:
 - Kopplingsfil mellan fallen och deras kontroller
 - Fil med individer som inte var folkbokförda vid sitt indexdatum
 - Fil med individer som var folkbokförda vid sitt indexdatum men där det inte gick att hitta några kontroller alls

I löpnummernyckeln som skapas brukar vi lägga till information om vilka som är fallindivider och vilka som är kontrollindivider.

Fall- och kontrolldragning

- Möjliga problem
 - Indexdatum före 19681231 (t.ex. Cancerregistret)
 - Indexdatum efter 20220101
 - Saknade/felaktiga indexdatum i fallpopulationen
 - Kontrollerna tar slut

Hur dras gränsen för en personuppgift?

Kort presentation + Diskussion

Personuppgift

Definition, GDPR art. 4(1)

Varje upplysning som avser en identifierad eller identifierbar fysisk person,

varvid en identifierbar fysisk person är en person som direkt eller indirekt kan identifieras

särskilt med hänvisning till en identifierare som ett namn, ett identifikationsnummer, en lokaliseringsuppgift eller onlineidentifikatorer eller en eller flera faktorer som är specifika för den fysiska personens fysiska, fysiologiska, genetiska, psykiska, ekonomiska, kulturella eller sociala identitet.

Personuppgift

GDPR, Beaktandesats 26 (utdrag)

”Personuppgifter som har pseudonymiserats och som skulle kunna tillskrivas en fysisk person genom att kompletterande uppgifter används bör anses som uppgifter om en identifierbar fysisk person.”

Pseudonymisering är (oftast) inte tillräckligt

Finns det en nyckel till identifierande uppgifter så är det personuppgifter

Personuppgift

GDPR, Beaktandesats 26 (utdrag)

”För att avgöra om en fysisk person är identifierbar bör man beakta alla hjälpmedel som, antingen av den personuppgiftsansvarige eller av en annan person, rimligen kan komma att användas för att direkt eller indirekt identifiera den fysiska personen.”

Personuppgift

GDPR, Beaktandesats 26 (utdrag)

”För att fastställa om hjälpmedel med rimlig sannolikhet kan komma att användas för att identifiera den fysiska personen bör man beakta samtliga objektiva faktorer, såsom kostnader och tidsåtgång för identifiering, med beaktande av såväl tillgänglig teknik vid tidpunkten för behandlingen som den tekniska utvecklingen.”

Sensitive personal data or anonymized data ?

	P nr	Number	ICD code	Date of diagnosis	Date of birth	Sex	Age st diagnosis
Example 1	720628-9999	No data	S720	2014-06-28	1972-06-28	Male	41.99863107
Example 2 pseudonymization	No data	14	S720	2014-06-28	1972-06-28	Male	41.99863
Example 3	No data	No data	S720	2014-06-28	1972-06-28	Male	41.99863
Example 4	No data	No data	S720	June 2014	June 1972	Male	42
Example 5	No data	No data	S720	2014	No data	Male	42
Example 6	No data	No data	S720	2014	No data	Male	40-45

Sensitive personal data or anonymized data ?

	P nr	ICD code	Date of diagnosis	Date of birth	Sex	Municip.	P nr
Example 4a	No data	S720	June 2014	June 1972	Male	Munkfors	720615-XXYX

The diagram illustrates the classification of data fields in the table above. Brackets and labels are used to group the data:

- Own data** (blue bracket): Encompasses the 'P nr', 'ICD code', and 'Date of diagnosis' columns.
- Quasi-identifiers** (purple bracket): Encompasses the 'Date of birth', 'Sex', and 'Municip.' columns.
- Navet** (orange bracket): Encompasses the 'P nr' column.

k-anonymitet genom omkodning

Table 1. A raw medical data set

	Quasi identifier			Other attributes			Sensitive attribute
	Gender	Age	Postcode				Diagnosis
1	male	25	4350	depression
2	male	27	4351	depression
3	male	22	4352	flu
4	male	28	4353	flu
5	female	34	4352	depression
6	female	31	4352	flu
7	female	38	4350	alcohol addiction
8	female	35	4350	alcohol addiction
9	male	42	4351	alcohol addiction
10	male	42	4350	alcohol addiction
11	male	45	4351	alcohol addiction
12	male	45	4350	alcohol addiction

Table 2. 4-anonymous view of Table 1

	Quasi identifier			Other attributes			Sensitive attribute
	Gender	Age	Postcode				Diagnosis
1	male	20-29	435*	depression
2	male	20-29	435*	depression
3	male	20-29	435*	flu
4	male	20-29	435*	flu
5	female	30-39	435*	depression
6	female	30-39	435*	flu
7	female	30-39	435*	alcohol addiction
8	female	30-39	435*	alcohol addiction
9	male	40-49	435*	alcohol addiction
10	male	40-49	435*	alcohol addiction
11	male	40-49	435*	alcohol addiction
12	male	40-49	435*	alcohol addiction

Utlämnande av data?

Hänsyn måste även tas till offentlighets- och sekretesslagen (2009:400)

Från hälso- och sjukvårdsverksamhet: 25 kap. 1 §

Från forskningsverksamhet: 24 kap. 8 §

k-anonymitet genom omkodning

Table 1. A raw medical data set

	Quasi identifier			Other attributes			Sensitive attribute
	Gender	Age	Postcode				Diagnosis
1	male	25	4350	depression
2	male	27	4351	depression
3	male	22	4352	flu
4	male	28	4353	flu
5	female	34	4352	depression
6	female	31	4352	flu
7	female	38	4350	alcohol addiction
8	female	35	4350	alcohol addiction
9	male	42	4351	alcohol addiction
10	male	42	4350	alcohol addiction
11	male	45	4351	alcohol addiction
12	male	45	4350	alcohol addiction

k-anonymitet genom omkodning

Table 2. 4-anonymous view of Table 1

	Quasi identifier			Other attributes			Sensitive attribute
	Gender	Age	Postcode				Diagnosis
1	male	20-29	435*	depression
2	male	20-29	435*	depression
3	male	20-29	435*	flu
4	male	20-29	435*	flu
5	female	30-39	435*	depression
6	female	30-39	435*	flu
7	female	30-39	435*	alcohol addiction
8	female	30-39	435*	alcohol addiction
9	male	40-49	435*	alcohol addiction
10	male	40-49	435*	alcohol addiction
11	male	40-49	435*	alcohol addiction
12	male	40-49	435*	alcohol addiction

**Mer information finns på:
www.socialstyrelsen.se**