

# Registerforskning i Danmark

Preben Bo Mortensen

Center for Registerforskning, Aarhus Universitet, Taasingegade 1, DK-8000 Aarhus C, Danmark

Telefon: 8942 6820    Telefax: 8942 6813    E-mail: pbm@ncrr.dk

## RESUMÉ

Registre udgør i dag én af de væsentligste forskningsressurser i Danmark. Der har i Danmark været gjort et stort arbejde for at afhjælpe nogle af de væsentligste barrierer imod udnyttelsen af registre til forskningsbrug. Det drejer sig om lovgivning, uensartede regler om data-adgang, uensartede priser, og andre forhold. De danske erfaringer viser i dag, at registre kan bruges i en lang række sammenhænge, samtidig med at borgernes privatliv beskyttes, og i fremtiden er der potentialer til at udbygge udnyttelsen, f.eks. i forbindelse med familiestudier, kobling mellem sociale og helbredsmæssige data, risiko-overvågning, geografiske studier, farmakoepidemiologiske studier og kobling til biobanker. Da de fleste barrierer mod udnyttelsen beror på lovgivning og politiske og administrative afgørelser, vil det i det store hele i lande som de nordiske udelukkende afhænge af politisk vilje, om potentialet skal udnyttes.

Mortensen PB. **Register-based research in Denmark.** *Nor J Epidemiol* 2004; 14 (1): 121-124.

## ENGLISH SUMMARY

At present, registers are some of the most significant resources in research in Denmark. In Denmark, great efforts have been made to remedy some major barriers in utilizing registers for research purposes. They were: legislation, different regulations for accessing data, differing prices, and others. Danish experience shows that registers can be utilized for a wide range of purposes, while protecting the citizens' privacy, and in the future there is great potential to expand the utilization, e.g., in connection with family studies, linkage between social and health related data, risk surveys, geographic studies, pharmaco-epidemiological studies and linkage to biobanks. Since most barriers are due to legislation, political and administrative matters, in the Nordic countries by and large this will depend solely on political will to take advantage of this potential.

Registre udgør i Danmark én af de vigtigste forskningsressurser overhovedet. Dette er selvfølgelig et postulat, som i sig selv kan være svært at bevise. Der er ingen officielle separate statistikker om registerbaseret forskning i Danmark, men et hastigt blik på Medline viser, at en endog meget stor andel af de danske artikler, som publiceres i internationale toptidskrifter er baseret på registre. Der er heller ingen statistikker over, hvor megen international forskningsfinansiering, der vedrører registerbaseret forskning, men der er næppe nogen tvivl om, at de danske befolkningsregistre er et væsentligt aktiv, når f.eks. amerikanske forskningsmidler og samarbejdspartnere skal tiltrækkes. I denne artikel vil jeg kort gøre rede for nogle af de problemer, der er og har været i dansk registerbaseret forskning, hvordan man har forsøgt at løse dem, og hvilke forskningsmuligheder vi nu står over for.

## LIDT HISTORIE

Officielle befolknings- og helbredsstatistikker har i Danmark, ligesom i andre lande, en flere hundrede år gammel historie, og individbaserede helbredsregistre har også en lang historie i form af f.eks. Det Psykiat-

riske Centralregister (1), Cancerregisteret (2), Dødsårsagsregisteret (3), osv. For mange praktiske formål er det imidlertid sådan, at år nul i Danmark er lig 1968, nærmere bestemt 1. april, da CPR nummeret blev indført. Det faktum, at alle borgere har en "unique identifier", som kan checkes logisk for fejl via en algoritme, således at fejllindtastninger undgås, er nøglen til de mange registerkoblinger, som muliggør kombination af mange forskellige typer data om den enkelte person, samtidig med at den enkelte person med stor sikkerhed kan følges over lang tid. Der har været drevet registerbaseret forskning i mange år, både vedr. psykiatriske sygdomme, cancer og meget andet. Registre som forskningsressurser har imidlertid haft en forskellig status. Oprindeligt var en del registerbaseret forskning mest vitalstatistiske opgørelser af ændringer af incidensrater over tid og lignende, mens mere analytisk orienterede epidemiologiske undersøgelser var mere sjældne. Lovgivningsmæssigt og politisk var klimaet også et andet. I mange år var der generelt en del skepsis, både hvad angik kvaliteten af den forskning, der kunne komme ud af brugen af registre, og bekymringerne vedr. sikkerhed kunne være store. I slutningen af 1980-erne var det ikke usædvanligt, at politikere ytrede sig meget bekymret eller ligefrem mente, at selve

brugen af registre til forskning var uetisk, og selve forskningsadgangen til nogle registre var under stærkt pres. Dette skiftede imidlertid radikalt i starten af 90-erne. I forbindelse med et forslag til EU-direktiv, som havde brugt den tyske lovgivning som skabelon, skete der et betydeligt politisk holdningsskift, formentlig hjulpet af en betydelig oplysningsvirksomhed over for politikere, drevet af danske epidemiologiske forskere. Lovgivningen ville indebære et krav om informeret samtykke fra hvert enkelt registrerede individ ved brugen af registre til forskning – noget, der selvsagt effektivt ville forhindre et hvilket som helst brug af registrene. I løbet af den dialog, som fulgte, gik det op for mange, at Danmark i kraft af registrene havde en unik mulighed for relativt billigt at gøre sig internationalt gældende inden for en række forskningsområder. Der var dog fortsat en lang række problemer i form af uensartet adgang til forskellige registre, høje priser, og en række andre problemer, som gradvis blev forsøgt løst op igennem 90-erne. En væsentlig aktør i denne proces var Danmarks Grundforskningsfond, som tog en række initiativer. Blandt andet finansierede man et Center for Registerforskning, først placeret ved Danmarks Statistik i København og siden år 2000 ved Aarhus Universitet, som har til opgave at rådgive og hjælpe forskere, der ønsker at bruge registre i deres forskning. Derudover finansierede man oprettelsen af en filial af Danmarks Statistik i Århus, der gjorde det muligt også for forskere i provinsen at få adgang til de individrelaterede oplysninger om sociale forhold og mange andre forhold, som Danmarks Statistik opbevarer. Herudover finansierede man en forskermaskine mhp. at gøre forskernes adgang til arbejdet med data billigere. Denne proces fortsatte op igennem 90-erne, og i år 2000 blev afholdt en konference med titlen "Registerdata – en enestående dansk forskningsresurse". Ved denne konference blev en række af de vigtigste problemer for dansk registerforskning diskuteret, og det blev efterfølgende besluttet at nedsætte et udvalg under det danske Forskningsministerium mhp. at fremkomme med anbefalinger og løsningsforslag. Dette udvalg med Mads Melbye som formand færdiggjorde sit arbejde i 2001, og rapporten kan hentes via Internettet: [http://www.videnskabsministeriet.dk/cgi-bin/doc-show.cgi?doc\\_id=82791&doc\\_type=37&keywords=Registerforskning](http://www.videnskabsministeriet.dk/cgi-bin/doc-show.cgi?doc_id=82791&doc_type=37&keywords=Registerforskning). Nogle af de væsentligste problemer, som udvalget skulle forsøge at anvise løsninger på, var: uensartede regler fra register til register, således at nogle registre slet ikke gav forskere adgang; andre, som f.eks. Danmarks Statistik, krævede, at man fysisk var placeret hos dem, mens analyserne blev udført; mens andre igen udleverede personhenførbare data til forskernes egne institutioner. Priserne var meget uensartede gående fra få hundrede til flere hundrede tusind kroner for datasæt, og nogle registeransvarlige myndigheder fastsatte deres egne regler med begrænsninger i f.eks. datasæts størrelse, antal af variabler, og andre regler, der fra forskerside kunne synes noget arbitrære. Samtidig med at udvalget

skulle forsøge at anvise løsninger på disse problemer, var det et væsentligt tema, at befolkningens tillid til forskeres brug af registre skulle bevares, således at sikkerheden om muligt blev styrket.

## SITUATIONEN NU

### *Data-adgang*

I dag er der stadig forskel på de enkelte registres model for data-adgang, men udvalgsarbejdet og andre forandringer, der samtidig er sket, har medført store fremskridt for forskerne. For det første er væsentlige registre som Lægemedelstatistikregisteret, dvs. indvidbaserede oplysninger om receptpligtig medicin, blevet gjort tilgængelig via Danmarks Statistiks forskerordning. Samtidig er Danmarks Statistiks forskerordning blevet liberaliseret således, at man nu fra forskningsinstitutioner, der godkendes og indgår kontrakt herom, kan opnå særlig sikker internet-adgang til indvidbaserede koblede, men afidentificerede registerdata – et meget stort fremskridt for de danske forskere. Ved et tilskud fra det danske Videnskabsministerium er priserne blevet betydeligt lavere, således at selve arbejdet med data nu er gjort gratis, og opbygningen af datasæt er blevet gjort betydeligt billigere. Samtidig er der blevet indført en genbrugsordning, således at andre allerede opbyggede datasæt skal kunne genbruges af andre forskere uden yderligere omkostninger. Det har i hele arbejdet været forsøgt at gøre det til et gennemgående princip, at omkostningerne, som forskerne skulle bære, udelukkende skulle være de marginale omkostninger forbundet med deres data-adgang. Generelt er det jo således i Danmark som i de fleste andre lande, at registrene ikke er opbygget for forskernes skyld, men for at løse en række administrative formål. Derfor har det været det gennemgående argument, at forskernes anvendelse af data er en ekstra gevinst, som samfundet kan høste, og at man derfor ikke skal stille økonomiske barrierer i vejen for dette.

### *Regler for data-adgang*

Det blev i udvalget diskuteret, om der var behov for yderligere organer, der skulle sikre ensartede adgangsbetingelser til de offentligt finansierede registre. Det blev konkluderet, at det regelsæt, Datatilsynet arbejder efter, er tilstrækkeligt. Den danske Persondatalov (se [www.datatilsynet.dk](http://www.datatilsynet.dk)) blev vedtaget i år 2000 og repræsenterer i sig selv et fremskridt for registerforskning. Den udgør implementeringen af et EU-direktiv, og står bl.a. derfor ikke umiddelbart over for at skulle ændres. Ud over at give udstrakt beskyttelse af den registrerede borger har den en væsentlig bestemmelse, nemlig bestemmelsen om, at forskere kan gives adgang efter nærmere givne retningslinjer mhp. undersøgelser af væsentlig samfundsmæssig betydning. Der er selvsagt en lang række restriktioner, både hvad angår datas opbevaring, forbud mod at kontakte registrerede enkeltpersoner, osv., og generelt er det selvsagt også

meget lidt berøring, forskeren har brug for at have med selve identiteten af de registrerede, når blot de nødvendige koblinger er gennemført. Indtil videre må man sige, at de danske regler har været en succes, idet der siden CPR nummerets indførelse i 1968 ikke har været noget tilfælde af misbrug af registerdata, der kan føres tilbage til forskere. Den teoretiske ekstra risiko, der er forbundet med at give forskere adgang, givet at man vil have registrering overhovedet, synes derfor at være forsvindende. Dette har givetvis også at gøre med en høj grad af selvjustits blandt forskere, idet de generelt vil være klar over, at selv enkelttilfælde af misbrug vil kunne ødelægge hele feltet for altid.

### *Resterende problemer*

Der er fortsat selvsagt problemer i forhold til, at der kan mangle relevant dokumentation, forskellige datasamlinger kan trænge til validering i forhold til konkrete projekter, osv., men generelt må det siges, at den nuværende situation i Danmark er bedre end nogensinde.

## **FREMTID**

Det er svært at se, hvor den nuværende udvikling vil stoppe, men forskere står i disse år over for en lang række nye interessante muligheder:

### *Familiestudier*

Det er nu muligt via det centrale befolkningsregister, CPR, at koble personer med deres førstegradsslægtninge, altså deres forældre, søskende og børn. Man kan også identificere grupper af personer via deres adresse, således at familie-enheder kan dannes. Dette giver mulighed for en lang række interessante undersøgelser, hvor familiære og miljømæssige risikofaktorer for sygdomme kan studeres – noget, der allerede er flere eksempler på (4). Koblingen af forældre til børn i befolkningsregistre er også nyttigt til klarlægning af f.eks. fertilitetsmønster, som kan bruges til studier af f.eks. kræftforekomst (5).

### *Sociale og helbredsmæssige data*

Muligheden for at koble sociale og helbredsmæssige data gør det muligt at studere samspillet af disse. Dette har f.eks. været udnyttet i selvmordsforskning (6,7). Det kan iøvrigt være vanskeligt at overskue det store antal registre der er til rådighed, men på Center for Registerforskningens hjemmeside findes en oversigt med links til en række væsentlige registre ([www.ncrr.dk](http://www.ncrr.dk)). Herudover kan henvises til Danmarks Statistik ([www.dst.dk](http://www.dst.dk)) og Sundhedsstyrelsen ([www.sst.dk](http://www.sst.dk)).

### *Risiko-overvågning*

Mediebilledet i dag er ofte præget af mistanke om sammenhænge mellem forskellige ekspositioner og sygdomsrisiko. Meget ofte er der stor opmærksomhed

og nogle gange angst forbundet dermed på basis af meget ringe empirisk grundlag. I sådanne tilfælde kan registre nogle gange være en af de eneste måder, hvorpå man hurtigt kan få afkræftet ubegrundede mistanker, som i sig selv kunne have negative helbredsmæssige konsekvenser. Et eksempel kunne være mistanken om sammenhængen mellem mæslingevaccine og autisme, hvor et stort studie kunne gennemføres – et studie, som formentligt havde været umuligt, hvis der ikke havde været adgang til individbaserede registre (8). Tilsvarende eksempler kunne være sammenhæng mellem sygdom og mobiltelefoni, og meget andet.

### *Geografiske studier*

I Danmark kan individers bopæl følges historisk, og dette kan også kobles til de geografiske koordinater, således at man f.eks. kan studere eksposition i form af opvækst i nærheden af højspændingsledninger, o. lign. Ud over at være relevant for den generelle sundheds-overvågning er det også, som ovenfor nævnt, nyttigt at kunne studere noget empirisk, både hvis der er en sammenhæng, men i mindst lige så høj grad hvis mistanken er ubegrundet.

### *Farmakoepidemiologi*

I de fleste lande er der betydelig kontrol med de lægemidler, der introduceres på markedet. Imidlertid er der kun meget ringe mulighed for at følge op med langtidsovervågning af mulige bivirkninger, når medikamenterne først begynder at blive brugt i den kliniske virkelighed uden for de kontrollerede undersøgelser. Muligheden for at koble receptdata til andre helbredsregistre giver unikke muligheder for dette – noget, der allerede har været udnyttet i mange undersøgelser (9).

### *Kobling til biobanker*

Dette er muligvis et af de største potentialer for brugen af registre i biomedicinsk forskning i fremtiden. Kombination af store biobanker og registre, der giver oplysninger om både sygdomme og slægtskabsforhold, giver helt unikke muligheder for undersøgelser af interaktion mellem arvelige og miljømæssige risikofaktorer for sygdom. Dette har bl.a. været baggrunden for en storstilet satsning ved opbygningen af en dansk national fødselskohorte (10), hvor et væsentligt aktiv er, at den store investering, der lægges i at indsamle blodprøver og interview-data under graviditet og ved fødsel, kan forudsiges at blive fulgt op med nærmest komplet follow-up via befolkningsregistre. Herudover er der talrige mere eller mindre systematiske biobanker, f.eks. ved patologiske institutter og andre steder, som kan være nyttige. En særlig datakilde, som kan være relevant i fremtiden, er samlinger af hælblodprøver indsamlet mhp. Føllings sygdom, hvor udviklingen af moderne laboratorieteknikker har muliggjort udnyttelsen af denne resurse (11).

## KONKLUSION

Som det fremgår af ovenstående, er forfatteren til denne artikel meget optimistisk og entusiastisk mht. de muligheder, der ligger i registerbaseret forskning. Det skal man jo være, når man er leder af et Center for Registerforskning. Men herudover mener jeg også, at det er realistisk begrundet. Jeg mener, at den model, man nu har fundet i Danmark, gør det muligt at udnytte noget af det kolossale potentiale, der ligger i registre, samtidig med at befolkningen får den beskyttelse, den har krav på, og ikke mindst at befolkningen på sigt får glæde af de forskningsresultater, som de igennem årene med eller uden deres vidende har leve-

ret data til. Som det også fremgår, er de fleste barrierer imod udnyttelsen udelukkende af lovgivningsmæssig, administrativ eller økonomisk karakter, og det er meget få af disse barrierer, der ikke vil kunne afskaffes, hvis man kan mobilisere den nødvendige politiske vilje, og en økonomisk investering vil være meget beskedent både i forhold til, hvad andre større forskningssatsninger vil koste og i forhold til de beløb, der i forvejen investeres i at vedligeholde og drive de administrative registre, det hele handler om. Der er derfor efter min opfattelse meget få gode argumenter for at unddrage de nordiske befolkninger de kolossale gevinster, registerbaseret forskning kan byde på.

## LITTERATUR

1. Munk-Jørgensen P, Mortensen PB. The Danish Psychiatric Central Register. *Dan Med Bull* 1997; **44** (1): 82-84.
2. Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry – history, content, quality and use. *Dan Med Bull* 1997; **44** (5): 535-539.
3. Juel K, Helweg-Larsen K. The Danish registers of causes of death. *Dan Med Bull* 1999; **46** (4): 354-357.
4. Mortensen PB, Pedersen CB, Westergaard T, Wohlfahrt J, Ewald H, Mors O, et al. Effects of family history and place and season of birth on the risk of schizophrenia. *N Engl J Med* 1999; **340** (8): 603-608.
5. Wohlfahrt J, Andersen PK, Mouridsen HT, Melbye M. Risk of late-stage breast cancer after a childbirth. *Am J Epidemiol* 2001; **153** (11): 1079-1084.
6. Qin P, Agerbo E, Mortensen PB. Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: a national register-based study of all suicides in Denmark, 1981-1997. *Am J Psychiatry* 2003; **160** (4): 765-772.
7. Agerbo E. Risk of suicide and spouse's psychiatric illness or suicide: nested case-control study. *BMJ* 2003; **327** (7422): 1025-1026.
8. Madsen KM, Hviid A, Vestergaard M, Schendel D, Wohlfahrt J, Thorsen P, et al. A population-based study of measles, mumps, and rubella vaccination and autism. *N Engl J Med* 2002; **347** (19): 1477-1482.
9. Vestergaard P, Olsen ML, Paaske JS, Rejnmark L, Sorensen HT, Mosekilde L. Corticosteroid use and risk of hip fracture: a population-based case-control study in Denmark. *J Intern Med* 2003; **254** (5): 486-493.
10. Olsen J, Melbye M, Olsen SF, Sorensen TI, Aaby P, Andersen AM, et al. The Danish National Birth Cohort – its background, structure and aim. *Scand J Public Health* 2001; **29** (4): 300-307.
11. Norgaard-Pedersen B. [The PKU registry and biobank at the National Serum Institute. Rules and applications]. *Ugeskr Laeger* 1998; **160** (15): 2266-2267.