

Bevaring og genbrug af forskningsdata fra sundhedsvidenskab

Professor Merete Osler, arkivar Lone Bredahl & arkivchef Steen Ousager

Syddansk Universitet,
Institut for Sundhedstjenesteforskning, Epidemiologi, og
Statens Arkiver, Dansk Data Arkiv

I de senere år har der været stigende fokus på, at det skal være muligt at få adgang til de data, som indgår i publicerede sundhedsvidenskabelige artikler. Således har flere lægevidenskabelige tidsskrifter, herunder også Ugeskriftet, i dag som betingelse for publikation, at de data, som er grundlaget for videnskabelige artikler, er til rådighed for andre uafhængige forskere, så de har mulighed for at kontrollere resultaterne [1].

Det betyder, at en forudsætning for publikation i anerkendte tidsskrifter er, at forskere skal kunne stille veldokumenterede data til rådighed for reanalyse. Kravet har i første omgang været mest aktuelt for klinisk kontrollerede undersøgelser og epidemiologisk kontraktforskning, men på længere sigt kan man forestille sig, at det kommer til at gælde mere generelt. Endvidere stiller flere fonde f.eks. National Institute of Health (NIH) og Wellcome Trust nu som betingelse for alle større bevillinger, at de data, som der bevilges midler til at indsamle og analysere, senere skal være tilgængelige for andre forskere. Samfundsvidenskab i Danmark og flere andre europæiske lande har en længere tradition for en sådan arkivering af forskningsdata og datadeling. På det sundhedsvidenskabelige område er Danmark dog på nogle områder på forkant med udviklingen, således vil en arkivering af data i Dansk Data Arkiv (DDA) opfylde kravene fra NIH.

Faktaboks

En forudsætning for publikation i anerkendte lægevidenskabelige tidsskrifter og tildeling af større forskningsbevillinger er, at forskere kan stille deres data til rådighed for reanalyse

Arkivering af forskningsdata i Dansk Data Arkiv/DDA Sundhed opfylder ovennævnte krav

DDA Sundhed medfinansierer også dokumentation og oparbejdning af undersøgelser direkte i de sundhedsvidenskabelige forskningsmiljøer

Forskningsrådenes initiativer for dataarkivering

Med baggrund i ovennævnte udvikling og i erkendelse af, at der anvendes store resurser på nye dataindsamlinger, som kun udnyttes i begrænset omfang, har Forskningsrådet for Sundhed og Sygdom (tidligere Det Sundhedsvidenskabelige Forskningsråd) i de seneste ti år gennemført forskellige initiativer for at støtte, at sundhedsvidenskabelige data arkiveres. I 1998 nedsatte Forskningsrådet et udvalg vedrørende sundhedsvidenskabelige data med det formål at foreslå initiativer, der kunne fremme dokumentation og arkivering af sundhedsvidenskabelige data [2]. Et andet initiativ var en bevilling fra Grundforskningsfonden i 1996, som gjorde det muligt at oprette Enheden for Registrering og Arkivering af Sundhedsvidenskabelige Data (ERAS) ved DDA. Enhedens overordnede formål var at sikre, at datamaterialer dokumenteres og arkiveres med henblik på såvel efterkontrol som ny udnyttelse af data. ERAS fungerede i perioden 1996-2002, og projektet viste, at det var muligt at få arkiveret flere data med et uudnyttet forskningspotentiale, og at der var positiv interesse i forskningsmiljøerne for at lade data arkivere [3, 4]. I 2004 sikrede en etårig bevilling fra Det Sundhedsvidenskabelige Forskningsråd en videreførelse af dette initiativ med oprettelsen af DDA Sundhed, og fra 2005 er enheden sikret økonomiske midler via forskningsreserven [5]. Det indebærer blandt andet, at Forskningsrådet for Sundhed og Sygdom i bevillingsskrivelsen for relevante projekter forudsætter, at data efterfølgende arkiveres i DDA Sundhed. DDA Sundhed modtager kopi af disse klausuleringskrivelser fra forskningsrådene og tilbyder forskerne hjælp til at tilrettelægge dataarbejdet på en sådan måde, at dokumentationen af data er i orden fra projektets begyndelse.

Bevaring af forskningsdata i Dansk Data Arkiv Sundhed

På nuværende tidspunkt er der identificeret og arkiveret over trehundrede datamaterialer, hvoraf nogle af de væsentligste er store danske databaser som f.eks. Bedre Sundhed for Mor og Barn, Befolkningsundersøgelserne i Glostrup, Det Danske Tvillingeregister og Skolelægejournalregisteret. Der er desværre også eksempler på, at datasæt, som ville have været interessante for yderligere opfølgende analyser, er gået tabt. Det drejer sig blandt andet om Sundhedsstyrelsens Sygdomsundersøgelse i 1951-1954 [6] og undersøgelsen af københavnske børns boligforhold og sygelighed i 1948-1952 [7, 8]. Arkivering i DDA sikrer ikke blot, at forskerne har en kopi (backup) af deres data, men også at data bevares i et systemuafhængigt arkivformat (OSIRIS), som er læsbart i fremtiden. DDA Sundhed opbevarer også personhenførbare data, og arkivering er

VIDENSKAB OG PRAKSIS | STATUSARTIKEL

Tabel 1. Dansk Data Arkiv Sundheds databeholdning fordelt efter adgang.

Adgangskategori ^a	Antal studier
Fri adgang	44
Fri adgang, men publikation kræver tilladelse	37
Adgang kræver tilladelse fra den forsker eller institution, der har afleveret data	191
I almindelighed ingen adgang	32
Andet	19
I alt	323

a) Adgangskategorien sættes af den forsker eller institution, der afleverer data. Den omfatter udelukkende udlevering af anonymiserede data.

det eneste alternativ til den sletning, som Datatilsynet fastsætter. Derudover kan datamaterialet herunder også personhenførbare data stilles til rådighed for andre til nye analyser.

Sidstnævnte efter godkendelse fra Datatilsynet.

Bevillingen til DDA Sundhed har gjort det muligt, at enheden også medfinansierer dokumentation og oparbejdning af undersøgelser direkte i forskningsmiljøerne. De undersøgelser, der i første omgang indgik i denne ordning, var nogle, som blev anset for at være specielt værdifulde. Midlerne til ekstern oparbejdning har indtil videre sikret dokumentation af 17 undersøgelser, hvoraf et par stykker har været særdeles nødlidende mht. datadokumentation, bl.a. fordi bemandingen er forsvundet ved forskningsprojektets afslutning, eller fordi arkiveringen har været vanskelig at sætte i system. Det gælder f.eks. Den Århusianske Fødselskohorte, der rummer et komplekst datamateriale, som har været underlagt omskiftelige forhold bl.a. pga. kohortens opstart som tidsbegrænset projekt. Udviklingen har i de seneste år givet bekymring blandt initiativtagerne til kohorten for, at »hukommelsen« omkring data skulle forsvinde. Datamaterialet er nu med DDA Sundheds mellemkomst bevaret for eftertiden.

Dette års bevilling til ekstern oparbejdning uddeles efter forslag fra et nyoprettet arkiveringsnetværk, som blev etableret i 2006 med det formål at sætte fokus på dataarkivering og synliggøre DDA Sundhed i forskningsmiljøerne. Netværket har i dag 64 deltagere fordelt på 54 forskningsenheder.

Genbrug af forskningsdata

En anden målsætning for DDA Sundhed er at stille forskningsdata til rådighed for genanvendelse og reanalyse. Der er i alt 323 datamaterialer i enhedens beholdning (Tabel 1). De kan findes og bestilles gennem det omfattende elektroniske søgekatalog på hjemmesiden (sundhed.dda.dk), hvor man kan søge på studie, emneord eller periode. Tabel 1 viser antallet af datamaterialer fordelt på adgangskategori. Adgangskategorien sættes af den forsker eller institution, der afleverer data. Den omfatter udelukkende udlevering af data i anonymiseret form og kun til videnskabelig eller statistisk brug. Ved bestilling af datamaterialer skriver man desuden under på, at datamaterialerne kun anvendes til de(t) formål, de er anmodet

om udlevering til, og at de ikke videregives til tredjepart. Desuden forpligter rekvirenter sig til at kreditere Dansk Data Arkiv og de personer, der oprindeligt varetog undersøgelsen, ved enhver form for publicering. Data skal slettes ved projektets afslutning.

Selve bestillingen sker ved at udfylde og fremsende en rekviritionsblanket. Bestillingen skal vedlægges en problemformulering eller beskrivelse af formålet med de analyser, man ønsker at fortage. Data udleveres i ønsket format f.eks. SAS eller SPSS, med elektronisk post eller på cd-rom. Det er planen i fremtiden at etablere onlineadgang til udvalgte datamaterialer. Indtil videre har erfaringer med udlevering af data til genbrug været begrænsede, formentlig fordi mange forskere fortsat får data direkte fra forskningsmiljøerne. Derudover er der formentlig ikke tilstrækkelig kendskab i de sundhedsvidenskabelige institutioner til de mange datamaterialer, som er til rådighed. Eksempelvis er der et stort uudnyttet potentiale for ph.d.- og specialestudierende. Datamaterialerne kan bl.a. anvendes til inspiration til nye forskningsprojekter, eller de kan anvendes direkte til nye analyseformål. Man kan løbende blive orienteret om nye materialer i DDA nyhedsbrev DDAInfo, som udkommer tre gange om året. Der er desuden mulighed for at abonnere på nyheder i elektronisk form via hjemmesiden.

DDA Sundheds fremtid afhænger af en forlængelse af bevillingen, og den forudsætter, at enheden bliver det naturlige valg i danske forskningsmiljøer for bevaring og formidling af danske forskningsdata på sundhedsområdet – det vil sige, at vi også inden for lægevidenskab bruger de mange muligheder som DDA Sundhed stiller frit til rådighed.

Korrespondance: *Merete Osler*, Institut for Folkesundhedsvidenskab, Københavns Universitet, DK-1014 København K. E-mail: m.osler@dadlnet.dk

Antaget: 21. juni 2007
Interessekonflikter: Ingen

Litteratur

1. Manuskriptvejledning for Ugeskrift for Læger. 22. udgave, januar 2006. Ugeskr Læger 2006;168:67-73.
2. Udvalget vedr. Sundhedsvidenskabelige data. Opbevaring og arkivering af sundhedsvidenskabelige forskningsdata. København: Statens Sundhedsvidenskabelige Forskningsråd, 2000.
3. Olsen J, Lyng E, Osler M et al. Hvad er fremtiden for arkivering af forskningsdata fra sundhedsvidenskab? Ugeskr Læger 2003;165:811.
4. Lyng E, Madsen M, Noack J et al. Arkivering af sundhedsvidenskabelige data. København: Danmarks Grundforskningsfond, 2002.
5. Wittendorff H, Bredahl L. Om DDA Sundhed – Et initiativ for sundhedsvidenskab. Metode & Data 2005;91:43-51.
6. Lindhardt M. Sygdomsundersøgelsen i Danmark 1951-1954. En statistisk analyse af den almindelige sygelighed samt enkelte andre forhold i forbindelse hermed. København: Munksgaard, 1960.
7. Christensen V. Boligforhold og børnesygelighed. En undersøgelse af hospitalsindlæggelsen under forskellige boligforhold. København: Munksgaard, 1956.
8. Andersen TF. Persistence of social and health problems in the welfare state: a Danish cohort experience from 1948 to 1979. Soc Sci Med 1984;18:555-60.