

## GIS og geodata i sundhedsforskning

Rikke Baastrup

*I 1800-tallets London brugte John Snow kortlægning af kolera-smittede personer til at udpege den brønd i Broad Street, som var årsag til koleraepidemien (Koch 2005), men hvordan bruges geografiske metoder, geodata og GIS i nutidens sygdomsforskning?*

### Introduktion

Det er velkendt, at menneskers helbred påvirkes af det omgivende miljø. En erkendelse som også fremgår af regeringens plan *Miljø og sundhed hænger sammen* fra 2003 (Regeringen 2003). Luftforurening, drikkevandskvalitet, ioniserende stråling og støj er blot nogle af de miljøfaktorer, som har indflydelse på vores sundhed. Miljøepidemiologer beskæftiger sig med at vurdere sundhedsrisikoen af miljøpåvirkninger, og i sådanne studier er det største problem oftest at kunne koble miljøfaktor og undersøgelsesbefolkning sammen, således at individuel eksponering kan bestemmes. Med de danske miljø- og sundhedsregistre er der gode betingelser for miljømedicinsk forskning i Danmark, men der findes ingen fælles nøgle mellem registrene (Poulstrup 2003). Derfor er geokodning og rumlig relation mellem registrene en god og effektiv metode til at knytte miljøpåvirkninger og sygdomstilfælde sammen. Med andre ord er den rumlige kobling mellem registrene essentiel for at undersøge sammenhæng mellem miljø og sundhed. I denne artikel vil jeg beskrive, hvordan kompetencer i geografi og geoinformatik er blevet brugt i et studie af arsen i drikkevand og risiko for kræft, et studie som er udført ved Kræftens Bekæmpelse.

### Baggrund

Af skadelige stoffer i drikkevand er arsen det, der er studeret mest omfattende, og det internationale agentur for kræftforskning ved WHO slår fast at arsen i drikkevand er kræftfremkaldende for mennesker (International Agency for Research on Cancer 2004). En række studier fra Asien og Latinamerika har fundet en sammenhæng mellem høje koncentrationer (op til flere hundrede µg/l) af arsen i drikkevand og udvikling af visse kræftformer (se f.eks. Marshall et al. 2007; Tsuda et

al. 1995; Wu et al. 1989). I lande med mindre forekomster i drikkevandet (100-1000 gange lavere) har undersøgelserne af sammenhængen mellem arsen og kræft ikke givet entydige svar (se f.eks. Karagas et al. 2001; Kurtio et al. 1999; Steinmaus et al. 2003). Den gennemsnitlige arsenkoncentration i det danske drikkevand er 1,3 µg/l. Hos forskningsprogrammet for *Miljø og Kræft* ved Kræftens Bekæmpelse var der derfor et ønske om at undersøge, om arsenkoncentrationerne i det danske drikkevand havde betydning for kræfttrisikoen, til trods for relativt små koncentrationer i det danske drikkevand. Formålet med studiet var derfor at undersøge, om arsenkoncentrationerne i det danske drikkevand giver anledning til en øget risiko for udvikling af kræft.

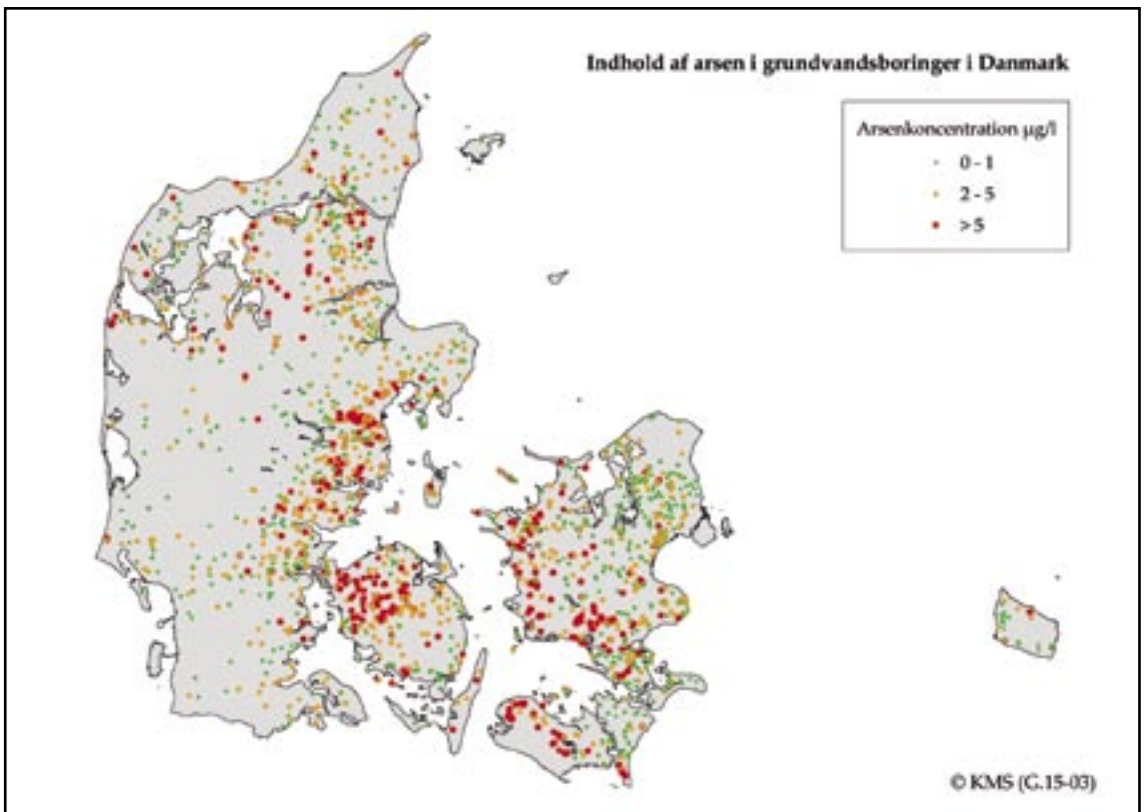
### Metode

Arsen forekommer naturligt i den danske undergrund, hvorfra det udvaskes til vores grundvand. Det findes især i geologiske aflejringer af tertiære marine lerbjergarter; typisk moræneler fra sidste istids gletscherfremstød (GEUS 2004), og derfor er der en geografisk variation i arsenforekomsten i det danske grundvand (se figur 1). Et EU-direktiv fra 2001 sænkede grænseværdien for arsen i drikkevandet fra 50 µg/l til 5 µg/l. Inden udgangen af 2003 skulle vandværkerne have opfyldt de nye krav, men det er stadig et problem for ca. 5 % af de danske vandværker, fordi de henter deres vand fra områder med naturligt forekommende arsen. Man forsøger at løse problemet ved at rense vandet bl.a. ved tilsætning af jern (Rasmussen 2006). Siden 2001 har det samtidig været obligatorisk for de danske vandværker at indberette koncentrationer af arsen i drikkevandet til *drikkevandskontrollen*, som varetages af GEUS. På den baggrund kunne vi kortlægge indholdet af arsen i drikkevandet hos de dan-

ske vandværker. Kendskabet til forekomsten af arsen i det danske drikkevand er historisk set meget begrænset, fordi det først sent blev inkluderet i måleprogrammet. Forskerne fra GEUS fandt det dog rimeligt at antage, at de nuværende målinger også afspejlede de historiske koncentrationer.

For at kunne undersøge, om indtagelse af arsen gennem drikkevandet øger risikoen for kræft, kræver det en undersøgelsesbefolkning, for hvilken arsenindtaget kan bestemmes. Ved Kræftens Bekæmpelse har man etableret en såkaldt kohorte ved navn *Kost, kræft og helbred*. Det er en undersøgelse, hvor man har indsamlet spørgeskemaoplysninger om kost og livsstil samt blod-, urin- og fedtvævsprøver fra 57.000 personer i Danmark. I epidemiologiske undersøgelser

er det vigtigt at kunne justere sine analyser for andre faktorer, som har betydning for den pågældende sygdom; det kunne eksempelvis være rygning. På den måde forsøger man at udelukke, at den effekt man eventuelt ville finde af arsen skyldes andre faktorer. Det betyder, at detaljerede oplysninger om den population, man vil undersøge, er vigtige for at gennemføre et validt studie. Derfor var det oplagt at bruge *Kost, kræft og helbred*-kohorten i studiet af arsen. Deltagerne i undersøgelsen blev rekrutteret fra Københavns og Frederiksberg Kommune, det meste af Københavns Amt, Århus Kommune samt Hinnerup og Hørning Kommune i Århus Amt. De var 50 – 64 år i 1993 – 1997, hvor kohorten blev etableret. Et kriterium for at blive inkluderet i undersøgelsen var, at deltagerne ikke havde eller havde haft kræft ved inklu-

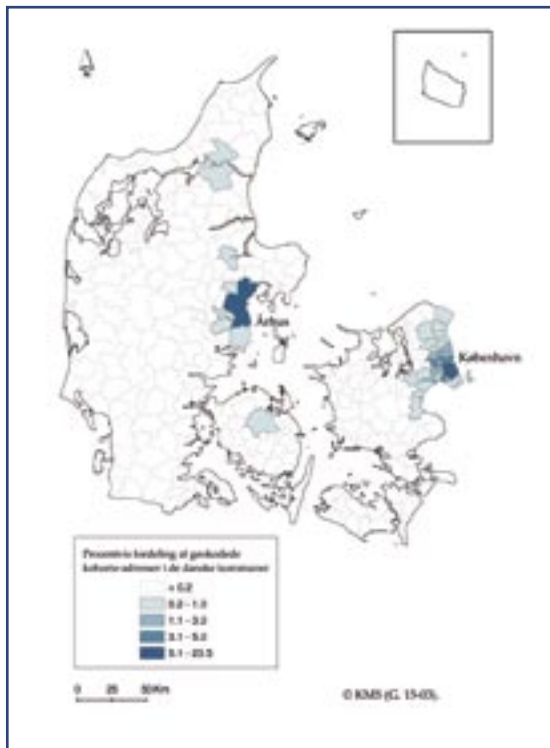


Figur 1. Arsenkoncentrationer i det danske grundvand. Kilde: grundvandsdata fra Jupiter-databasen, GEUS.

sionen i kohorten, på den måde fik man en population, hvori det i den efterfølgende tid var muligt at følge kræftudviklingen. Det sker ved, at man løbende udtrækker oplysninger fra det danske cancerregister for at følge op på hvilke deltagere, der har fået kræft. Cancerregistret var opdateret indtil 2003, da dette studie blev gennemført, så studieperioden strakte sig indtil 2003.

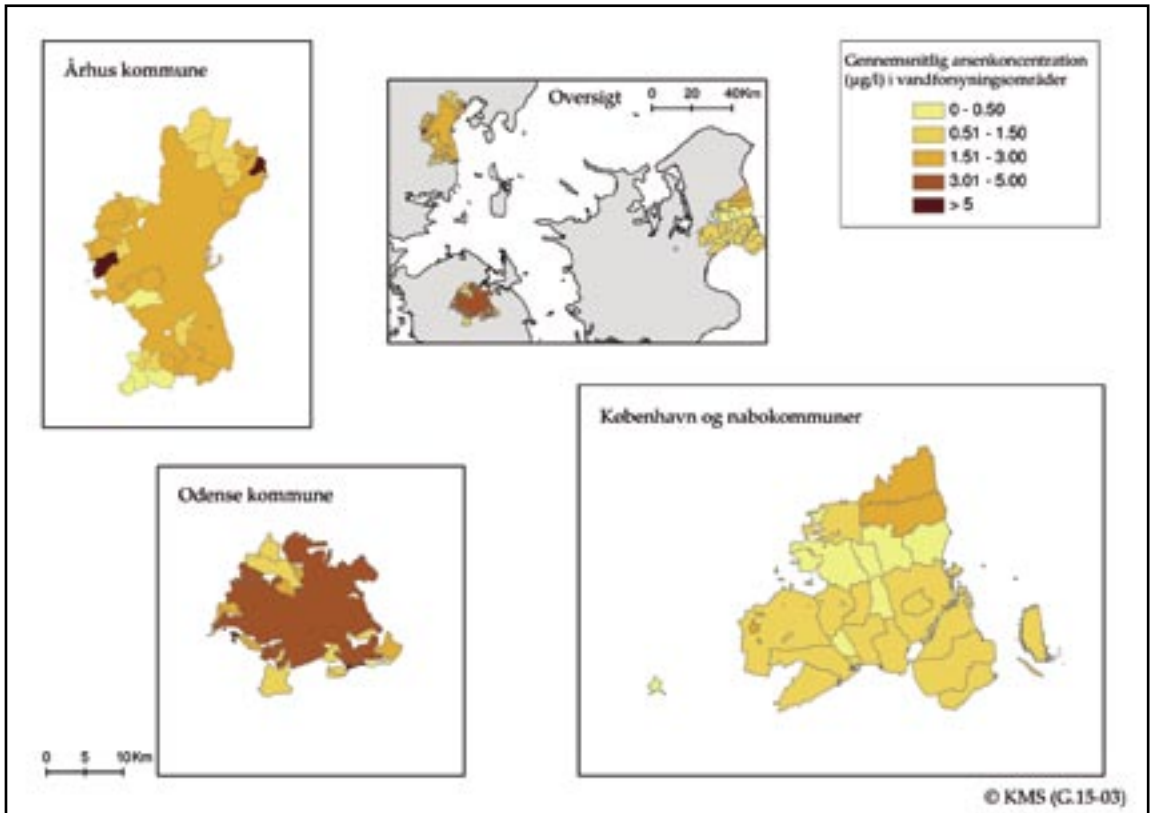
For at kunne sammenkæde deltagerne med arsenkoncentrationerne i drikkevandet blev oplysninger om deres bopælsadresser udtrykt fra det Centrale Person Register (CPR) vha. deres personnummer. Det resulterede i 202.339 forskellige bopælsadresser både i og udenfor optageområderne, som deltagerne havde haft i perioden 1970 – 2003. CPR blev oprettet i 1968 (Møller 2002), så derfor var det ikke muligt at spore deltagernes adresser længere tilbage i tiden. 1970 blev derfor valgt som studiets starttidspunkt. Adresserne blev geokodet ved hjælp af OSAK-adresserne. Vi dannede en adressenøgle ved at sammenstille kommunekode, vejkode, husnummer og evt. bogstav, så nøglen kunne bruges til at sammenkæde kohortemedlemmernes adresser med adressekoordinaterne i OSAK. 91,9 % af adresserne i kohorten fik således tilført geografiske koordinater. Det lykkedes at placere yderligere 6,3 % af adresserne ved at modificere adressenøglen på forskellige måder, f.eks. ved at fjerne bogstaver fra nøglen eller ved at lægge to til eller trække to fra husnummeret. 1,8 % adresser kunne ikke geokodes bl.a. fordi en del adresser havde vejkode i intervallet 9900 – 9999, og disse koder anvendes til at registrere personer uden fast bopæl, hjemløse samt udsendte f.eks. soldater og diplomater (Lind 2006). Figur 2 viser, hvordan de geokodede kohorte-adresser fordeler sig på de danske kommuner.

Oplysninger om vandforsyningsområder blev indsamlet fra kommuner og vandværker, for at kunne bestemme tilhørsforholdet mellem bopælsadresser og vandforsyning. Da der ikke var ressourcer til at indsamle oplysninger om vandforsyningsområder for hele lan-



Figur 2. Fordeling af geokodede kohorte-adresser ( $n=198.758$ ) i de 271 "gamle" danske kommuner. Proportionerne er beregnet som antal af geokodede kohorte-adresser i hver kommune delt med det totale antal geokodede kohorte-adresser.

det, koncentrerede dataindsamlingen sig primært om de områder, hvor størstedelen af kohortemedlemmernes adresser var lokaliseret – nemlig i de to rekrutteringsregioner (Århus og København) samt kommuner, hvor vandforsyningsdata var lettilgængelige (Odense kommune). I alt blev der indsamlet 94 forsyningsområder fra 24 kommuner. I nogle kommuner var der en eller anden form for GIS implementeret i forvaltningen, så forsyningsområderne kunne udleveres i digitalt GIS-format. I andre tilfælde blev områderne digitaliseret ud fra trykte kort i vandforsyningsplaner eller andet materiale fra kommunen. Figur 3 viser de indsamlede forsyningsområder klassificeret efter arsenkoncentrationen på det/de vandværker, som forsyner området.



Figur 3. Kortlagte forsyningsområder klassificeret efter arsenkoncentration.

Vi brugte *Spatial Join*-funktionaliteten i ArcGIS til at etablere relationen mellem vandforsyningsområder og kohortemedlemmernes adresser. 84% af de geokodede adressepunkter fik således tildelt attributterne fra den område-polygon, som de var placeret indeni. 14% af adresserne blev knyttet til det vandværk som lå nærmest, og de ca. 2% adresser, som ikke kunne geokodes, blev tildelt en "manglende" værdi for arsen.

På baggrund af det link mellem bopælsadresse og vandforsyning, som blev skabt i GIS, kunne vi for hvert medlem af kohorten opsummere den individuelle arseneksponering over tid. Arsenkoncentrationen ved hver adresse blev vægтет i forhold til hvor lang tid personerne havde boet på adressen og i forhold til hvor stort dagligt indtag af postevand kohortemed-

lemmerne havde rapporteret i spørgeskemaet. Vi beregnede to mål for arseneksponeringen; en tidsvægtet gennemsnitlig eksponering og en kumuleret eksponering. Begge mål inkluderede alle kendte adresseperioder for hvert medlem af kohorten.

De statistiske analyser blev udført i den statistiske programpakke SAS, hvor en regressionsmodel blev brugt til at beregne om de medlemmer af kohorten, som havde indtaget mest arsen gennem drikkevandet, havde en statistisk signifikant større risiko for at udvikle kræft end de mindre eksponerede deltagere. Analyserne blev justeret for alle de kendte kræfttrisikofaktorer, som der fandtes oplysninger om i spørgeskemaerne, for at udelukke at resultaterne skyldtes andre faktorer end arsen.

## Resultater

Analyserne viste, at der var en signifikant omvendt sammenhæng mellem arseneksponering og risiko for non-melanoma (ikke ondartet) hudkræft, hvilket kan tolkes som en beskyttende effekt af arsen i forhold til udvikling af non-melanoma hudkræft. For de andre kræftformer, som var inkluderet i analyserne (lunge-, blære-, nyre-, lever-, prostata-, tarm- og brystkræft samt melanoma (modermærke) hudkræft), fandt vi ingen sammenhæng mellem arsen og kræftrisiko. Yderligere analyser viste at den inverse sammenhæng mellem arseneksponering og non-melanoma hudkræft kunne skyldes andre regionale forskelle mellem personer rekrutteret fra Århusområdet og Københavnsområdet. Samtidig så vi en positiv sammenhæng mellem arseneksponering og risiko for brystkræft, som dog kun var signifikant blandt kvinder rekrutteret fra Århusområdet, og når vi brugte det ene af de to eksponeringsmål. Resultaterne stemmer ikke overens med tidligere studier fra Asien og Latinamerika, men enkelte tidligere studier (Karagas et al. 2002; Snow et al. 2005) har antydnet, at små doser af arsen kan have en beskyttende effekt. Vi er ikke bekendt med, at andre studier har fundet en sammenhæng mellem arsen og brystkræft, og den sammenhæng, vi fandt i dette studie, var kun signifikant for kvinderne i Århus for det gennemsnitlige eksponeringsmål. Derfor er det sandsynligt at dette resultat skyldes tilfældigheder frem for en reel sammenhæng. Vores resultater skal tolkes med forsigtighed, og der skal flere studier til, for

at bekræfte at arsen muligvis mindsker risikoen for non-melanoma hudkræft og øger risikoen for brystkræft.

## Konklusion

Studiet viste at de globalt set beskedne mængder arsen i det danske drikkevand ikke øgede risikoen for de kræftformer, som var inkluderet i projektet, bortset fra brystkræft blandt kvinder i Århusområdet, men arsen kan muligvis have en beskyttende effekt imod non-melanoma hudkræft. Dette bør dog undersøges nærmere.

Alternativet til at bruge GIS i dette studie ville have været at indhente kundeoplysninger fra alle landets vandværker for at kunne koble arsendata med personerne i kohorten eller at foretage biologiske målinger af arsen i f.eks. negle eller urin fra alle 57.000 personer. Disse løsninger ville dog være økonomisk og tidsmæssigt højest urealistiske. Den vigtigste konklusion man kan drage af projektet er derfor, at GIS var et essentielt værktøj til at estimere arseneksponeringen gennem drikkevandet, fordi det uden brug af geografiske nøgler og rumlige relationer ikke havde været praktisk muligt at sammenkæde data fra Kost, kræft og helbred-kohorten med arsendata fra de danske vandværker.

Hele studiet bliver publiceret i *Environmental Health Perspectives* vol. 116, februar 2008. Foreløbig kan den videnskabelige artikel findes online på følgende adresse: <http://www.ehponline.org/docs/2007/10623/abstract.html>.

## Referencer

- GEUS. 2004. Grundvandsovervågning 1998-2003. København:GEUS, Miljøministeriet.
- International Agency for Research on Cancer. 2004. Some Drinking-water Disinfectants and Contaminants, including Arsenic.
- Karagas MR, Stukel TA, Morris JS, Tosteson TD, Weiss JE, Spencer SK, Greenberg ER. 2001. Skin cancer risk in relation to toenail arsenic concentrations in a US population-based case-control study. *Am J Epidemiol* 153:559-565.
- Karagas MR, Stukel TA, Tosteson TD. 2002. Assessment of cancer risk and environmental levels of arsenic in New Hampshire. *Int J Hyg Environ Health* 205:85-94.
- Koch T. 2005. John Snow: The London epidemics. In: *Cartographies of Disease: Maps, Mapping and Medicine* Redlands, California:ESRI Press, 75-104.
- Kurttio P, Pukkala E, Kahelin H, Auvinen A, Pekkanen J. 1999. Arsenic concentrations in well water and risk of bladder and kidney cancer in Finland. *Environ Health Perspect* 107:705-710.
- Lind M. 2006. Specialkonsulent, National Geodatabank, Kort & Matrikelstyrelsen. Personlig kommunikation.
- Marshall G, Ferreccio C, Yuan Y, Bates MN, Steinmaus C, Selvin S, Liaw J, Smith AH. 2007. Fifty-year study of lung and bladder cancer mortality in Chile related to arsenic in drinking water. *J Natl Cancer Inst* 99:920-928.
- Møller JØ. 2002. Personnummeret i det centrale personregister (CPR). <http://www.cpr.dk/publikationer/pnr-notat%20ny%20skrift.htm>.
- Poulstrup, A. 2003. Forprojekt om samkøring af registerdata for miljø og sundhed. København:Miljøstyrelsen.
- Rasmussen MB. 2006. Naturen forgifter drikkevandet i Asien. *Miljø Danmark* 1:35-37.
- Regeringen. 2003. Miljø og sundhed hænger sammen. København:Miljøstyrelsen.
- Snow ET, Sykora P, Durham TR, Klein CB. 2005. Arsenic, mode of action at biologically plausible low doses: What are the implications for low dose cancer risk? *Toxicol Appl Pharmacol* 207:557-564.
- Steinmaus C, Yuan Y, Bates MN, Smith AH. 2003. Case-control study of bladder cancer and drinking water arsenic in the western United States. *Am J Epidemiol* 158:1193-1201.
- Tsuda T, Babazono A, Yamamoto E, Kurumata-ni N, Mino Y, Ogawa T, Kishi Y, Aoyama H. 1995. Ingested arsenic and internal cancer: a historical cohort study followed for 33 years. *Am J Epidemiol* 141:198-209.
- Wu MM, Kuo TL, Hwang YH, Chen CJ. 1989. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *Am J Epidemiol* 130:1123-1132.

## Om forfatteren

Rikke Baastrup er forskningsassistent ved Institut for Epidemiologisk Kræftforskning, Kræftens Bekæmpelse, hvor hun har skrevet speciale. Hun er kandidat fra Geografi ved Københavns Universitet. e-mail: baastrup@cancer.dk