

## En ny spiller på den danske GIS-arena: Statens Serum Institut

*Martin Rudbeck Jepsen, Steen Ethelberg og Gert Nicolaisen*

*Statens Serum Institut (SSI) har længe beskæftiget sig med overvågning af smitsomme sygdomme i Danmark. Mange sygdomstilfælde indrapporteres til SSI med den smittedes cpr-nummer, og vedkommendes adresse bliver noteret. Imidlertid er disse data hidtil blevet aggregeret for at skabe et overblik over en given sygdoms aktivitet på nationalt niveau.*

*De sidste par år er GIS langsomt begyndt at vinde indpas på SSI. Vi er p.t. to afdelinger, Epidemiologisk afdeling og Center for Biologisk Beredskab, som benytter GIS-værktøjer til overvågning af den rumlige fordeling af sygdomme i Danmark og analyse af potentielle modværgforanstaltninger. På Epidemiologisk afdeling er vi ved at etablere en dynamisk sygdomskortlægning ved hjælp af ArcGIS server. Denne kortlægning vil sikre en næsten realtime overvågning af alle rapporterbare humane sygdomme i Danmark, med henblik på tidligt at opdage eventuelle epidemiske udbrud.*

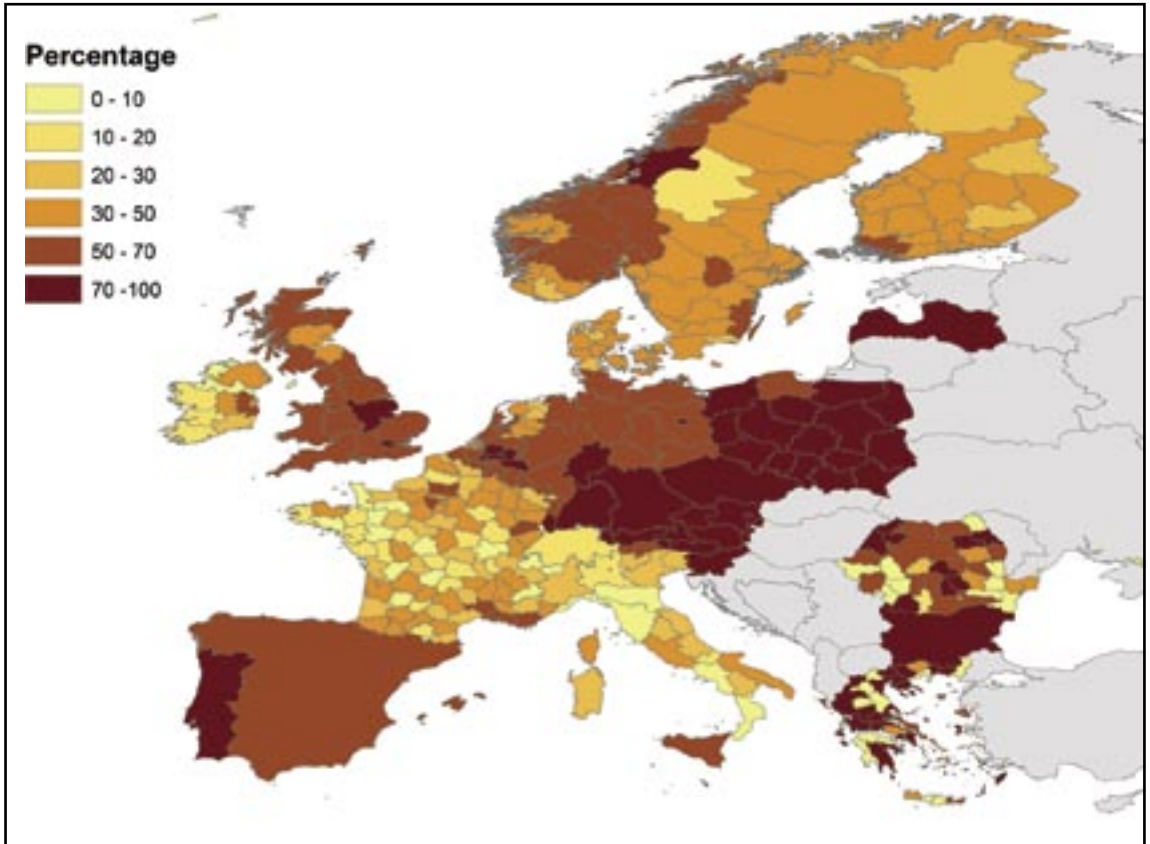
### Introduktion

Statens Serum Institut (SSI) er en sektorforskningsinstitution, der dækker mange aktiviteter bl.a. analyser af prøver taget fra patienter landet over, produktion af vacciner, overvågning af forekomsten af smitsomme sygdomme samt undersøgelser af udbrud mhp at afdække smitekilder. Sidstnævnte relateres ofte til den første praktiske anvendelse af GIS, nemlig den engelske læge John Snows opklaring af et koleraudbrud i 1850'ernes London (Vinten-Johansen et al 2003). Dr. Snow var en snedig rad, som bekymret over et stigende antal koleratilfælde kortlagde patienternes bopæle og fandt dem koncentreret i et afgrænset område. På basis af den indsigt gennemførte John Snow en epidemiologisk undersøgelse og fandt, at de tilspurgte patienter alle hentede drikkevand fra samme brønd. Det kan derfor synes ejendommeligt at SSI og den danske sundhedssektor generelt først i de seneste år har fattet interesse for anvendelsen af GIS i overvågning og analyse af forskellige sygdommes forekomster - men nu er vi på vej! Vi vil i det følgende præsentere de aktiviteter vi har igangsat på Epidemiologisk afdeling, hvor vi overvåger udbredelsen af en række smitsomme sygdomme, og på Center for Biologisk Beredskab (CBB), som står for beredskabet vedrørende biologiske kampstoffer og bioterrorisme.

### Salmonella Atlas

I forbindelse med EUs sjette rammeprogram (FP6) blev der givet midler til etablering af et

forskningsnetværk, MedVetNet (MVN 2008) indenfor sygdomme som smitter fra dyr til menneske, såkaldte zoonoser. MedvetNet omfatter ca 30 arbejdsplaner, der hver for sig beskæftiger sig med en række projekter. Den sjette arbejdsplan, WP6, beskæftiger sig med anvendelse af GIS i forbindelse med overvågning og analyse af zoonoser, og WP6 ledes fra Epidemiologisk afdeling på SSI. Ét af de første projekter i WP6 har været at opbygge et digitalt Salmonella Atlas som viser den årlige forekomst af salmonellainfektioner hos mennesker i forskellige EU lande. Atlasset dækker perioden 1994 - 2006, og indeholder geografiske informationer på sub-nationalt niveau. Det grundliggende datasæt stammer fra et centralt europæisk overvågningsnetværk, Enternet (Enternet 2007) som modtager rapport om påviste salmonellatilfælde fra medlemslandene. Nogle lande rapporterer på et administrativt niveau svarende til en meget detaljeret rumlig skala, mens andre lande rapporterer på noget grovere niveauer. Fælles for alle lande er, at der ikke anvendes standardiserede koder for de administrative inddelinger af de respektive lande, og en stor del af arbejdet med opbygningen af Salmonellaatlasset har derfor været at rense og geokode de rapporterede data, så de passer til de officielle NUTS-koder (Eurostat 2008) som gælder for EU. NUTS-koderne spænder fra det nationale niveau (NUTS 0) til regionale inddelinger (NUTS 3) som f.eks. de (historiske) danske amter. Hvor data ikke har kunnet geoko-



Figur 1. Skærmdump fra [www.epigis.dk](http://www.epigis.dk). Billedet viser forekomsten af *S. enteritidis* i 2004.

des på regionalt niveau er data summeret på nationalt niveau.

Atlasset er produceret med en standard ArcGIS 9.1TM desktop, hvor vi har produceret kort for hvert år i perioden 1994-2006 (13 år) for hver af de 10 mest hyppigt forekommende salmonellatyper. Yderligere er kortene produceret som nationale kort og som kort på NUTS 1-3 niveau (afhængigt af rapporteringsniveauet hos de medvirkende lande). Resultatet er 260 kort som vi har gemt i et almindeligt billedformat og publiceret på adressen <http://www.epigis.dk>. Kortene viser for hver af de 10 salmonellatyper forekomsten (incidensen) per 100.000 indbyggere i de respektive regioner/lande.

En del af den forskel i salmonellaforekomsten mellem lande, som kan observeres på sal-

monellaatlasset, skyldes forskelle i landenes indberetningssystemer og indberetningskultur. I visse lande er man mindre tilbøjelige til at tage og indsende prøver fra patienter med symptomer på en salmonellainfektion, eller måske mindre tilbøjelig til at indrapportere data til en central instans end i andre lande, og salmonellaforekomsten vil for førstnævnte steder synes lavere. For at udviske denne tendens har vi derfor også produceret en række kort på regionalt niveau som viser en given salmonellatypes forekomst som andel af det samlede antal indrapporterede tilfælde for en given region (Figur 1).

Salmonellaatlasset har været et godt udgangspunkt for anvendelsen af GIS ved SSI, og har vakt en del opmærksomhed i det

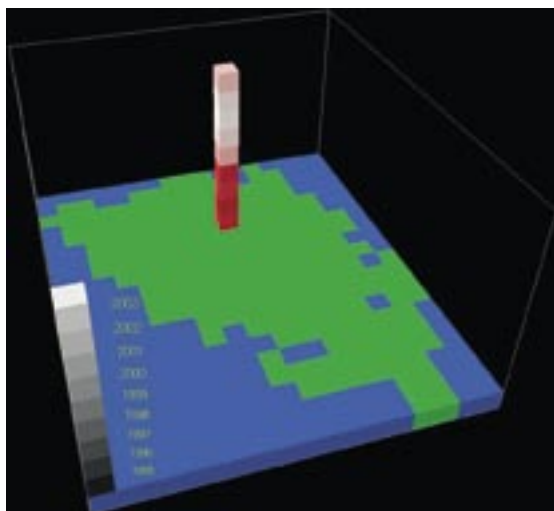
europæiske miljø for zoonotiske sygdomme som hidtil kun har haft begrænset tradition for at anvende GIS. Anvendelsen af ArcGIS desktop til produktion af statiske billedfiler, som efterfølgende publiceres på internettet, har vist sig en driftsikker løsning, som dog involverer en del manuelt arbejde. Vi har derfor indkøbt en ArcGIS Server for at kunne producere en webgisløsning, hvor internetbrugeren selv specificerer for hvilke typer og år, der skal genereres kort.

### WebGIS-baseret dansk sygdomsovervågning

Det er planen, at vores nyindkøbte ArcGIS server skal leve et travlt liv. Et nyt, spændende phd-projekt på epidemiologisk afdeling benytter bl.a. data fra de danske vagtlæger til at overvåge forekomsten af influenzalignende sygdomme i Danmark. Læge og phd-studerende Katja Qureshi modtager dagligt elektroniske data fra vagtlægernes rapporteringssystem. Datasættet indeholder en record for hver tilset patient, og en angivelse af, om patienten har en influenza-lignende sygdom. Data er anonymiserede, men angiver i hvilket postdistrikt patienten er hjemmehørende, og vi kan derfor joine disse data til et vektorlag over de danske postdistrikter og dermed skabe en næsten real-time kortbaseret overvågning af den rumlige fordeling af influenzalignende sygdom i Danmark. Det er hensigten, at dette overvågningsystem skal supplere de nuværende informationer om influenzaaktivitet på SSIs hjemmeside (SSI 2008) og tilføje øget funktionalitet som f.eks. design af kort af influenzaforekomsten indenfor brugerspecificerede perioder. Det er også meningen, at dette system skal tjene som prototype for yderligere systemer til overvågning af de øvrige sygdomme, som epidemiologisk afdeling overvåger.

### 3D analyse af clusters

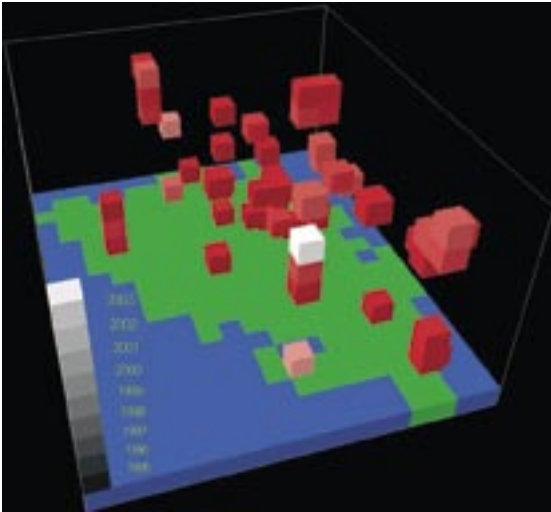
I lighed med den europæiske overvågning af salmonella modtager SSI også danske indberetninger om humane infektioner med campylobacter. Campylobacterbakterien er oftest fødevearbåren, men menes i nogle tilfælde også at smitte via 'økologiske' vektorer, f.eks.



Figur 2. 3D repræsentation af campylobacterdata for Odense. Hvide celler viser studieområdet (Fyn og omkringliggende øer). X- og Y-akserne repræsenterer rum, mens Z-aksen repræsenterer tid. Data er aggregeret i celler med dimensionerne 5\*5 km \*1 år. Her ses Fyn fra Syd-sydvest, med øhavet i forgrunden.

ved badning i kontamineret vand. Hvis man antager, at fødevearbårne 'udbrud' af campylobacter vil være punktføremønstre i tid og rum mens infektioner grundet økologiske forhold vil være mere persisterende i tid og rum vil vi derfor kunne identificere de 'økologiske infektioner' som klynger af høje campylobacterforekomster. Som input til denne analyse har vi taget et udtræk fra en database med bopæl på alle personer, for hvem der er indsendt en fæcesprøve til analyse, samt bopæl for alle de personer hvor prøven er testet positiv for campylobacter. Vi har begrænset udtrækket til personer fra Fyn og perioden 1995-2003, og vi har aggregeret data per år og 5\*5 km celler. Resultatet er et tredimensionelt rum hvor fladen udgøres af x- og y-akserne mens z-aksen repræsenterer tid. Figur 2 viser studieområdet og en 'stak' af data for den 5\*5 km celle der repræsenterer Odense.

For at finde områder hvor campylobacter forekomsten over tid og/eller rum har været højere end gennemsnittet anvender vi



Figur 3. 3D-visualiseringen og clusteranalyserne er udviklet i Netlogo (Wilensky 1999)

Moran's lokale indeks (Mitchell 2005), hvor vi har udvidet algoritmen til at inkludere tid. Figur 3 viser celler identificeret som del af et kluster (høj indekxsværdi) med en z-værdi på 2. Her er kluster-analysen foretaget på antallet af indsendte prøver per indbygger i cellen, og fortæller, at der i disse områder indsendes flere prøver til analyse end man ville forvente. Det bemærkes, at de fleste clusters har en udstrækning i tid fremfor rum (de orienterer sig langs z-aksen) hvilket måske peger på, at der disse steder findes praktiserende læger som er mere tilbøjelige til at indsende prøver end lægerne andre steder på Fyn.

### Agent-baseret modellering

I periferien af SSIs GISaktiviteter har vi beskæftiget os med dynamisk, agent-baseret modellering af sygdomssmitte. Agent-baseret modellering kan være interessant set fra et geografisk perspektiv, fordi det er muligt at anvende stedbunden information, f.eks. til modellering af trafikanters adfærd i rum, og fordi agent-baserede modeller er oplagte til modellering af netværk, spændende fra trafiknetværk til sociale netværk. På epidemiologisk afdeling har vi anvendt Netlogo til at mo-

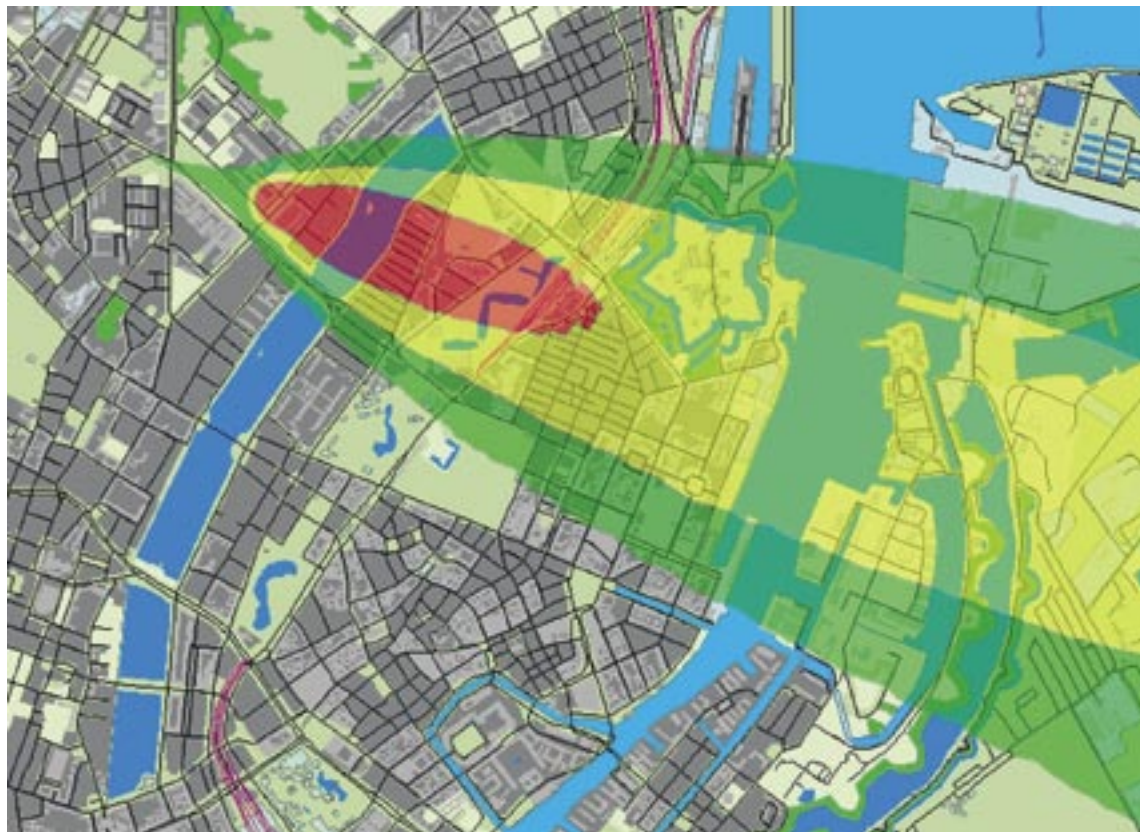
dellere sociale netværk, og set på smittespredningen mellem individerne. Data til op-sætningen af modellens sociale netværk har vi taget fra Danmarks Statistik, bl.a. frekvensfordelinger af husholds- og arbejdspladsstørrelser i hovedstadsområdet. Modellen kan prøves på Center for Biologisk Beredskabs hjemmeside (CBB 2007). Vi har også opbygget en model over spredning af en seksuelt overførbart virus, HPV, som blandt andet kan forårsage livmoderhalskræft. Denne model indgår i Sundhedsstyrelsens vurdering af de nyligt lancerede vacciner mod HPV (Sundhedsstyrelsen 2007).

### CBB's anvendelse af GIS

Center for Biologisk Beredskab er oprettet i 2001 for at kunne imødegå biologiske sikkerhedstrusler. Centrets opgaver omfatter truselsvurderinger, iværksættelse af forebyggende foranstaltninger og modforanstaltninger overfor truslen fra biologiske våben rettet mod landets borgere. En væsentlig del af centrets arbejde er rettet mod prøvetagning, diagnostik og anbefaling af modforanstaltninger f.eks. i forbindelse med pulverbrevshændelser. Ved disse hændelser udarbejder centrets analysestab rutinemæssigt en computerbaseret spredningsmodel, som afbilder det berørte område med et plume (ESRI). Plumet genereres af et program, som CBB har licens til igennem det amerikanske forsvar, der betjener sig af samme program. Det genererede plume kan i form af en shapefil overføres til ArcMap 9.2.

Med plumet i ArcMap åbnes der for visualisering af de geospatiale data på projektor og kort, der kan printes ud og bringes med til møder og på skadestedet, så man ad den vej kan lette kommunikationen af hændelsens udstrækning, og så andre faggrupper hurtigt kan orienteres om begivenhedens omfang uden at den store faglige indsigt er en forudsætning.

Desuden åbnes der med ArcMap for en række analytiske muligheder. Danmarks Statistik tilbyder eksempelvis det såkaldte kvadratnet (Danmarks Statistik 2008), hvor landet er

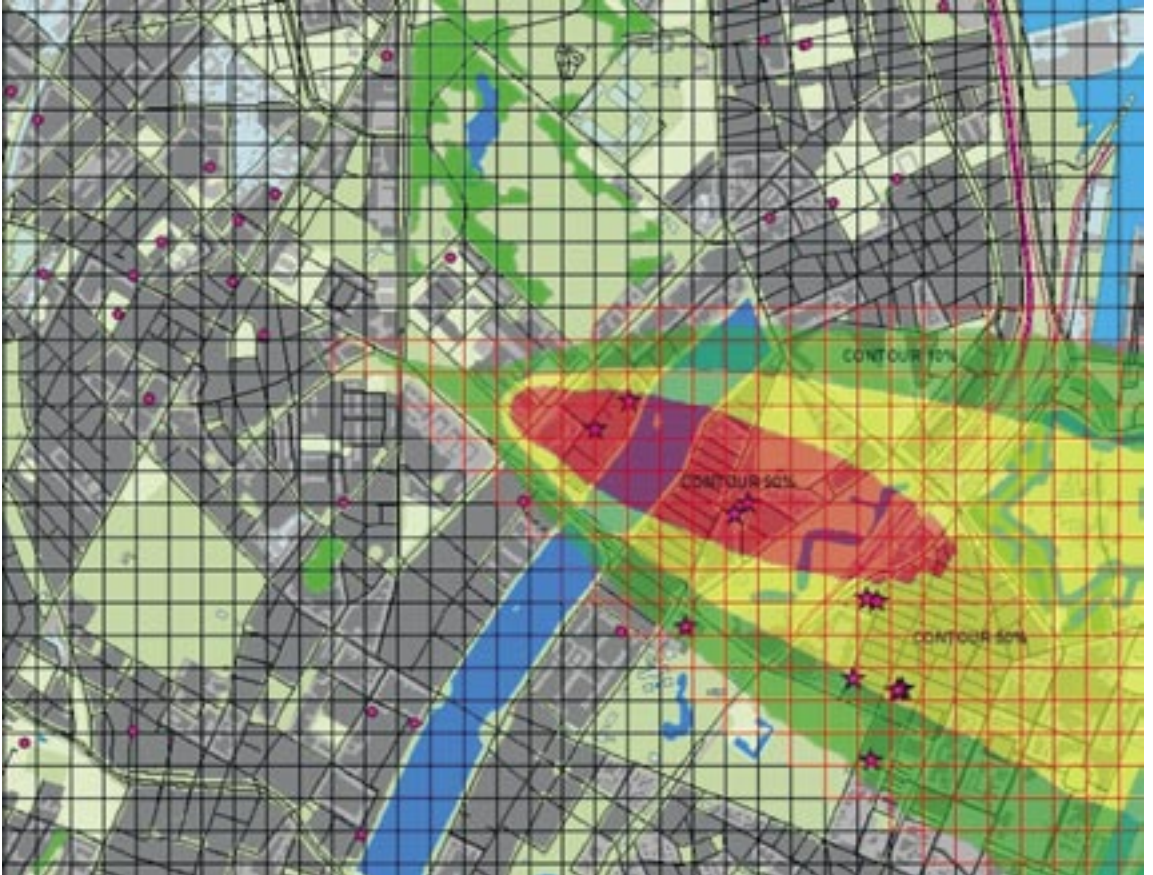


Figur 4. Plume for hændelse ved Rigshospitalet i København overført til ArcMap. Rød zone betegner 90% infektionsrisiko, gul 50% og grøn 10%

delt ind i kvadrater som geospatiale data kan henføres til. Det mindste af kvadratnettene er på 100m x 100m, dette net kan indlæses i ArcMap som shapefil. Til hvert af kvadraterne kan der knyttes en række persondata, som blandt andet giver mulighed for, at få udskrevet hvor mange personer der kan forventes at opholde sig indenfor et givet kvadrat ved dag og ved nat, ligesom personernes køn, alder m.m. kan aflæses. På samme måde kan CVR data over danske virksomheder gøres tilgængelige som shapefiler, og hermed åbnes en vigtig mulighed for at identificere steder af særlig interesse, så som biograf, skoler, stationer og andre lokaliteter, hvor mange personer må forventes at være samlet. Stedernes præcise adresse, telefon-

numre, antallet af ansatte, samt en lang række andre væsentlige data kan aflæses, ligesom lister med de ønskede data kan printes ud, så der kan tages kontakt til de ramte adresser i området og de nødvendige foranstaltninger kan sættes i værk.

Med en toolboks udvælges de kvadrater og specifikke virksomheder, som ligger inde i plumet. Antallet af personer, der kan forventes at opholde sig i de enkelte kvadrater ved dag og ved nat, fastslås, og på samme måde findes navn, adresse, telefonnumre, antallet af ansatte mm. for nøglevirksomheder, institutioner og arbejdspladser inden for plumet. På denne måde skabes datagrundlaget for beredskabets indsats.



Figur 5. Plume i Arc Map viser hændelse ved Rigshospitalet i København med indtegning af 100m x 100m kvadratnet. Vuggestuer i området er indlæst som røde pletter.

### Konklusion

Samlet set anvender SSI nu GIS til mange formål, både i den daglige overvågning af smittsomme sygdomme og til analyse af historiske data. Vi vil fremover især arbejde videre med webgis, fordi vi dermed kan servere georefererede data let tilgængeligt i en web browser for det sundhedsfaglige personale på SSI, som ikke har tradition for at betragte datas rumlige fordeling endsige foretage rumlige analyser. Vi håber derved at kunne integrere analyser af data med den daglige overvågning. Særligt vil vi udnytte, at vi via cpr har adgang til bopælsadresse for patienter til at

arbejde med rumlig statistik og se, om der kan findes rumlige mønstre i den måde en given sygdom spredes på over tid, om patienter bor tættere på hinanden end det kan forventes etc.

## Referencer

- MVN (2008): <http://www.medvetnet.org>. Sidst besøgt 21/01 2008
- Enter-net (2007): [http://www.hpa.org.uk/hpa/inter/enter-net\\_menu.htm](http://www.hpa.org.uk/hpa/inter/enter-net_menu.htm). Sidst besøgt 10/12 2007.
- Eurostat (2008): [http://ec.europa.eu/eurostat/ramon/nuts/introduction\\_regions\\_da.html](http://ec.europa.eu/eurostat/ramon/nuts/introduction_regions_da.html). Siden sidst besøgt 21/01 2008
- SSI (2008): <http://www.ssi.dk/sw2796.asp>. Sidst besøgt 21/01 2008
- Wilensky, U. 1999. NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.
- CBB (2007): <http://www.bioberedskab.dk/model/model.html>. Sidst besøgt 10/12 2007
- Sundhedsstyrelsen (2007): [http://www.sst.dk/publ/Publ2007/MTV/HPV/HPV\\_vaccination.pdf](http://www.sst.dk/publ/Publ2007/MTV/HPV/HPV_vaccination.pdf)
- P. Vinten-Johansen, H. Brody, N. Paneth, S. Rachman & M. Rip (2003): Cholera, Chloroform, and the Science of Medicine: A Life of John Snow. New York : Oxford University Press.
- Mitchell, Andy (2005): The ESRI guide to GIS analysis. Volume 2: Spatial measurements and statistics. Redlands: ESRI press.
- ESRI (2004): <http://gis.esri.com/library/userconf/feduc05/docs/pap222.pdf>. Sidst besøgt 24/01 2008
- Danmarks Statistik (2008): <http://www.dst.dk/HomeDK/TilSalg/doga/Standard/kvadrat.aspx>. Sidst besøgt 24/01 2008.

## Om forfatterne

Martin Rudbeck Jepsen (mjp@ssi.dk), Steen Ethelberg (set@ssi.dk), Gert Nicolaisen (gmn@ssi.dk). Alle Statens Serum Institut.