PUBLIC HEALTH GIS NEWS AND INFORMATION

April 1997 (No. 15)

Dedicated to CDC/ATSDR scientific excellence and advancement in disease control and prevention using GIS

I. Public Health GIS (and related) Events

☞ ACSM-ASPRS 1997 Annual Convention and Exhibition: Washington State Convention and Trade Center, Seattle, April 7-10; For further information, contact Denise Calvert, ACSM (301) 493-0200. [ACSM- American Congress on Surveying and Mapping, ASPRS- American Society for Photogrammetry and Remote Sensing]

☛ Forty-Sixth Annual Epidemic Intelligence Service (EIS) Conference, Centers for Disease Control and Prevention, Atlanta, April 14-18; For further information, contact Laura Coker, Epidemiology Program Office, CDC, (404) 639-3878.

☞ MapInfo Worldwide User Conference: MapWorld '97, Renaissance Orlando Resort, Orlando, Florida, April 27-May 1. Contact Sally Shapiro, MapInfo, e-mail sally shapiro@mapinfo.com or mapworld@mapinfo.com, http://www.mapinfo.com.

☛ Third International Conference on Health Effects of Low Dose Radiation: Challenges for the 21st Century, Statford-upon-Avon, United Kingdom, May 11-14; For further information, contact Rachel Coninx, BNES, One Great George Street, London SWIP 3AA, Uk or FAX: +44 (0) 171 233 1743.

☞ NCHS Cartography and GIS Guest Lecture Series, "The National Atlas of the United States of America," presented by U.S. Geological Survey staff, Hyattsville, May 19; For further information, contact C. Croner, Office of Research and Methodology, NCHS, CDC, (301) 436-7904, ext. 146 (see below: Section II. A. 1.).

☛ HazMat International '97, 15th Anniversary of the HazMat Environmental Conference and Exhibition, New Atlantic City Convention Center, Atlantic City, NJ, June 17-19; For further information, call 800-331-5706 or (218) 723-9130 or visit www.advanstar-expos.com and www.esonline.com.

☞ 1997 Annual Educational Conference, National Environmental Health Association, Arlington, VA, June 28-July 2; For further information, contact Reggie Moore, NEHA, at e-mail <NEHA.org@juno.com>.

☛ Fifth International Symposium on Spatial Databases (SSD 97), Berlin, Germany, July 15-18; For further information, contact Institut fuer Informatik, Freie Universitaet Berlin, Takustr. 9, 14195 Berlin, Germany. Fax: ++49-30-838-75-109 or e-mail <ssd97@inf.fu-berlin.de>.

☞ 1997 Joint Meeting of the Public Health Conference on Records and Statistics, and the Data Users Conference, Washington, D.C., July 28-31; For further information, contact Barbara Hetzler, NCHS, (301) 436-7122, ext.148.

II. News from GIS USERS

(Please communicate directly with colleagues on any issues)

A. General News

1. From **Jay Donnelly**, U.S. Geological Survey: THE NATIONAL ATLAS OF THE UNITED STATES OF AMERICA. In 1970 the U.S. Geological Survey (USGS) published The National Atlas of the United States of America. It was a 400-page, oversize, 12-pound collection of maps. In 1998, a new and innovative National Atlas will be published. This new

edition will include both electronic and paper map products and will exploit information management, access, and delivery technologies that did not exist in 1970.

A New National View. The National Atlas of the United States of America will provide an intuitive, map-like view into the enormous wealth of data collected by the Federal government. The new National Atlas will: deliver authoritative views of scientific, societal, and historical information; provide easy-to-use tools to display, manipulate, and query National Atlas data so that customers can produce their own relevant information; make this information more accessible to individual Americans; provide a showcase for the geospatial data collected by Federal agencies; provide links to current and real-time events and to other Federal producers of geospatial information, and; provide a national framework of well-maintained and documented base cartographic data for use by other Federal data contributors.

A Different and Improved National Atlas. How will a new National Atlas differ from the earlier edition? There will be many dissimilarities between the two, but three are noteworthy. First, though high-quality maps will continue to be a product of The National Atlas of the United States of America, the USGS will concentrate on producing compelling and useful electronic products. Second, the audience for this atlas is quite different. The USGS will cooperatively produce an atlas that becomes an essential reference for all home computer users. And finally, the new National Atlas will exploit information access and delivery technologies such as CD-ROM and the World Wide Web (WWW) that did not exist in 1970.

New Products. The National Atlas of the United States of America will include four distinct products. In addition to high-quality, small-scale maps, the atlas will include authoritative digital national geospatial and geostatistical data sets on CD-ROM. Examples of geospatial data include soils, boundaries, volcanoes, and underground water bodies. Crime patterns, population distribution, and incidence of disease are examples of geostatistical data. This information is tied to specific geographic areas and is categorized and indexed using different methods, such as county, state, and zip code boundaries or geographic coordinates like latitude and longitude. These data will be collected and integrated to a consistent set of standards to support analysis. The atlas will also include easy-to-use software for data display, query, and custom information and map making. The product will include hot links to atlas sites on the WWW. More up-to-date, real-time, and regional data will be served via the Web. Descriptive information and marketing products and software programs will also be delivered via the WWW.

New Customers. In the early 1970s, The National Atlas of the United States of America was typically found in the reference collections of libraries across the United States. Educators and government organizations were also primary customers for the original publication. At a sales price of $100, not many Americans were adding the Atlas to their home libraries. The new National Atlas is designed for individuals who own powerful home computers. This is not a customer segment that is typically addressed by USGS programs. Therefore, the USGS is seeking strategic alliances with commercial partners to gather and analyze customer information and to assess the market for the National Atlas. Efforts in 1997 are concentrated on identifying customers, determining their expectations, and using this information to sharpen product definitions. The USGS will also pursue partnerships to make products that are responsive to the needs of secondary markets like education, business, and libraries. For more information on The National Atlas of the United States of America please contact: National Atlas, U.S. Geological Survey, 508 National Center, 12201 Sunrise Valley Drive, Reston, Virginia 20192. [**Editor**: CDC/ATSDR staff GIS Users will need to arrange for offsite ENVISION to view this presentation]

2. From **Lois Dean**, HUD: HUD DEVELOPS THE NEXT GENERATION OF COMMUNITY PLANNING SOFTWARE. The U.S. Department of Housing and Urban Development (HUD) has announced that it will release a new generation of community planning software in April, 1997, that will help citizens analyze and participate in decisions that

affect the ways local governments use local, state and federal resources. The software has been developed in collaboration with Caliper Corporation of Newton, Massachusetts. Conceived by HUD's Office of Community Planning and Development (CPD) under the name Community Planning Software Plus (CPS+), the package is a special edition of Caliper's Maptitude Geographic Information System (GIS) for Windows.

The new software does everything that the first generation of software (CPS) did, but it is much easier to use for persons not proficient with computers and is compatible with any database management or word processing system on the market. HUD Secretary Andrew Cuomo has called the new software a "home run," adding that it is part of "a future where no one is left behind and everyone has access to technology which allows them to reach as high as their talents and hard work will take them." Local governments apply for funds from HUD by submitting a five year Consolidated Plan to CPD. The Consolidated Plan is a new community development planning process instituted by the Clinton Administration to improve efficiency in the grant application and review process. The Consolidated Plan combines planning, application, and reporting requirements for four major HUD formula programs into one comprehensive plan. These programs include Community Development Block Grants, HOME, Emergency Shelter Grants, and Housing Opportunities for Persons with AIDS.

Citizens, businesses, and neighborhood organizations are included in and affected by the Consolidated Planning process locally. CPS+ is also a new way for HUD to share information with public interest groups, non-profit organizations, businesses, and residents. CPS+ will allow local and state governments as well as their citizens to quickly see and understand information about their communities, and to plan HUD projects that meet their concerns. CPS+ is designed especially to provide access to users who are not familiar with the Geographic Information System. Special software features including a "MAP LIBRARY" will guide users through a wide selection of customized map displays featuring colors, patterns, and charts, while a system of interactive data screens will provide streamlined access to project descriptions, project funding and more.

Special features are also built into this edition to make it easy for users to create their own maps that display a wide range of characteristics about their community and store them in their "MAP LIBRARY." Through this powerful combination of mapping software and data bases, CPS+ will enable end-users to: Create maps that display housing conditions, economic indicators and population characteristics; Review project descriptions from HUD Grantees; Understand the impact of program initiatives on neighborhoods and communities; Propose their own projects and display affected areas on a map; and Submit Consolidated Plans and annual updates electronically to CPD. CPS+ will leverage the strengths of Maptitude by providing access to complete mapping and visualization tools; convenient packaging of supporting geographic, demographic, and HUD program data; plus effective end-user documentation.

A basic component to this is involvement of the citizens in the Consolidated Planning process, and subsequent Action Planning. We have developed CPS+ so that local organizations can have the same software and information base as the city, at a subsidized cost, and become stronger participants in the local development process through greatly enhanced access to the facts, demographics, housing conditions, etc. about their community. They can also compare their city's strategies to those of other cities through the WWW; HUD has posted the executive summaries of Consolidated Plans developed by 1000 major US cities to its HOME Page at http://www.hud.gov. Each City's Consolidated Plan also contains a number of maps, showing conditions and also the location of the proposed projects.

Inquiries concerning the uses and applications of the software may be addressed to Dick Burk at (202) 708-2096, <Richard Burk@hud.gov>.

[**Editor**: Caliper reports- Internet maps help businesses tap empowerment zone grants- Businesses looking to share in $4 billion of federal grants, loans, and tax incentives can now turn to an interactive map on the Internet to see if they are located in a qualifying area. The mapping application was created by Caliper Corporation for HUD. It allows businesses to easily determine whether their site is located within the

boundaries of one of the 105 Empowerment Zones or Enterprise Communities (EZ/EC) that are scattered across the country and are targeted for job creation and economic development over the next eight years (source: http://www.caliper. com/news.htm)]

3. From **Katherine Heck**, NCHS (through below source): PROJECT SCOPE announces PPGIS-SCOPE, an open forum for discussion of the issues and applications involved in the implementation, effective use and future direction of public participation GIS (Geographic Information Systems/Computer Mapping)...Project SCOPE Incorporated, a Boston-based nonprofit public service organization, is pleased to announce ppgis-scope, a new listserv created to provide an open discussion forum regarding all aspects of community use of Geographic Information Systems (commonly referred to as GIS or Computer Mapping).

The ppgis-scope listserv is intended to serve the needs of community-based organizations, GIS users and advocates for GIS in the community. The ppgis-scope list near-term goal is to facilitate the exchange of information and experiences which will advance the effective GIS proliferation for public use. Geographic information is essential for local organizations to participate in the improvement and advancement of their neighborhoods. HUD, a resource for GIS public participation news and information to strengthen community connections which will lead to a stronger base of public participation in the local development process, will contribute to this list. The ppgis-scope listserv will provide news and support in 5 areas: 1) Community projects using GIS/mapping; 2) What trainers are learning about mapping software; 3) Cities that package GIS data for community mappers; 4) Groups involved in addressing the hardware issues, and; 5) HUD's new CPS+ mapping software and who uses the package.

The ppgis-scope listserv needs your input on information in the following areas: 1) GIS-related conferences, public forums, training opportunities; 2) Data, online training and other GIS resources available on the Internet; 3) Cities supporting public use of GIS mapping with programs, data, training or technical assistance; 4) Students who are helping to introduce and implement GIS into communities, and 5) What works...new directions, GIS innovations, examples of leadership. The ppgis-scope listserv seeks your participation for the following: 1) Pinpointing critical GIS data needs; 2) Describing the pitfalls and challenges to GIS use; 3) Examples of partnerships which do/do not work; and, 4) Sharing your experiences and expertise in GIS applications for public services, dealing with crime, accidents, fire, brownfields, environmental justice, location, of development, social service delivery, community culture, historic preservation, bus routing, local infrastructure projects, housing services and development, urban design, repair programs, community policing and more.

Project SCOPE will archive material at http://www. projectscope.org and develop a database of resources generated as a result of the ppgis-scope discussions. Thank you for joining us, and we welcome your participation and your involving others in the list who would benefit from and contribute to the listserv process. To join ppgis-scope, send an e-mail message to "majordomo@igc.org" with the words "subscribe ppgis-scope" in the message area. Do not include the quotes, leave the subject line blank and include nothing in the message except "subscribe ppgis-scope". If you have any difficulty, please e-mail me directly for assistance. Carolyn Manjourides, President, Project SCOPE Incorporated, 25 Huntington Avenue, Suite 331, Boston, MA 02116 USA; phone (617) 424-1121or e-mail <cmanjourides@ project scope.org>.

B. Technical News

4. From **Iris Shimizu** and **Chris Moriarity** (NCHS): For your information, the December 96 issue of The Survey Statistician (published by ISI) includes an article by Jim Lepkowski and Judy Bowles entitled "Sampling Error Software for Personal Computers." They catalogue eight commercial or documented free-ware statistical software packages that are currently available for use by the general survey data analyst. Each package accounts for clustering and other complex features of stratified multistage samples. A Web site, devoted to complex survey

analysis software (description, links, etc.), was set up recently by Alan Zaslavsky of Harvard on behalf of the Survey Research Methods Section of ASA. The address is http://www.fas.harvard.edu/~stats/ survey-soft /survey-soft.html. For further information, Jim can be reached at the Institute for Social Research, 426 Thompson Street, Ann Arbor, Michigan 48103 or by e-mail at jimlep@umich.edu. [Editor: GIS can assist cost effectively in the drawing of complex sample survey designs]

5. From **Tom Richards**, PHPPO: Just a note to keep you posted on some GIS-related activities related to the Atlanta Regional Commission (ARC), the official planning agency for the 10-county Atlanta region: 1) ARC offers a three day training course each month on "Introduction to ArcView GIS 3.0. ARC is an "approved training center" by ESRI. They have spaces for 8 students each session (total of four computers; students work in pairs). The first two days covers the ArcView GIS Version 3.0 workbook. On the third day, the students are provided an introduction to ARC's CD ROM for the 10 county Atlanta region, work on a project, and prepare slides in Power Point. Cost for public/non profit agency employees is $350. Contact person is Jennifer Fredenburg (404-364-2513). 2) ARC has developed a Economic Development Information System (EDIS) CD-ROM. This includes information such as streets and highways, census tracts, and extensive demographics, which is updated on an annual basis. The cost is $490. Contact phone number 404-364-2521. EDIS web site is: http:// www.atlanta-data.com.

The next updated version will be released around May 1997. Future versions for the information on the EDIS CD-ROM will be designed to be read by Business Map and/or ArcView GIS 3.0. Business Map is an ESRI product , and is available at stores such as COMP USA (cost about $99). Apparently, if shapes files can be created in ArcView GIS 3.0, then Business Map would have the capability of reading those files. 3) ARC has "experimented" with the amount of RAM and performance of ArcView GIS 3.0. Apparently, they found optimal performance was somewhere around 24 MB RAM, and performance may decline some with higher amounts of RAM.

C. Internet News

6. **Landscape Ecology**: The twelfth annual symposium of the United States Regional Association of the International Association for Landscape Ecology (US-IALE) was hosted by the Nicholas School of the Environment at Duke University in Durham, North Carolina, March 16-19. Landscape ecology focuses on spatial heterogeneity: how to characterize pattern, how pattern originates, and its implications for populations, communities, and ecosystem processes. Increasingly, landscape ecologists are concerned with how landscapes change through time and the forces that govern these dynamics. Clearly, an understanding of the agents and trajectories of landscape change is crucial to our ability to anticipate future landscape conditions and the implications of these for resource management and land use planning [see http://fagus.env.duke. edu/iale97/].

7. **GeoComputation 97**: The 2nd International Conference on GeoComputation will be held 26-29 August 1997 at the University of Otago, Dunedin, New Zealand. Themes for the meetings include: Artificial intelligence (expert systems, neural and fuzzy modelling); Fractals and chaos; Visualization, virtual reality, and multimedia; Process based modelling; Developing distributed computing environments for geographical data processing; Quality control techniques for data and models; Exploratory data analysis and data mining; Integration of geographical tools and techniques, particularly those requiring high performance computing; Macro-modelling, micro-modelling, and scaling; Advances in geographical information systems, particularly in the area of spatial analysis; Cellular automata; and, Statistical modelling (predictive and descriptive). For more information, see http:// divcom.otago.ac.nz: 800/conferences/geocomp97/.

8. **Analysis of Spatial Data Research Network**: The European Union is funding a research network on Statistical and Computational Methods for the Analysis of Spatial Data. The network is funded under the EU 'Training and Mobility of Researchers' (TMR) initiative, and will last for almost four years, during

which time it will employ postdoctoral researchers in seven different research centres in various EU countries. The network will also sponsor a number of workshops on areas of Spatial Statistics. The statistical analysis of spatial data is complicated by the need for reasonable modelling of the spatial structure. The necessary complexity of models means that analytic solutions to statistical procedures are rarely available, and therefore it is important to develop efficient computational techniques for the analysis of these models.

The purpose of this network is to bring together leading researchers and train young researchers in the modelling, statistical analysis, and development of computationally efficient procedures for the analysis of spatial data. The project will be structured into five areas, which will overlap considerably: (1) stochastic geometry and the development of spatial models, particularly for spatial point patterns; (2) extensions of geostatistical methods to non-Gaussian spatial and spatio-temporal data; (3) statistical and computational procedures for the analysis of latent structures in data, including graphical and hierarchical models; (4) the theoretical analysis of Markov chain Monte Carlo algorithms, computational procedures which underpin the other areas of application; (5) the development and statistical analysis of spatial models for extremes.

The research team brings together experts with a wide range of areas of expertise. These range from the theoretical (for example the probabilistic properties of stochastic models for point processes, and the theoretical properties of the algorithms to be implemented), to the analysis of spatial data in industrial and environmental applications. The network will be coordinated by Gareth Roberts (Cambridge), at e-mail gareth@statslab.c; The homepage of the network, with links to the other nodes, is being established at the Web site: http://www .statslab.cam. ac.uk/~gareth/ss.html. [Source: <owner-ai-geostats@gis.psu.edu>]

9. **TIGER Field Report** (TIGER® updating at the Westchester Field Office): There is definitely a twinkle in Scott Deuel's eye. He knows something most of the rest of us don't. And what he knows (what he and his colleagues are building) will have enormous and lasting impact on desktop mapping. Scott's business card says only "Geographer", but in fact he leads the TIGER® updating effort for the states of Illinois, Indiana and Wisconsin. He and his staff work out of the U.S. Census Bureau's Field Office in Westchester, Illinois, southwest of Chicago. His is among the smaller of twelve such offices around the country. In addition to himself, the staff includes three update clerks, one lead geographic assistant, one geographic specialist, and three geographers. Their job: making the Master Address File in TIGER® as accurate and comprehensive as possible in aid of the administration of the 2000 Census. And as a nice byproduct, we in desktop mapping get national street coverage files unrivaled anywhere else in the world.

Deuel and his staff update TIGER® using Bureau-developed proprietary GUSx software running on spanking new Silicon Graphics Indigo 2 machines, at 200 MHZ with 120 MB of RAM! They are linked live to the repository for all of TIGER® which is in Charlotte on a DEC minicomputer. They TELNET to "check out" a county's TIGER® files to work on them. While checked out, the files are read only to anyone else. When edits are finished the files are "checked in". When I visited Westchester recently, one of the staff was reviewing a 4 foot by 6 foot map that had just been received from the planning department of the city of Wilmette, Illinois. It was submitted as part of TIPS (TIGER® Improvement Program), the program in which local entities, sometimes counties and sometimes cities, aid in improving TIGER® . Here's how it works: Periodically (Deuel expects this to move to a monthly tempo soon) the U.S. Postal Service sends updated Delivery Sequence Files to the Census Bureau. This is essentially a list of every deliverable address in the United States. These addresses are then geocoded against the current TIGER®. This results in an exception list, a list of valid addresses NOT matched by TIGER®. This list, organized by Census Tract, is sent to participating municipalities along with a street map with Census Tract Boundaries. The city then resolves as many addresses as possible, noting the appropriate address RANGES on the map and returning it to the Bureau Field Office. Often, other information, such as street network changes, are also

passed along by the municipality.

By way of example, the map I was shown covered about a one square mile area east of Edens Plaza. Several new address ranges were filled in; an off ramp that no longer exists was crossed off; and two tracts had "blow-ups" attached with new subdivision roads and address ranges filled in. In Scott's Illinois/Wisconsin/Indiana region there are already 1,200 municipalities participating in this effort; 135 in Cook County alone! Scott reports the municipalities are very eager to participate; he hasn't really pushed to recruit them yet. TIPS is only one of the ways the Field Office staff update TIGER® every day. In a program called Master Address File Geocoding Office Resolution (MAFGOR), field office personnel use printed maps to determine address ranges and street segments to be added to TIGER®. Soon, the Digital File Exchange program will begin gearing up. This program, already field-tested, allows municipalities with their own GIS systems to exchange digital files with the Bureau's field offices. [Source: http://www.wessex.com/TIGER/tigerfield.htm]

III. GIS Outreach

(Editor: All solutions are welcome and will appear in the next edition; please note that the use of trade names and commercial sources that may appear in *Public Health GIS News and Information* is for identification only and does not imply endorsement by CDC or ATSDR)

☎ **Use of GIS for HIV Public Health Interventions** (from **Janet L. Heitgerd**, NCHSTP/Program Evaluation Research Branch and Richard L. Melchreit, MD, AIDS Program, Conn. Dept. Of Public Health): There is increasing interest among CDC staff, State partners, and community planning groups in using GIS to help plan, implement and evaluate HIV prevention programs and activities at the local, state and national level. We are looking for information about peoples' experiences using GIS for making decisions about public health interventions aimed at preventing HIV transmission and the risk factors associated with transmission. This could include information about a single program or determining the appropriateness of an array of programs given local circumstances. Additional information on data and methodological issues would be most welcome. [Janet can be reached by e-mail at jbh0@oddhiv1.em.cdc.gov and Richard at melc100w@wonder.em.cdc.gov].

☎ **Homicides Clustering and Spatial Autocorrelation**: I'm conducting a spatial analysis of homicides and I'm looking for software that will provide the types of measures often used to describe the level of racial segregation within a region. Is anyone familiar with other uses of the measures listed below outside of the arena of racial segregation? Does your field use similar measures that I might want to apply to the analysis of violent crimes, gangs, or homicides? I use SpaceStat quite frequently but would like to move beyond Moran's I and or Geary's C to determine the seriousness of my clustering and spatial autocorrelation. Specifically, I'm interested in computing the following: index of dissimilarity - how many homicides do I have to move to have an equal distribution; distance-decay isolation index - how far apart are tracts with homicides; spatial proximity index - measure of clustering. Thank you for you time, George E. Tita, Carnegie Mellon University, The Heinz School of Public Policy, (412) 268-2186.

☎ **Cancer Clustering**: I am a Masters student at Johns Hopkins University in Environmental studies. This semester I am taking an introductory GIS course. I am working on the cancer cluster application of the system. I eventually have to gather layers and produce a poster session using GIS but I also have to write a paper on a specific application. I mainly am interested in mapping cancer clusters around environmental conditions such as a waste dump, incinerator, power plant, spill site, etc. Anything that you could share with me would be greatly appreciated. Thank you very much for your time. Kerrie Jones, phone 410 550-5439, fax 410 955-0767, 3311 Rueckert Ave, 2nd floor, Baltimore, Md 21214.

☎ **Biosafety, Genetic Change and GIS**: Hello GIS users. I'm doing some study about how GIS can be applied in medical research. So, I'm looking for some information, articles, or research that has been done on applying GIS in the bio-safety issue or other medical research. Can we utilize GIS to the medical application

that dealt with genetic change? Is there any example that this kind of application has been done? I appreciate any comment and suggestions you may have. Thank you in advance. Penroong Bamrungrach, Research Associate, Thailand Environment Institute, Environmental Information Center, Bangkok, Thailand, Tel: 66-2-331-0047 ext. 4173, Fax:66-2-3324873, e-mail: roong@tei.or.th.

☎ **CDC WONDER and TIGER?**: A contact at the Census Bureau is offering us at CDC WONDER the opportunity (or perhaps, asking us to take on the responsibility) to take over the Tiger Mapping System data and mapping engine. Forgive me for being a little shaky on the details but I believe we are talking about: (1) having digitized census tract boundaries from which we could produce maps (2) to which, with additional data of currently unknown origin, we could add streets, zip code boundaries, city boundaries (3) on which, using tract level census data, we could fill in fairly detailed demographic information. All of this would be happening via CDC WONDER on the Web.

I am far from being a GIS expert at this point, and am not entirely conversant in the various Census Bureau data sets, and even less so am I qualified to talk about what would really make a difference to people out in the front lines of public health. I know that we have been asked more than once in the past 6 months to make detailed Census data available over WONDER. So at this point I am tossing this out to all of you as a question. Would this be something worth pursuing, not as a gee-whiz space age Web application, but as something of value to public health? It will take a not-insignificant commitment of resources at our end to pursue. Any thoughts you have on this subject would be appreciated - even if they consist of "forget it - it's not worthwhile". I'd be very interested in having any and all w/ GIS interest comment. Thanks for your time. **Robb Chapman**, Acting Chief, EPO/DPHSI/PHISB (a.k.a."CDC WONDER")

[**Editor**: Robb's inquiry was sent to our GIS Users Group on March 3 and generated much response. Special thanks are extended to **JohnCrellin, John Mann, Fred Seitz, Susan Katz, Richard Spindler, Roger Friedman, Sandy Rothwell, Bill Henriques, Richard Coles, Marian MacDorman, Steve Campbell, Kevin Liske, Larry Killen, Charlie Rothwell, Monroe Sirken, Tom Richards, Roger Rochat, Andy Autry, Alfredo Vergara, Steve Guptill, Brian McCarthy, Mark Eberhardt, Jim Gaudino, Chuck Hamilton, Chris Johnson and Rita McWilliams**; Selected responses (unduplicated) are included below]

Responses

- **John Crellin**, ATSDR: I strongly encourage you to add the TIGER Data to WONDER. As an environmental epidemiologist with 7 years of experience at ATSDR and 7 years at a state health department, I have found geographic and demographic data extremely useful in evaluating health problems. Placing these data on WONDER would make them available to much broader range of health agencies and investigators.

- **Steve Guptill**, USGS: Thanks for the e-mail about the CDC WONDER and Tiger Map Service. I know a fair amount about the Tiger Map Service and its underlying data bases (and probably more importantly know the people at the Census Bureau who build the databases -- not the same folks who run Tiger Map Service). Anyway, what I suspect is that CDC would want the capabilities in Tiger to display reference map data, and add capabilities to display demographic and epidemiological information by Census tract, or other area. I do not think that CDC wants to get in the business of running the general Tiger Map Service, which is a very heavily used Web site by folks that want a map of some particular area.

- **John Mann**, ATSDR: I don't have a simple answer to your problem. I recommend that you form a workgroup of experienced volunteers from PHS agencies, as well as, any state and local agencies that are interested. If you wish to test the waters of state interest, ATSDR could make an announcement of the problem during our Partners meeting (state and tribal representatives) at the meeting scheduled during the last week of March. I foresee two benefits of Tiger datasets available thru Wonder. One, availability of Tiger datasets might encourage epidemiologists at

CDC to use some spatial analyses in their research rather than rely solely on statistical analyses. Two, PHS , state, and local health programs that routinely use current Tiger datasets could quickly update their datasets when new Census data are available.

However, I expect the cost of maintaining the system to be quite high both in computer resources and trained specialists. I believe the benefits of such a program would be increased if there was some subunit of EPO or NCHS that served as a central focal point for all types of information and datasets for health applications of GIS. The subunit would be a clearing house for multiple spatial datasets (nonconfidential) acquired by various Public Health Service, state, and local health programs in their work. Essentially, the subunit would build on the foundation of coordination and cooperation created and encouraged by Chuck Croner. Health scientists could avoid purchasing duplicative datasets and software that might be available from another agency for free. Obviously, I'm full of opinions but I'm also willingly to help with the hard work of acquiring good information for your decision. If you believe I can be helpful, please give a telephone call at 404-639-0633. Good Luck!.

- **Steve Campbell**, Maryland Dept. Of Health: 1) The Census Bureau (CB) probably has invested a lot of money in putting geocoded information up on its own web site - see http://www. census.gov/datamap/ www/ 24.html for Maryland's Profile, for example - and there are a number of web-interactive maps on their web site, which probably cost them quite a bit of money to construct and maintain, never mind the physical storage (hard disk space, etc.) that they're probably using to keep the pages active on the Web. In short - why duplicate their existing efforts, since CDC would have to invest major bucks to get a server, etc., capable of supporting this? Let the CB keep maintaining it, and have CDC point users in the CB's direction. (Or, contribute towards the server and maps' upkeep, and have CB add CDC data to its Web pages.)

2) Individual states are currently maintaining statistics and geocoded information on their own web pages, including pertinent Census data. Maryland is a good example for examples of state-coded information. Again, why duplicate the wheel? If there's pertinent/unique data, have the CB (or some state health association) add the data to the various state maps, and maybe pay a lot less money/time/effort in supporting such data efforts.

3) There's a lot of effort involved in supporting a Tiger Map System, as is evidenced by the fact that it took Census at least 2 years to get it straight - and they still don't have it quite right, as their system isn't designed to produce statistics based on user requests and the other (non-CB) systems designed to extract information from their data don't work quite as well as they might (owing to extreme memory overload and graphics intensive systems). I'm not sure that CDC has the manpower or technical expertise to maintain a Tiger Mapping System, and I would have said that Congress wouldn't give any more money to CDC to set up such a system.

4) Perhaps it would be better to collaborate with other agencies already possessing Tiger Systems (see point 1) - but extend it beyond Census to other Federal agencies which either need or have geocoded data. This might save money for the government and more effectively combine data which could then be used analytically and more effectively. I guess I'd say, summing up, that it's probably not worth CDC investing major bucks and effort in developing a Tiger Mapping support office/network, etc. It'd be a better investment to work with existing systems and agencies to put CDC data out on THEIR geocoded maps, and then put pointers on both their web pages and CDC's web pages to the correct locations. This would reduce the amount of duplicated effort and create a more powerful data system.

- **Richard Spindler**, Wisconsin Dept. Of Health: I'm from Wisconsin in the Childhood lead program, and I think target areas (states) would find it useful. CDC-Childhood Lead seems to be moving toward targeted screening. This might mean, for example, older housing, which a GIS system would clearly be useful for. I'm just beginning to use the Census data here at Wisconsin. I wonder how many states are ahead of CDC on this.

- **Roger Friedman**, NIP: I, too, am no expert with TIGER data, but my understanding is that the

commercial digital map vendors have had to do a significant amount of cleanup of TIGER data to fix both gross errors (feature at wrong coordinates) and usability problems (polygon start-end points not connected). However, it was my understanding that it already included street information. Also, what is the future of TIGER for the 2000 Census? If I recall their schedule correctly, all Census subdivisions should be defined and provided to local governments for review some time in 1998. Therefore, their new map should be ready in a preliminary way sometime this year.

- **Sandra Rothwell**, NCHS: I have two comments, one sort of pro, one sort of con. Also, I want to be sure you know that I have very little practical knowledge about the tiger files, only a theoretical idea of what they are. Comment 1: Detailed maps have historically been very useful in public health, so the idea is not one to be rejected out of hand. Comment 2: On the other hand, linking any of our data with detailed geographic information raises serious confidentiality issues - especially in an arena such as WONDER where so many people have access. CDC would have to be very careful how this application was used.

- **Bill Henriques**, ATSDR: I saw a demo of ESRI's product Map Objects last week. It operates on a web server and allows you to 'do GIS' on the Internet. It was quite impressive. The functionality of ArcView will also be coming soon on the Internet as well. New interfaces like the Spatial Data Engine (SDE) allow direct connections to relational databases on Oracle or Sybase. Thanks for allowing us the opportunity to comment on this venture. We are continually asked by our customers (the public and state/local health departments) where they can get health information for a particular location. When possible, we develop maps to depict demographic and health information. We also develops maps for sensitive subpopulations (the very young, the elderly, and women of childbearing age) around every hazardous waste site the ATSDR is working on. Efforts like the Empowerment Zones work that HHS is doing will require us to look at many different data sets for a defined area. I'd be more than happy to discuss this with you in the future.

- **Brian McCarthy**, NCCDPHP: In my opinion, the cutting edge of reproductive health epidemiology for the 21st century is going to be our ability to link GIS systems and public health data bases with individual record data, and other data bases which can be linked together. This might put us right where we need to be in having access to the data.

- **Marian MacDorman**, NCHS: I am not very familiar with this particular data set either, but anything that can be done to promote geocoding and small-area analysis within CDC I think is extremely useful.

- **Rita McWilliams**, NCEH: So far as Public Health usefulness is concerned: In early 1996, I attended a GIS meeting at Rutgers University in NJ where several county health departments reported using GIS to track plumes from waste dumps and to track Rabies. I believe the state department of health and the state DEP were working together using GIS on that. I also am aware that the department of health cancer surveillance program intends to use GIS for its work. I believe GIS has many public health applications: 1) identification of locations of services such as hospitals and trauma care centers to ensure that the population is covered and that resources are not duplicated; 2) surveillance; 3) tracking of zoonotic diseases like rabies, lyme disease, hanta virus, malaria, dengue fever, etc.; 4) tracking for epiaids when an unknown outbreak like Legionella is encountered; 5) tracking and fine-tuning prevention efforts over large geographic areas e.g., vaccinations, and; 6) tracking environmental disasters and their repercussions such as oil spills, chemical releases, contamination of water sources, etc.

- **Kevin Liske**, ATSDR: Still sounds like an investment in hardware and personnel committed to keeping things going and working with the Census Bureau to make sure that it runs smoothly and keeping it updated as they do new work.

- **Charlie Rothwell**, NCHS: I believe CDC should seriously consider taking on this project ... for the following reasons: 1) by itself Census data displayed geographically is of considerable use to public health

... 2) also CDC could play a leadership role for many states currently not involved in GIS ... to push them into this area ... 3) however many states are much more advanced than CDC in actively using GIS .. ie South Carolina, New York, North Carolina , Colorado, etc.. thus if we were to proceed ... might be best to form an outside public health advisory group to provide us direction; 4) this activity should not be just a WONDER project ... although WONDER could be useful to states and communities who will never have this capability ... we need to recognize that states will and do have access to more detailed geographical information which they should be encouraged and trained to use at their end ... and thus this encouragement, training, etc., could be sponsored by CDC.

Finally I hope the originating e-mail didn't mean that Census is dropping its responsibility for the TIGER engine ... but only sharing its use .... for health is only one issue in the use of GIS and the TIGER files and the Census is the appropriate place to build the demographic files, maintain the geocoding necessary for conducting the census, determining demographic small area shifts for block grant allocations, congressional representation and targeting federal programs.

- **Roger Rochat**, NCCDPHP: I think this would be an extraordinary opportunity.. it is essential that databases of relevant vital and health data are geocoded at levels that would permit analysis at levels below county level. Can CDC/NCHS make vital records tapes available with local geocoding e.g., zipcodes, census tract data, etc.? If birth files had even zipcodes, then e.g., national patterns of teen births, out-of-wedlock births, premature births, infant deaths, etc. could be examined and compared with other local census data. Some states and local health districts currently can do this, but many lack the tools and skills as well. Such data should be made available to public health schools, social scientists, and other professionals who work with health or geographic data.

- **Monroe Sirken**, NCHS: What other information in addition to the 1990 population Census data is included in the Tiger Mapping System. Will Census be willing to update Tiger by the year 2000 Census data? Starting with the year 2000 census, the Census Bureau plans to establish a continuing address file. How will that address file relate to the Tiger System?

- **Tom Richards**, PHPPO: My comments about having Tiger Data on Wonder System would be as follows: 1) As part of potential background information about relevance to public health, I have developed a draft abstract and Fact Sheet for a project entitled "Pilot project to develop a geographic information systems-based sampling frame for national surveys of local health departments and local boards of health" (available upon request). 2) My general impression from preliminary work on this project is that "geographic information systems" are a potential growth area for public health applications, at state and local levels as well as at the national level. So -- anything you could do in the WONDER system to facilitate this would be well worthwhile. 3) Within CDC/ATSDR in Atlanta, Bill Heinriques and co-workers at ATSDR are actually using TIGER. I would suggest that you talk directly to Bill. He could show you TIGER, and he also most likely would have some very helpful comments.

To date, 4) I myself have not yet been able to successfully access TIGER data -- though I have made an effort to do this. So far, I have found that: there are 6 TIGER-LINE CD Roms for 1995; when decompressed, they include about 42 gigabytes of data; total cost for all 6 CD Roms is about $1,500. The documentation is about 250 pages in length, and well worth reviewing as part of understanding TIGER. An important "catch-22" is actually reading the TIGER data (i.e., converting the TIGER data into a form convenient for ordinary users). If you want to do this, you most likely need a computer with specifications something as follows: somewhere between 2 to 4 gigabytes hard drive, 200 mmHz Pentium or 200 Mhz Pentium Pro (maybe), Windows 95 or Windows NT (maybe), 12x CD ROM, 128 MB RAM, 4MB WRAM Video Card, and a good color printer. Also, you will need ArcInfo or some other computer software that allows TIGER-LINE data conversion (e.g., Maptitude). ArcView GIS 3.0 itself does not provide the ability to convert TIGER-LINE data. Another

Catch-22: ArcInfo may not be that easy for the "casual intermittent GIS user."

Another potential problem, 5) is that some projects (e.g., my current project) require "national data" rather than "information limited to one county or census tract." The TIGER-LINE CD Rom files are currently structured so users need to decompress records for individual counties or small units. Of course, if I had to do this for 3000 counties, this would represent a considerable investment of time. Thus, for my project, I am starting to explore the possibility of using commercially available files (derived from TIGER-LINE data) that specifically focus on the subsets of information that I most need, and that is already compatible with my computer software program (ArcView GIS 3.0). Specifically, since ArcView GIS 3.0 comes with a county boundary file, the data I need is a "national street data" CD ROM (so that I can code local health department and local board of health addresses to "street address" level of precision") and a "minor civil division" boundary file (for boundaries of cities and townships). 6) The other alternative that I am starting to explore is what data is posted on the Internet at the state level, and whether it might be possible to download that data.

- **Chris Johnson**, NCCDPHP: As long as the data are available, I don't think it matters if it is via WONDER or a Census homepage. For modem-oriented WONDER, I imagine access would be impossible because of the size of the datasets to be downloaded.

- **Susan Katz**, NIST:Currently I am employed at NIST, another DOC agency (like Census) which also does some GIS work (mostly synchronization of GIS standards). There are MANY commercial applications of GIS in addition to many useful PH applications, so I don't think CDC should be responsible for the whole burden. Perhaps an interagency arrangement would be appropriate (although more complex to manage, I think it would yield the best coordination and combination of technical expertise). I am currently working under an interagency agreement with CDC/PHPPO. These can easily be extended. I would be glad to help connect you with DOC people if you would like to investigate this route.

-**Alfredo Vergara**, NCEH: I'm an epidemiologist with the Radiation Studies Branch at NCEH. I have used a program called Landview II which was put together by the Census Bureau and EPA, as part of President Clinton's initiative to make data available to communities (Environmental Justice). This program, which you can download from the Census home page and also can purchase in CD's, combines Tiger files for all of the US; several EPA databases on reportable chemical handling, manufacture, releases, and superfund sites; and census data for state, county, congressional district, census track and block. The CDC also has contributed mortality by county (Gib Parrish, NCEH) and the National Oceanic and Atmospheric Agency contributed wind data. All data is in dBase format. It is a very useful descriptive tool, even though as a GIS it is not sophisticated. You can do some thematic mapping, but you cannot combine variables easily and you can not do statistical analysis. It is however very user friendly.

Our branch has been working with a citizen group in Ohio to encourage the community to use this program for environmental health monitoring (home grown). Anyway, their proposal sounds a bit like this, and I was wondering if you knew about Landview. I have contact persons and e-mails for the people at Census and EPA who work with Landview. At the end of last year, one of them mentioned that they would be releasing a windows version of the program that was even easier to use and had updated data files, so it sounded to me like it was not a dead project. If users would like more information I would be happy to talk to them (770 488-7621) and provide contacts.

- **Mark Eberhardt**, NCHS: I saw this notice and wanted to add one piece of information. It seems like having this system would be important. I know that some of the data systems at NCHS are considering collecting global positioning information on survey respondents. Issues such as confidentiality and cost will impact on the collection of this information. If the data are collected, it seems that the TIGER Mapping System would be a valuable tool to use this information, especially if other systems are not readily available.

****

[Editor- **POSTNOTE** from Robb: I called Fred Broome (Geography Division, Census Bureau) and got some good perspective on the whole question. I was already leaning this way but as a result of this conversation have pretty much decided at this point not to pick up the mapping engine, because (1) there are a number of commercial Web Server add-ins we could use (2) I lean away from getting stuck with maintaining home-made code when there are good alternatives (3) we are exploring using JAVA applets to do the actual mapping, which we think may end up being far faster and more efficient than the Tiger Mapping System approach (doing all of the mapping at the Web Server). However, I am still keenly interested in providing "one stop" access of the Census Bureau mapping and demographic data for CDC, state and local HDs, and other PH users. I believe it would be valuable to either maintain a copy of this data or to maintain reasonably fast access to this data on behalf of the PH community. Deciding how best to do this may take some doing. Census is working on a mechanism to provide access to this data over the Internet; simultaneously, we are in the early stages of designing a mechanism by which dissimilar systems in different agencies could exchange data with each other. Such an approach doesn't exist yet but in the long run could save us the considerable effort of maintaining our own copy of the data here. Fred agreed to share their dissemination system design plans with us, which ideally could lead to compatible systems. Thanks again (GIS Users) for your very useful involvement. **Robb Chapman**]

IV. Special Reports

(Submissions are open to all)

❖ **US EPA's Aerometric Information Retrieval System (AIRS)**: AIRS is a computer-based repository of information about airborne pollution in the United States and various World Health Organization (WHO) member countries. The system is administered by the U.S. Environmental Protection Agency (EPA), Office of Air Quality Planning and Standards (OAQPS), Information Transfer and Program Integration Division (ITPID), located in Research Triangle Park, North Carolina. AIRS is installed on the IBM computer system at the EPA's National Computer Center (NCC) in Research Triangle Park, North Carolina. Any organization or individual with access to the EPA computer system may use AIRS to retrieve air pollution data.

The **Air Quality Subsystem (AQS)** contains measurements of air pollutants and meteorological data from about 10,000 monitoring stations operated by EPA, state and local agencies, and WHO member countries. EPA uses this data to assess the overall status of the nation's air quality and to prepare reports to Congress as mandated by the Clean Air Act. EPA also uses the data to identify areas where improvements in air quality are needed. Locations where air pollution exceeds federal standards are called Non-Attainment Areas.

The **AIRS Facility Subsystem (AFS)** contains emissions, compliance data, and permit data for stationary sources regulated by the U.S. EPA and state and local air pollution agencies. This information is used by states in preparation of State Implementation Plans (SIPs), to track the compliance status of point sources with various regulatory programs, and report emissions estimates for pollutants regulated under the Clean Air Act. States are required by the Code of Federal Regulations (CFR) to report to EPA annual emissions estimates for point sources emitting greater than or equal to 100 tons per year of volatile organic compounds (VOCs), nitrogen dioxide (NO2), sulfur dioxide (SO2), particulate matter less than 10 microns in size (PM-10); 1000 tons per year of carbon monoxide (CO); or 5 tons per year of lead (Pb). States are also required by the Clean Air Act Amendments to report emissions data for point sources in areas where air pollution exceeds federal standards (nonattainment areas). AFS contains data for nearly 150,000 air pollution point sources monitored by the U.S. EPA and/or state and local air regulatory agencies.

The **GeoCommon Subsystem (GCS)** is a repository of reference data and control information shared by the AQS and AFS subsystems. The data is comprised of codes and code descriptions used to identify places, pollutants, and processes, geographic information, and values such as air quality standards and emission factors. Reference data consists mostly of the codes and text descriptions used in AIRS to

identify places or things, e.g., FIPS state/county/city codes, pollutants and meteorological measurements, collection and analysis methods for air samples, sampling intervals, units of measurement, source category codes, standard industrial classification codes. Geo-Common also contains valid combinations of items, such as pollutant, sampling interval, method of collection and analysis, and units. Numerical values, such as the minimum detectable value of pollutant concentration using particular collection and analysis methods, are also present. EPA maintains this reference information.

The Clean Air Act requires EPA to set National Ambient Air Quality Standards for pollutants considered harmful to public health and the environment. The Clean Air Act established two types of national air quality standards. Primary standards set limits to protect public health, including the health of "sensitive" populations such as asthmatics, children, and the elderly. Secondary standards set limits to protect public welfare, including protection against decreased visibility, damage to animals, crops, vegetation, and buildings. EPA has concluded that the current primary standard is not adequate to protect the public from adverse health effects.

Under the Clean Air Act, EPA is required every five years to review and revise, as necessary, national ambient air quality standards (NAAQS) for six air pollutants -- ground-level ozone, carbon monoxide, particulate matter, sulfur dioxide, nitrogen dioxide, and lead. EPA is in the midst of reviewing, and determining whether to revise, the national ambient air quality standards for ground-level ozone and particulate matter (PM-10).

*****

Ozone: Exposure to ambient ozone concentrations has been linked to increased hospital admissions for respiratory causes, such as asthma. Studies conducted in the Northeastern United States and Canada show that ozone air pollution is associated with 10-20 percent of all of the summertime respiratory-related hospital admissions. Repeated exposure to ozone can make people more susceptible to respiratory infection and lung inflammation, and can aggravate preexisting respiratory diseases, such as asthma.

Children are most at risk from exposure to ozone because they are active outside, playing and exercising, during the summertime when ozone levels are at their highest. For example, summer camp studies in the eastern U.S. and southeastern Canada have reported significant reductions in lung function in children active outdoors. Adults who are outdoors and moderately active during the summer months, such as construction workers and other outdoor workers, are also among those most at risk. These individuals, as well as those with respiratory illnesses, such as asthma, can experience a reduction in lung function and increased respiratory symptoms, such as chest pain and cough, when exposed to relatively low ozone levels during periods of moderate exertion. Long-term exposures to ozone can cause repeated inflammation of the lung, impairment of lung defense mechanisms, and irreversible changes in lung structure, which could lead to chronic respiratory illnesses such as emphysema, chronic bronchitis, and/or premature aging of the lungs.

*****

Particulate Matter: EPA has similarly proposed revisions to the primary and secondary national ambient air quality standards (NAAQS) for particulate matter (PM). EPA believes these changes are necessary to protect public health and the environment.

The characteristics, sources, and potential health effects of larger or "coarse" particles (from 2.5 to 10 micrometers in diameter) and smaller or "fine" particles (smaller than 2.5 micrometers in diameter) are very different. Coarse particles come from sources such as windblown dust from the desert or agricultural fields and dust kicked up on unpaved roads from vehicle traffic. Fine particles are generally emitted from activities such as industrial and residential combustion and from vehicle exhaust. Fine particles are also formed in the atmosphere from gases such as sulfur dioxide, nitrogen oxides, and volatile organic compounds that are emitted from combustion activities and then become particles as a result of chemical transformations in the air.

Coarse particles can deposit in the respiratory system and contribute to health effects such as aggravation of asthma. EPA's "staff paper" concludes that fine particles, which also deposit deeply in the

lungs, are more likely than coarse particles to contribute to the health effects (e.g., premature mortality and hospital admissions) found in a number of recently published community epidemiological studies. These recent community studies find that adverse public health effects are associated with exposure to particles at levels well below the current PM standards for both short-term (from less than 1 day to up to 5 days) and long-term (from generally a year to several years) periods. These health effects include premature death and increased hospital admissions and emergency room visits (primarily among the elderly and individuals with cardiopulmonary disease); increased respiratory symptoms and disease (among children and individuals with cardiopulmonary disease such as asthma); decreased lung function (particularly in children and individuals with asthma); and alterations in lung tissue and structure and in respiratory tract defense mechanisms.

EPA believes that the current standards do not adequately protect the public from the adverse health effects of particles and need to be revised. In addition, EPA also believes that there are welfare effects from particles for which the current PM10 secondary standards do not provide adequate protection. Chief among those is visibility impairment. Particles primarily in the fine range are responsible for visibility impairment because of their ability to scatter and absorb light effectively.

EPA's "staff paper" concludes that fine particles are a better surrogate for those components of PM most likely linked to mortality and morbidity effects at levels below the current standards, while course fraction particles are linked to effects such as aggravation of asthma at higher concentrations. The Clean Air Scientific Advisory Committee made a near unanimous (19 of 21 members) recommendation that new standards for PM2.5 be added while retaining PM10 standards at an indicator for course fraction particles.

*****

Action Plan: On November 27, 1996, EPA proposed to revise the national ambient air quality standards for ozone and PM. EPA is under a court order to issue a final decision on whether to retain or revise the PM standards by June 28, 1997. The Agency intends to follow the same schedule for issuing the final ozone standard, and also intends to propose a regional haze program in mid-1997.

Under the Clean Air Act, once EPA revises an ambient air quality standard, the Agency must establish deadlines for attaining the standard and also set forth the actions that cities and states must take to meet the new standards by those deadlines. In the case of ozone, PM and regional haze, EPA will develop a two-phased implementation strategy. Phase I of the strategy will focus mainly on matters related to designation of areas with regard to their attainment status for new or revised standards. Phase II will address remaining implementation plan requirements. EPA anticipates proposing Phase I for public comment in June 1997 and taking final action on Phase I in June 1998. Phase II is scheduled to be proposed in June 1998 and finalized in June 1999.

The proposed rules were published in the Federal Register on December 13, 1996. Proposed Rule and Federal Register Cite: Ozone NAAQS, 61 FR 65716; Particulate Matter NAAQS 61 FR 65638; Particulate Matter Monitoring Requirements, 61 FR 65779; Interim Implementation Policy, 61 FR 65752; Advance Notice of Proposed Rulemaking on Implementation, 61 FR 65764. [Anyone with a computer and a modem can download the new proposals and fact sheet from the Clean Air Act Amendments bulletin board of EPA's electronic Technology Transfer Network (TTN) by calling (919) 541-5742 (look under "Recently Signed Rules"). For further information about how to access the board, call (919) 541-5384. The TTN can also be accessed through EPA's homepage on the Internet, http:// ttnwww.rtpnc.epa.gov]

V. Public Health GIS Literature

(This section may include literature citations, abstracts, syntheses, etc., and submissions are open to all)

Selected Abstracts: 1997 Symposium on Statistical Methods, Statistical Bases for Public Health Decision Making: From Exploration to Modeling, CDC, January 1997

❑ "GIS Displays for Multivariate Statistical Models of Toxic Response," Hertzberg, R. and White, C.:

Graphical displays of multivariate data often are designed to portray one or more response variables along with several covariates. Graphics for multivariate models, such as available with geographic information systems (GIS), use many visual cues to show changes in several response variables as functions of two spatial coordinates. We have adapted GIS displays to show multivariate toxicity models along with goodness-of-fit indicators. The primary covariates are dose and exposure duration, and the response variables include toxic severity, animal test species and model goodness-of-fit. Advantages of such displays are the impression of trends in both the predicted response as well as model fit across the covariate domains. Disadvantages of using canned GIS software include the automatic smoothing that obscures unevenly spaced covariate values. [e-mail: hertzberg.rick@epamail.epa.gov]

❒ "The GIS Based Spatial Network Models of Schistosomiasis Infection," (poster) Zhou, Y., Maszle, D., Gong, P. and Spear, R.: It is known from experience of field epidemiologists that spatial factors play an important role in schistosomiasis transmission. This study uses statistical methods to identify the spatial risk factors and models the spatial process of disease transmission. A Geographic Information System (GIS) is applied to assist the spatial analysis and modeling of schistosomiasis transmission in a mountainous region of Sichuan Province in southwestern China. Spatial network autocorrelations, spatial regressive and autoregressive models are used to study the spatial aspect of schistosomiasis transmission. The results show statistically the significant spatial autocorrelations of risk factors and the simple spatial regressive and autoregressive models predict the risk of schistosomiasis infection well. The GIS based modeling approach is shown to be useful in understanding schistosomiasis transmission process and promising for the development of public health intervention strategies and control programs. [e-mail: zhouyi@rfs.63.berkeley.edu]

Other Symposium Papers with Geographic and Statistical Mapping Themes: (1) "Smoothing and Simultaneous Adjustment of Cancer Rates: Melanoma Rates Among Whites in U.S. Counties," Kafadar, K. [e-mail: kk@math.cudenver.edu];(2) "Spatial Artifacts in Current Methods of Mapping Parameter Estimates," Gelman, A. and Price, P.N. [ph: (212) 854-3652]; (3) "Exploring Spatial Patterns of Mortality Data: The New NCHS Atlas," Pickle, L.W., Mungiole, M. and White, A.A. [e-mail: lwp0@nch09a.em.cdc.gov]; (4) "Predicting Pollutant Concentrations Using Spatial Modeling with Covariates and Monitoring Data," Price, P.N., Nero, A.V. and Gelman, A. [e-mail: pn price@lbl.gov]; (5) "Analyzing Spatial Patterns in Health Outcome Surveillance Data," Merrill, D.W. and Selvin,S. [e-mail: dwmerrill@lbl.gov];(6) "Application of a Weighted Headbanging Algorithm to Mortality Data Maps," Mungiole, M., Pickle, L.W., Simonson, K.H.and White,A.A.[e-mail:mim4@nch09a.em.cdc.gov];(7)"Small Area Estimates of Overweight Individuals in the U.S.," (poster) Davis, W. and Malec, D. [e-mail: wbd1@nch09a.em. cdc.gov], and ; (8) "A Surface-Model Approach to the Exploration of Spatial Patterns of Fertility," (poster) Ruiz, M.O. [e-mail: mruiz@coss.fsu.edu].

<u>Selected Abstracts: 92nd Annual Meeting of the Association of American Geographers, April 1996</u>

❒ "Exploratory Spatial Data Analysis with GIS: Operational Issues," Anselin, L., Bao, S. and Smirnov, O.: This paper discusses operational issues encountered when implementing state of the art techniques for exploratory spatial data analysis (ESDA) in commercial GIS software. The specific focus is on recently developed ESDA methods for measuring local spatial association that are available in the SpaceStat software for spatial data analysis, such as LISA statistics and the Moran scatterplot. Different degrees of close and loose coupling of an ESDA module with the ArcView GIS software are considered and evaluated in terms of design complexity, implementation effort, computational efficiency and user friendliness. Keyword: spatial statistics, geographic information systems, spatial autocorrelation; Luc Anselin's (W.VA U.) e-mail <luc @lambik2.rri.wvu.edu>.

❒"Field Screening of Volatile Organic Compounds in Landfill Gas Emissions," Emerson, C.W.: Potentially

hazardous concentrations of toxic volatile organic compounds (VOCs) have been measured in the gas produced by the degradation of buried waste in landfills. The occurrence of these compounds is often unevenly distributed in the filled area of a solid waste disposal site. The high cost of detailed laboratory analyses of volatile organic compounds in landfill gas limits the number of samples that can be taken as part of a site audit. When mapping the spatial distribution of toxic compound concentrations at a given landfill, the small number of available point concentrations must therefore represent large areas of the site. The high level of uncertainty in this map often leads to an unacceptable level of uncertainty in the estimated risk to surrounding populations caused by emissions of toxic compounds from a particular solid waste disposal site. Inexpensive field screening techniques can extend the spatial coverage of a sampling effort at minimal costs and can reduce the variance of toxic emissions estimates by limiting detailed, but expensive laboratory analyses to those samples that have a high probability of containing measurable amounts of a group of targeted VOCs. An evaluation of a field screening strategy performed on a synthetic dataset derived from measurements obtained in California's statewide audit of VOCs in landfill gas shows that a large number of field screening samples provides a more precise estimate of the mass loading of VOCs in landfill gas emissions than does a small number of laboratory analyses. Keyword: landfills, risk assessment, screening; Charles Emerson's (U. of Iowa) e-mail <charles-emerson@uiowa.edu>.

❑ "Spatial Clustering of Breast Cancer on Dead-end Streets: A GIS Analysis," Gilman B. and McLafferty, S.: The high rate of breast cancer in Long Island, NY has received a great deal of publicity in the past 5 years. Researchers suspect that the higher-than-normal rates of breast cancer on Long Island result from a combination of demographic, socio-economic, and environmental factors. However, the roles and contributions of these factors are poorly understood. A previous study found no evidence of spatial clustering of breast cancer in West Islip, NY, a Long Island community with a known high prevalence rate (Timander 1995). Despite this, community residents feel there may be clustering along dead-end streets, at the ends of water mains, where pesticides and other contaminants may accumulate. This research presents an exploratory analysis of spatial clustering of cases in West Islip at the ends of these streets. Using a geographic information system we modeled the street network to determine residential location in relation to dead-end streets. Data from a community survey of breast cancer in local residents were mapped in reference to this street network in order to assess spatial clustering. We used statistical methods to incorporate known risk factors such as family history and confounders such as length of residence. This research illustrates the use of geographic information systems in community-based analysis of urban environmental health problems. Keyword: breast cancer, geographic information systems; Brett Gilman's (Hunter College) e-mail <bgilman@everest.hunter.cuny.edu>.

❑ "Isarithmic Mapping of Childhood Lead-Soil Exposure Hot Spots," Padgett, D.A.: The banning of lead-based paint during the late 1970s has led to abatement measures aimed at reducing children's exposure risk. However, recent research indicates that childhood lead poisoning continues to be a serious problem, especially for minority children living in urban areas. Lead contaminated soil in children's play areas is now considered a primary exposure source. Current methods for identifying exposure points primarily utilize choropleth maps with hot spots identified as census tracts or block groups exhibiting socio-economic and physical characteristics indicative of high risk. The predictive ability of choropleth maps is somewhat limited by constraints of time and place. This study illustrates the development of a GIS-based isarithmic mapping technique for predicting childhood lead-soil exposure hot spots. Hot spots are defined using a combination of soil sampling, groundtruthing, and expert systems derived data. Keyword: hazardous waste, geographic information systems, risk assessment; David Padgett's (Austin Peay State U.) e-mail <padgettd@lynx.apsu.edu>.

❑ "Measuring Distance to Care Using GIS Across the Rural-Urban Continuum," Savitz, L.A., Hanchette,

C.L., Ethridge, W.D., Ricketts, T.C. and Gesler, W.M.: The U.S. Office of Rural Health Research funded a study to assess the influence of maternal employment location in obstetrical health care seeking behavior in North Carolina. Primary research questions were: (1) To what extent is obstetrical care bypassing behavior influenced by maternal employment; and (2) To what extent is adequacy of obstetrical care influenced by maternal employment. Two hundred and ten women were interviewed postpartum for this retrospective, cross-sectional study. Calculation of distance to care from various points of origin involved several steps-selection of several distance measure(s), database development, address matching, and distance calculation. Computed distances were used to statistically test hypotheses using crude and adjusted odds ratios generated by logistic regression. It was found that employed women were 2.2 times more likely to bypass obstetrical providers located closest to their homes than were unemployed women overall with education acting as a strong effect modifier. No significant differences in adequacy of care were detected between employed and unemployed women. Results from this study may contribute to policy decisions related to health care delivery for employers, insurers, and government officials as well as provide indications for future data collection efforts in similar studies using GIS to calculate distance to care in rural and urban areas. Keyword: geographic information systems, distance to care, bypassing; Lucy Savitz (Cecil G. Sheps Center for Health Services Research, UNC Chapel Hill, NC 27599-7590).

❑ "Regionalizing Communicable Diseases in North Carolina," Wilson, J.L.: Two hypotheses with respect to county-level communicable disease patterns in North Carolina are examined. The first hypothesis is that groups of communicable diseases will exhibit meaningful spatial patterns. For example, sexually transmitted diseases will tend to have similar spatial distributions. The second hypothesis asserts that there will be spatial correspondence between the mapped patterns of communicable diseases and certain socio-economic variables like educational attainment and income levels. Morbidity and population data are derived from a combination of state and federal (US Census) sources. The group membership for communicable diseases is determined using principal components analysis and clustering techniques found in *SPSS-Windows*. Counties categorized by component scores or cluster membership are mapped using *Atlas GIS for Windows*. Spatial clusters or regions are delineated using the software *CLUSTER* and the B-W statistic for spatial autocorrelation. Spatial clusters for socio-economic variables are determined in the same manner as the communicable disease regions. Difference mapping between these morbidity regions and socio-economic variable layers is carried out using Atlas GIS and their spatial correspondence is assessed by examining the residual map. The results show that there are distinct clusters of associated communicable disease groups and that there is some correspondence to the map pattern of socio-economic variables. This type of cartographic analysis can be useful for targeting public health measures at a state regional level. Keyword: clustering techniques, difference mapping, communicable diseases; Jim Wilson's (East Carolina U.) e-mail <gewilson@ecuvax.cis.ecu.edu>.

❑ "Environmental Equity Models for Siting Facilities in Densely Populated Areas," Falit-Baiamonte, A. and Osleeb, J.P.: Neighborhoods within a city vary with respect to obnoxious facilities that are located within its boundaries. The impact on the neighborhood results from a combination of the size and number of facilities, the types of air and water emissions that are being released as well as the traffic generated. To assess the vulnerability of a specific population to environmental hazards, the BAEL (Baseline Aggregate Environmental Loads) GIS profile for the Greenpoint/Williamsburg community in Brooklyn has been developed. This community suffers from mixed land use zoning that permits hazardous material generators to be located adjacent to residences. Within this community can be found a waste water treatment facility, an incinerator, a low level radio active waste repository, approximately 20 EPA Toxic Release Inventory sites and over 200 hazardous material processors. Most mathematical models which have been developed to assist planners in locating noxious facilities either maximize the distance between the facilities or minimize the total population impacted by

these facilities. This paper describes a model which explicitly addresses the issue of spatial equity by measuring a census block group's exposure to environmental hazards. The spatial distribution of this measure provides a basis upon which to assess the equity of the system. The model can be used to determine the potential locations of noxious facilities by simultaneously minimizing the vulnerable population effected by the facilities and maximizing the overall equity of the system. Keyword: mathematical modeling, environmental equity, GIS; Anthony Falit-Baiamonte's (Hunter College) e-mail <baiamont@everest.hunter.cuny.edu>.

❑ "Evaluation of the Youth Protection Ordinance in Charlotte, NC," Groff, E.R.: Charlotte is one of several major cities to institute a Youth protection Ordinance. At its inception, the Youth Protection Ordinance enjoyed broad community and police support because of its emphasis on preventing juvenile crime and victimization. To achieve these goals the Youth Protection Ordinance relies on a curfew for youth under 16. However, there has been little research on the effectiveness of curfews at reducing juvenile crime and victimization. The purpose of this research is to quantify the City of Charlotte's attempt to reduce the victimization of juveniles and the incidence of juvenile crime through the application of a curfew within the city limits. Several facets of the curfew's impact are studied. First, a profile is developed of curfew violators. Next, through the use of a Geographic Information System (GIS), geographic concentrations of violations and violators are identified. These areas can then be targeted for both prevention and enforcement activities. Finally, in an attempt to answer the question of curfew effectiveness, juvenile arrest and victimization rates in these user defined areas are examined. In addition, the offense rates of crimes frequently associated with juvenile offenders are compared with the same period last year. Results show a definite clustering in the locations where curfew violators are picked up and among the home addresses of violators. Distance from home address to location of violation varies by race and sex with white females having the largest average distance and black females the shortest. Keyword: crime geography, curfew, GIS; Elizabeth Groff, Charlotte-Mecklenburg Police Department, Charlotte, NC 28202.

❑ "The Disease Ecology of Childhood Lead Poisoning in North Carolina: Predicted and Actual Patterns of Risk," Hanchette, C.L.: During the past two years, a geographic information system (GIS) has been used to model and map childhood lead poisoning risks in North Carolina counties and communities through the use of a cartographic modeling technique known as suitability analysis. The model has utilized census data to identify potential lead poisoning "hot spots" at the county and sub-county levels. Lead poisoning prevalence data from nearly 150,000 children screened through North Carolina's Childhood Lead Poisoning Prevention Program has indicated that although high risk areas for lead poisoning generally coincide with those predicted by the model, there are some interesting differences. After a brief presentation of the modeling results, this paper will describe the disease ecology of lead poisoning in North Carolina with special reference to demographic, historical and cultural factors that contribute to the presence or absence of the disease. Carol Hanchette's (N.C. State Center for Health and Environmental Statistics) e-mail <chanchet@gis.sches.ehnr.state.nc.us>.

❑ "Mapping Native American Health Statistics in Lancaster County, Nebraska," Kroll, K.C. and Echavarria, F.R.: The purpose of this study is to map spatial patterns of various health statistics for Native Americans in Lancaster County, including Lincoln, Nebraska. Health care data of Native American clients of the Nebraska Urban Indian Health Coalition (NUIHC) were combined with 1990 U.S. Census data and georeferenced to TIGER files using a GIS. The maps will help the NUIHC evaluate the effectiveness of its current outreach, as well as design effective health care programs in the future. The results from this research have implications for the delivery of health services to other under-represented groups in and beyond Nebraska. Keyword: GIS, health, Native Americans; Kurtis Kroll's (U. of Nebraska, Lincoln) e-mail <kurtisk@andes.unl.edu>.

❑ "Mapping Violence Through Time: A Local

Example," LeBeau, J.: Various observations and generalizations regarding the spatial patterns of violence have relied on databases pertaining to a single time period, usually a year. The advent of GIS and the availability of police calls for service databases allows one to examine the spatial distribution of different forms of violence through multiple time periods. This paper discusses and illustrates the spatial distribution of violence across 537 areal units in Charlotte, North Carolina, from 1984-1993. Generalizations are formed regarding the expansion and contraction of violence with urban growth, transportation planning, drug markets, and police intervention. Keyword: violence, GIS, North Carolina; Jim LeBeau's (Southern Illinois U. at Carbondale) e-mail <lebeau@siucvmb.siu.edu>.

❑ "Applying GIS to Environmental Equity Analysis in Louisiana," Liu, L.: The issue of environmental equity has increasingly received attention in Louisiana. The state of Louisiana has the third largest percentage of minority population in the entire country, behind District of Columbia and Mississippi. A large number of facilities in Louisiana release toxic materials to air, water and soil, causing pollution and affecting human health. In fact, the Toxic Release Inventory (TRI) database of the United States Environmental Protection Agency (EPA) has listed more than four hundred such facilities in Louisiana. Many communities are affected by these TRI facilities. This research aims to raise questions related to environmental equity in two aspects. One is to investigate whether minority populations and poor communities are disproportionately exposed to potential toxic chemicals. The other is to examine whether there is a spatial correlation between the location of the thirteen Superfund sites in Louisiana and the distribution of minority populations. GIS is used as a tool for spatial analysis in this research. Topographical overlay, proximity analysis, and summary statistics are performed using ArcView, while regression analysis is done utilizing Systat. Preliminary studies have revealed the following results: 1) there is a significant negative correlation between the percentage of minority population and the distance to TRI facilities; and 2) there is a positive correlation between the percentage of minority population and the distance to Superfund sites in some communities. Keyword: environmental equity, GIS, Louisiana; Lin Liu's (U. Of New Orleans) e-mail <lliu@cs.uno.edu>.

❑ "When are Disease Clusters for Real?," Rushton, G.: Disease data for small geographic areas- such as cases of infant mortality in Des Moines, Iowa that I examine here- are notoriously difficult to evaluate for the presence of significantly high rates. Yet new geographic surveillance techniques, made possible by advances in GIS and methods of spatial analysis, are likely to identify many more cases of putative disease clusters in the future than in the past. I make multiple tests of statistical significance for high rates of infant mortality in Des Moines, Iowa, based on reference distributions computed by Monte Carlo simulation methods. I conclude that some key characteristics of the spatial pattern of areas of "significantly" high rates distinguish "real" clusters from false clusters. Keyword: disease clusters, GIS, spatial analysis; Gerry Rushton's (U. of Iowa) e-mail <gerard-rushton@uiowa.edu>.

❑ "The Integration of Multilevel Modeling Techniques and GIS to Examine Access to Care," Wittie, P.S.: Geographers and public health researchers continue to look at ways to model explanations of disease and health-related behavior, especially when public health applications have been found to challenge conventional methods. The admonition of Reuben C. Warren, Associate Director for Minority Health at the Centers for Disease Control and Prevention, serves as a reminder for public health investigators to coalesce traditional venues of public health concern, such as influences of income, education, and individual behavioral influences, with the longtime focus of medical geographers -geographic location- to discern neighborhood influences on individual health outcomes. This blending of the spatial analytic techniques and tools of medical geographers with the individual approach of epidemiologists may facilitate the development of techniques to examine place-based health concerns. This paper discusses a relatively unused methodological genre, multilevel modeling, where multi-leveled health and socioeconomic data are

statistically and spatially analyzed to examine the socio-structural or various place-based pressures on individuals. In contrast to conventional methods of analysis, multi-level modeling facilitates an analysis of both neighborhood and individual-level effects to determine system-related problems with access to care. The technique has relevance for developing health model simulations and for understanding the role of place in health and health behavior, especially when used in conjunction with a GIS where modeling may identify other similarly characterized areas which deserve further study. Keyword: health, multilevel modeling, GIS; Peggie Wittie's (UNC, Chapel Hill) e-mail <uncpsw@geog.unc.edu>.

❒ "Implementing Spatial Measures of Segregation in GIS," Wong, D.W. and Chong, W.: Previous research has indicated that many traditional measures of segregation fail to account for spatial distribution of minority population. As a result, distribution of a minority in a checker board pattern and a two-region pattern (each region dominates by one group) will yield the same level of segregation. Interactions between different ethnic groups across the enumeration boundary are not taken into account to measure segregation. Spatial measures have been proposed to tackle the deficiencies in traditional measures, but they are rather difficult to implement because they require spatial information which is quite difficult to obtain. Spatial components include adjacency information and the length of common boundary are required. But these components can be extracted from most, if not all, GIS packages. Even if the spatial information is available, the computation of many of these indices is more complicated than the traditional measures such that ordinary spreadsheets can no longer handle the situation. Unfortunately, GIS are not quite capable in supporting this type of computation. We have to rely on self-written programs or powerful mathematical modeling packages to accomplish the job. In this paper, we discuss how spatial information pertinent to segregation calculations can be extracted from Arc/Info and how these components can be combined with attribute data stored in GIS to derive various spatial measures of segregation. The latter task is accomplished with the help of S-Plus, a very powerful statistical and mathematical modeling package, and GISLink, which is a data bridge transferring data between S-Plus and Arc/Info. Keyword: segregation measures, GIS, mathematical-statistical modeling; David Wong's (George Mason U.) e-mail <dwong@gmu.edu>.

****

Journal Article: Perlin, S.A., Setzer, R.W., Creason, J. And Sexton, K. (1995). "Distribution of Industrial Air Emissions by Income and Race in the United States: An Approach Using the Toxic Release Inventory," *Environmental Science & Technology*, Vol.28, No.1, pp. 69-80. Abstract: There currently is a scarcity of scientific information to guide public policy decisions about issues of "environmental justice"; broadly defined as the goal of achieving adequate protection from the harmful effects of environmental agents for everyone, regardless of age, culture, ethnicity, gender, race or socioeconomic status. This paper highlights several key methodological issues that need to be addressed as part of ongoing efforts to strengthen the scientific foundation for informed decision-making regarding environmental justice. Specifically, careful thought must be given to the selection of appropriate (1) statistical tests, (2) geographic unit(s) of analysis, (3) exposure estimators, and (4) comparison (reference) populations. These methodological issues are examined in the context of a nationwide study looking at the differences by ethnicity/race and household income in county-level air emissions of industrial chemicals. National and regional comparisons are made for 1990 using emission estimates from the Toxic Release Inventory, demographic data from the Census, and income data from the Donnelley Marketing Information Services.

New Book: *Visual Explanations*, by Edward R. Tufte. Graphics Press, 1996. P.O. Box 430, Cheshire, CT 06410, $45 Postpaid. This is the third in a series of books written and published by Tufte on the design of the display of information. In the introduction to this book Tufte tells us how these three books are related. "The Visual Display of Quantitative Information" is about pictures of numbers, how to depict data and

enforce statistical honesty. "Envisioning Information" is about pictures of nouns (maps and aerial photographs for example, consist of a great many nouns lying on the ground). Envisioning also deals with visual strategies for design: color, layering, and interaction effects. "Visual Explanations" is about pictures of verbs, the representation of mechanism and motion, of process and dynamics, of causes and effects, of explanation and narrative. Since such displays are often used to reach conclusions and make decisions, there is a special concern with the integrity of the content and the design. "The Visual Display of Quantitative Information" was in the list of 10 bestsellers in earth's biggest bookstore: Amazon.com. Tufte's previous books have won prizes for their own design and you will find his new book also a visual delight. From his previous books, we also know that we will find wonderful historical examples of the right way to display information and their modern counterparts.

In Tufte's first book we had the graphic display of Napoleon's defeat by the Russian army and winter in the year 1812. Tufte remarked: "it may well be the best statistical graphic ever drawn." The annual weather map of the New York Times provides a modern example of a very informative graph. In "Visual Explanations" we saw Louis Bretez's wonderful "Plan de Paris," showing every building of this city and a modern counterpart, Constantine Anderson's three dimensional map of midtown Manhattan. In this new book, we see how John Snow, with his map of the area surrounding the Broad Street Pump, showed the evidence that the cholera epidemic in England in 1884 was caused by the drinking water. He used this map to persuade the authorities to remove the pump handle, an action that was credited with ending the epidemic. Tufte contrasts this with a modern example of the incorrect display of graphic information. He shows us the charts providing information on the failure in tests of the O-rings on space shuttles. These charts played an important role in allowing the Challenger space shuttle to go despite the cold weather. He shows that, if the information on 0-rings had been properly presented graphically, the danger to the O-rings of cold temperatures would have been evident and could have prevented the 1986 Challenger disaster.

A novel chapter, co-authored by Jamy Ian Swiss, a professional magician, illustrates how magicians disguise information and shows that we can learn from them what not to do. This even includes a few hints for our lectures: magicians never explain exactly what they are going to do -- a lecturer should. Magicians never repeat a trick -- an important idea in a lecture is worth repeating. By now we are supposed to have learned from Tufte to let the data speak for itself and not to litter the graphics with a lot of junk. Alas, as Tufte observes, many of the web-page designers have not learned this. They clutter their homepage with large buttons, flashing lights and other paraphernalia that disguise what in on their web sites. As usual, Tufte is not content to just make critical remarks -- he gives us good examples and shows what should be done.

A common theme of Tufte's books is the proper use of color. Here again we see what a difference it makes, in conveying information, when colors are subtly integrated into the graphics instead of distracting us by their gaudiness. It would be an interesting study to see if Tufte's impressive books have "made a difference". For example, in his first book his own study showed that, among all the wonderful graphical ways we have to convey information, only maps, time-series plots and bar and pie charts are typically used in newspapers. One seldom sees graphs that relate two or more variables despite the wide use such graphs in scientific publications. Are things any different 20 years later? Our own impression is that the answer is no for the newspapers, but yes for scientific journals, especially medical journals. In this latest book Tufte shows us what we should be doing when we want to convey information that assesses change, dynamics, and cause and effect, which Tufte regards to be at the heart of thinking and explanation. He does it in a book that is a joy to look at, to read, to think about it, and even to review. [Source: Picked up from the Internet and forwarded by **Donna Stroup**, EPO]

News Article: The article "Hantavirus Risk Assessed by GIS," appeared in the Supplement to *Government Technology*, Dec., 1996. It describes a collaborative

project directed by **James Childs**, Division of Viral Rickettsial Diseases, CDC, to model geographic variation in rat bite risk, in New York City, using GIS. The GIS work was done at Hunter College, Department of Geography, under the direction of **Sara McLafferty**. The purpose of the project is to determine the possibility of Hantavirus infection from rat-bite cases among high-risk populations. Using rat bite data at the census block level, along with randomly selected control (no bite) blocks, logistic regression models were used to determine which variables were most significant in predicting the likelihood of rat bites occurring in a particular block. In addition, by modeling each borough independently through a series of equations, the number of variables were reduced. A more accurate fit of the data was achieved for models of each borough. Preliminary findings suggest highest predictors of rat bites include distance to subways and parks, median income, census block population density and persons under 18 years of age.

VI. Related Census, DHHS and Other Developments

Excerpts from the September 18-19, 1996 meeting of the NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS, Public Health Service, Washington, D.C.

EXECUTIVE SUMMARY: The Committee heard reports on the HHS Data Council, the National Center for Health Statistics, the HHS Research Planning Initiative, HCFA, and AHCPR, as well as the Health Insurance Portability and Accountability Act of 1996 (Kassebaum/Kennedy) and its implications for the Department and the Committee.

The Kassebaum/Kennedy bill received considerable attention because of its challenging demands and mandates for both the Committee and the Department. The new NCVHS charge in the legislation and attendant timeline were reviewed in detail and served as a major reference point for the Committee's discussion of its work plan. The new responsibilities of NCVHS under the bill are: two additional members, named by Congress; annual report on status of implementation, beginning one year after implementation, and; various forms of assistance and advice to the Secretary on standards and privacy. A task given almost entirely to NCVHS is the development within four years of recommendations on standards for medical record information and its electronic data interchange. The possible revision of certain aspects of the timelines was discussed, particularly so that privacy protections are in place before standards are promulgated.

It was noted that the priorities mandated by the Kassebaum/Kennedy bill are data standards, privacy, and the electronic medical record. The Committee agreed that the first two are the immediate priorities while addressing the third can be deferred for a few months.

WELCOME FROM THE DEPARTMENT AND HHS DATA COUNCIL: Mr. Jack Ebeler, Deputy Assistant Secretary for Planning and Evaluation (ASPE), began by being the first of many speakers to thank those leaving the Committee and to welcome those just joining it, together with the outgoing and incoming Chairs. He noted the Committee's reputation for stature and effectiveness, and said that it and the Department have a lot of work to do.

The Data Council is formally composed of the senior leadership of HHS, who report to the Secretary, including all HHS heads of operating and staff divisions, all assistant secretaries, and all agency heads. Some agency heads delegate their roles to the person in the agency in charge of information issues. The Council was created to advise the Secretary and serve as the Department's focal point on data policy issues. In its role as the liaison for NCVHS, the Council has recently participated in rechartering the Committee and selecting 10 new members from nearly 200 nominees.

Mr. Ebeler and HCFA Administrator Bruce Vladeck have co-chaired the Data Council, which meets monthly. The Council's Work Plan is on the home page, listing six major themes:1.HHS data collection strategy- Major activities here are continuing work on the Survey Integration Plan and the HHS Initiative on Research Planning (on which Dr. Raub will report later in this meeting). The Council is also working on strategies for state-level data. 2.Data

standards-This involves both the Department's own standards activities and its participation in external bodies working in this area. 3. Work with NCVHS-The Council is setting up a parallel staffing structure for the two bodies, and serves as the vehicle for receiving and responding to NCVHS recommendations, e.g., Dr. Starfield's recent presentation of the core data element recommendations. Dr. Detmer will participate in the meetings of the Council. 4.HHS focal point for privacy-related issues 5.Focal point on National Information Infrastructure (NII) activities 6. Standing Department-wide forum for other issues as they arise

Mr. Ebeler noted that in general, the Department needs the National Committee to keep it abreast of the rapid changes in the health field. Asked for his thinking about the Committee's role in respect to the Council's research planning, the NII, and national standards groups, he said this should flow largely from the Committee's own priorities and work plan. At a minimum, the Committee will be asked to comment on the research plan, and it is the formal external advisory body in the area of external relations. Dr. Detmer said he planned soon to meet with the Executive Subcommittee to develop a work plan, taking the Data Council's work plan into consideration. As for the information infrastructure initiative, Mr. Ebeler and Mr. Scanlon said the Committee would be given copies of the Department's reports to the Vice President, in some areas of which (privacy and standards) the Committee is already involved. The Committee can decide how much it wants to get involved with the other two, telemedicine and enhanced consumer information. Mr. Ebeler said the Data Council has working groups on survey integration, telemedicine, privacy, health data standards, racial and ethnic data, and international health data collaborations.

Outgoing Chair Judith Miller Jones observed that the Committee is entering a new era of direct communication with Congress, as a result of the Kassebaum/Kennedy bill's provision for two Congressional appointments to the Committee. Outgoing Committee Member Dr. Nicholas Zill urged the Committee to monitor the Department's ability to implement the ambitious plans it has articulated, e.g., the redesign of the National Health Interview Survey and the maintenance of a vital statistics system that is crucial to the success of performance partnerships.

HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT OF 1996: ROLES AND RESPONSIBILITIES -- ADMINISTRATIVE SIMPLIFICATION: Mr. James Scanlon, HHS Executive Staff Director, said the Health Insurance Portability and Accountability Act of 1996 (Kassebaum/Kennedy), which was signed into law in August, contains a number of health insurance reforms, as well as administrative simplification provisions.

The administrative simplification provisions outline a national framework for the adoption of standards for electronic data interchange in the health arena. The provisions focus on standards, code sets and classifications, security issues, unique identifier numbers, electronic signatures and authentication, and privacy issues. The overall purposes of the bill are, first, to improve the Medicare and Medicaid programs and the effectiveness of the health care system through standards for the electronic transmission of health information, and, second, to protect the security and privacy of that information.

Mr. Scanlon outlined the challenging timeline for the various provisions of the bill, all time periods which begin with the date of enactment. It requires Congressional appointment of two members to NCVHS within 60 days, the Secretary's recommendation of privacy standards and NCVHS reports on implementation to Congress within 12 months, and the Secretary's adoption of the majority of transaction and security standards within 18 months. Further actions by both NCVHS and the Secretary are required in successive 6-month stages, to 54 months. Within 18 months, standards are to be adopted for health claims or equivalent encounter information, coordination of benefits information, electronic data interchange of health care payment and remittance advice, first report of injury, and health claim status, among other things.

HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT OF 1996; ROLES AND

RESPONSIBILITIES -- PRIVACY: In response to Mr. Steinwald's comment, Ms. Renee Landers, Deputy General Counsel at HHS, suggested that perhaps Congress wanted to affirm the authority of the legislative branch to address privacy issues, rather than leaving their resolution to the discretion of an administrative agency. She characterized the action also as a gesture in the direction of federalism, which leaves intact state laws that are more stringent than the federal law.

The statute requires the Secretary to make detailed recommendations to Congress regarding the privacy of individually identifiable health information, to be completed within 12 months. The statute specifically provides guidance regarding what the recommendations are to cover. If there is no legislation on privacy standards within 36 months, the Secretary is required to issue regulations on electronic information privacy standards within 42 months. This regulatory authority appears to apply only to electronically-transmitted claims information, not to health records generally. Ms. Landers reiterated that the privacy regulations would not supersede any more stringent provisions of state law.

In conclusion, Ms. Landers stressed that the Department is very aware of the implications of the bill's standard setting provisions for making individually identifiable health information available more readily, thus raising the urgency of the need for adequate privacy protections.

WELCOME AND PERSPECTIVES FROM SENIOR ADVISOR FOR HEALTH STATISTICS: Dr. Detmer welcomed Dr. Ed Sondik, Director of NCHS and the Secretary's Senior Advisor on Health Statistics. Dr. Sondik framed his remarks in terms of the challenges facing the Department and the Committee, calling this the most challenging time in history in health statistics and health information systems. One challenge is in area of coordination and integration, in which the Department is trying to assess its evolving role in the health care field. He stressed the importance of working across all departments of government to deal with the changes in the health, welfare, and information systems. In this area both the Data Council and the Committee have critical roles, working closely together. Another area is in survey integration, where joint efforts and elimination of duplication are important but where it is also prudent not to over-integrate and over-centralize.

The Department also needs the Committee's guidance in evaluating the rapid changes in the health care system and building cooperation both within the Department and between it and outside stakeholders. He noted that the National Institutes of Health, which he recently left, is "constantly seeking outside guidance" on its programs and directions. He then cited several examples of partnerships and consultations, beginning with recent wide-ranging discussions with the field about how to make the best use of the National Health and Nutrition Examination Survey (NHANES). Partnership is also being developed with the National Library of Medicine, and will be cultivated with the Department of Defense and the Veterans Administration. He invited the Committee's perspectives on how to facilitate such partnerships.

Another challenge is to make the fullest use of existing technology and plan for that which is coming down the line. Dr. Sondik reported that the NCHS Vital Statistics Program has changed the way vital statistics are processed and reported. The result of this and a country-wide effort is that the 1995 statistics will be released several months earlier than in the past. The new approach uses continuous processing and automated systems.

The Kassebaum/Kennedy bill represents another set of challenges in which the Committee's advice to the Department will be essential. Implementation of ICD-10 is another area, one in which NCVHS has had a longstanding role. In this area as in others, the Department's linkages to the community are critical.

Turning to the National Center's major challenges, Dr. Sondik began with monitoring and evaluating the changes in health care, prevention, and the health care system. He noted the need to deal with both the forest and the trees -- i.e., to produce and sustain a picture of what is happening in the country in general, while also producing data on the local picture. Much of the action, he noted, is now at the local level, be it states or smaller areas, but national trends must

also be sustained. This raises questions about the best allocation of the information budget and what kind of structure would be most useful to localities.

Dealing with the genetic revolution is another challenge, in which thought must be given to how to use large scale studies for information while protecting individual rights. He noted that consideration is being given to using NHANES III, a representative national sample, to study the prevalence of various genes.

A challenge that faces both the Center and NCVHS is how to get valuable information into the hands of those who can use it. Dr. Sondik noted that both person-to-person communication and the use of technology have a role to play here. As an example of the latter, he called attention to the White House Home Page, which has a Social Statistics Briefing Room with an area on health statistics. A notable facet of this service is that those at NCHS who are responsible can post whatever they think is important without having to clear it with anyone. All information on the home page includes linkages to sources of additional information.

REPORTS FROM [Selected] SUBCOMMITTEE CHAIRS

Subcommittee on Minority and Other Special Populations: Outgoing Committee Member Dr. David Williams said he would highlight ongoing issues for this Subcommittee, beginning with the translation of the National Health Interview Survey (NHIS) into Spanish. This is especially important in view of the oversample of the Hispanic population and the need for data quality through consistent use of one translation. Cost-based objections have been raised, but none seems insurmountable. The Spanish translation should be done before use of the CAPI program begins for the NHIS. Another ongoing issue is the introduction of sociocultural data in Department surveys.

The need to monitor the impact of changes in the health care system on vulnerable populations, and to assess the possible need for new data systems to monitor developments, is growing in importance. The Subcommittee is particularly concerned about mechanisms to monitor the transition into managed care of Medicare and Medicaid populations, and the ability to monitor the effects of these changes on minority and other low income populations.

Another issue of continuing concern is potential changes to OMB Directive 15, about which the Subcommittee has been routinely briefed. It is very concerned about the quality, validity, and continuity of race and ethnicity data. A focal issue is the public pressure for a multi-racial category in classifications. State practices in this regard are being monitored.

The Subcommittee's study of racial disparities in insured populations also needs to be monitored, and may be something on which the NCHS Minority Health Statistics Grants Program can focus. Dr. Williams noted that the Subcommittee has provided oversight and advice to this program, which is currently reviewing its role. The Subcommittee also has followed the Department's plans for the Report on Social Class and Health, which needs further followup. Another issue is the health and data needs relating to the heterogeneous Asian and Pacific Islander American population.

Dr. Sondik said that in the interest of quality, NCHS is revisiting the question of translating the NHIS questionnaire into Spanish. He added that the Data Council has established a standing work group on race and ethnicity data, which will serve as the focal point for federal-wide data standards in that area. Ms. Jones pointed out that the purview must extend beyond race and ethnicity to socioeconomic status. She commended to the Committee a survey of these issues and their interrelation written by Dr. Williams, and urged the Committee to help make the issues more intelligible to Congress and the general public.

Subcommittee on State and Community Health Statistics: Dr. Zill stressed the importance of providing state-level estimates to get at such things as state variations in Medicaid waivers and the quality of Behavioral Risk Factor Survey data. He asked Dr. Sondik about plans to produce state level estimates. Dr. Sondik said the Center will experiment with expanding the National Immunization Survey and using it as a kind of "survey engine." The hope is that at a minimum, this would be a resource states could use. He noted that the local estimates generate significant buy-in from the community because of their

credibility, and this is a spur to local action. Questions remain about the cost of doing this across the country, but his hope is to find a cost-effective mechanism. Dr. Detmer affirmed the value of local statistics as a stimulus for social action.

Mr. Scanlon said the Department is rethinking various strategies for State-level data and would like the Committee's advice in this regard. The areas being explored include using national surveys to develop state estimates, linking state telephone surveys to interview survey information, and analytic linkages with administrative data.

HHS RESEARCH PLANNING INITIATIVE : As to specific recommendations, the work group stressed the importance of the basic surveys in both health and other areas, together with the need for continued investment in their refinement to ensure that they address the right questions. The report also emphasizes that national level information is being lost through shifts to managed care, devolution, block grants with reduced information reporting requirements, and other factors. Another area of concern is the desperate need in every agency for subnational data, including both smaller geographic areas and selected population groups. Oversampling within national surveys is often prohibitively expensive, and less expensive ways of getting higher resolution must be found.

In addition, some funds have been allocated for research and evaluation to pursue some of the themes of the report. One example is a NCHS project to investigate the use of telephone surveys to get subnational information, using the infrastructure of the National Immunization Survey to ask additional questions. Among other things, researchers will be evaluating the weaknesses related to the fact that key segments of the population do not have telephones.

WELCOME AND PERSPECTIVE FROM AGENCY FOR HEALTH CARE POLICY AND RESEARCH: Dr. Clif Gaus, AHCPR Administrator, said he would focus his remarks on AHCPR's work on the Medical Expenditure Panel Survey (MEPS) and on data standards. The MEPS grew out of this integration effort. It consists of five component surveys. The Household Survey component collects data from a nationally representative subsample of households that participated in the 1995 National Health Interview Survey. The new design uses an ongoing survey, with each household surveyed for two years. The other component surveys supplement the Household Survey with information about health care use and costs from nursing homes, medical providers, health insurance providers and employers. The survey is integrated with the Medicare Current Beneficiary Sample, thereby eliminating duplication and unnecessary costs associated with dual surveys.

The panel survey now in the field is providing national estimates of health care utilization, health care expenditures, sources of payment, health insurance coverage, and health status. The subsidiary surveys come off of the initial household sample of 9,000 households. One is the Medical Provider Survey, which yields information on the provider and verifies information on services rendered to household members. The insurance component is analogous. The survey also obtains detailed information through interviews with individual employers or insurance companies and from abstracts of health insurance information from health plans. There is a separate frame of employers, which will be in the field next year. The sample of 30,000 establishments will permit disaggregation down to a state level to track changes throughout the country.

Dr. Amare asked for Dr. Gaus' views on how well the surveys and data sources cover issues of ethnic and racial classification and socioeconomic status. He noted that AHCPR and ASPE took stock of that question and found a great lack of uniformity among agencies. He favors changing the categorization system, and finds the current system inadequate for public policy and clinical decisions in terms of the racial/ethnic categorization that is vital to the quality of care. At his request, Dr. Ross Arnett reviewed the limitations of the surveys in respect to ethnic subgroups, and said the Department is looking for ways to zero in on subpopulations. Mr. Scanlon noted three things: a directory of all the HHS data collection resources related to race and ethnicity that is included in the HHS Home Page, an active HHS working group on race and ethnicity, and a new standing group on race and ethnicity in the Data Council. [**Editor**: For

more information see HHS home page //aspe.os.dhhs.gov/datacncl/index.htm>] <http://www.os.dhhs.gov> and the DHHS Data Council's home page (via ASPE home page) <http:

****

Net Site of Interest for this Edition: http://ttnwww.rtpnc.epa.gov [see IV. Special Report, The US EPA's Aerometric Information Retrieval System (AIRS) in this edition]

Final Thought

Do you know the meaning of 'Bhattacharyya distance'?: A technique used in Digital Image Processing that measures the statistical separability of spectral classes, giving an estimate of the probability of correct classification.

Charles M. Croner, Ph.D., Editor, ***PUBLIC HEALTH GIS NEWS AND INFORMATION***, Office of Research and Methodology, National Center for Health Statistics <cmc2@nch09a.em.cdc.gov>

Spring of '97 ... stay in GIS touch