

# PUBLIC HEALTH GIS NEWS AND INFORMATION

December 1997 (No. 19)

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

\*\*\*\*\*

ANNOUNCING: ATSDR's 1998 "GIS IN PUBLIC HEALTH" CONFERENCE!

August 17-20, Town and Country Resort and Convention Center, San Diego

[Requests for Preliminary Announcement and Call for Papers may be sent by e-mail to [GIS@CDC.GOV](mailto:GIS@CDC.GOV)]

\*\*\*\*\*

**Selected Contents:** Conferences (p.1); News from GIS Users (pp.1-6); GIS outreach (pp.6-9); Special reports (pp.9-10); GIS literature (pp.10-12); NCVHS news (pp.12-19); Net site of interest (p.19); Final thought (pp.19-20)

## I. Public Health GIS (and related) Events

☞ HazWaste World/Superfund XVIII Conference, E. J. Krause & Associates and the Environmental Industry Associations, Washington, D.C., December 2-4, 1997 [Contact: <http://www.ejkrause.com/enviroshows> or Susan Cantor at voice (301) 493-5500]

☞ Land Satellite Information in the Next Decade II: Sources and Applications, American Society for Photogrammetry and Remote Sensing and the North American Remote Sensing Industries Association, Washington, D.C., December 3-5, 1997 [Contact: <http://www.asprs.org/asprs/index2.html> or voice (301) 493-0290]

☞ International Conference and Workshop on Interoperating GIS, National Center for Geographic Information and Analysis and the Open GIS Consortium Inc., Santa Barbara, December 3-6, 1997 [Contact: e-mail [i20@ncgia.ucsb.edu](mailto:i20@ncgia.ucsb.edu) or voice (805) 893-8224]

☞ NCHS Cartography and GIS Guest Lecture Series, "A Study of Homicide in Washington, D.C. Using Desktop GIS," [see enclosed abstract] by DeWitt Davis, University of the District of Columbia, at NCHS Hyattsville, January 28, 1998, 2:00-3:00 PM; please make Envision arrangements now at offsite locations [Contact: Chuck Croner at e-mail [cmc2@cdc.gov](mailto:cmc2@cdc.gov)]

☞ 1998 ACM (American Congress on Surveying and

Mapping) Annual Convention and Exhibition, Baltimore, MD, February 28-March 4, 1998 [Contact: <http://www.landsurveyor.com/acsm/> or Denise Calvert at voice (301) 493-0200]

☞ 4th International Conference on Hemorrhagic Fevers and Hantaviruses, Centers for Disease Control and Prevention, Atlanta, March 5-7, 1998 [Contact: Amy Corneli, Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases at e-mail [akc8@cdc.gov](mailto:akc8@cdc.gov) or voice (404) 639-1510 or see <http://www.cdc.gov/ncidod/diseases/hanta/hant conf.htm>]

☞ International Conference on Emerging Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, March 8-11, 1998 [Contact: voice (202) 942-9248 or see <http://www.cdc.gov/ncidod/eidann.htm>]

☞ 94th Annual Meeting of the Association of American Geographers, Boston, March 25-29, 1998 [Contact: AAG, 1710 16th Street NW, Washington, DC 20009-3198 or voice (202) 234-1450]

## II. News from GIS USERS

(Please communicate directly with colleagues on any issues)

### A. General News (and Training Opportunities)

1. From **Karl Sieber**, NIOSH: Steve Spaeth and myself from NIOSH attended GIS-LIS '97, Oct. 28-31, 1997 in Cincinnati. There were a variety of short courses, ranging from an introduction to GIS to more advanced topics. (Both of us took the intro

course-upon reviewing it and the notes from the course you [Chuck Croner] presented at the CDC symposium on small area statistics in 1995, your course was definitely more useful in giving specifics on use of TIGER files, etc. and other data useful for GIS). The conference was more a trade show, and I learned about many products and data sources available on the web. While there, I met several folks who work for the EPA and actively use GIS techniques, as well as folks involved in mapping nuclear disposal sites and water quality.

2. The National Center for Health Statistics (NCHS), CDC, commemorated the 10th annual NCHS/National Geography Awareness Week (GAW) celebration with a presentation by **Lee De Cola**, Research Physical Scientist, US Geological Survey. His talk, "Visualizing Demographic Statistics through Map Animation: Implications for Human Disease Databases" (retitled) sent a powerful message about the marketability of dynamic mapping. In less than one minute, the growth of San Francisco was portrayed by decade over a 150 year period. The response by television news media in SF, in 1994, was exceptional with all stations providing coverage (and creative drama) to their respective presentations. Maryland Governor Glendening has shown similar support for Lee's 1997 animation of population growth over 200 years between Washington, DC and Baltimore. Lee used Mathematica and GIF Animator for the animation, though the original data were compiled in ARC/INFO, and some output was done using ERDAS. If you have yet to see this animation, you can visit web site [http://geog.gmu.edu/gess/classes/geog590/gis\\_internet/ldecola/baltwash](http://geog.gmu.edu/gess/classes/geog590/gis_internet/ldecola/baltwash). Eight CDC locations joined Lee's presentation through Envision. The GAW program is part of the NCHS Cartography and GIS Guest Lecture Series. Editor

3. January 28, 1998 NCHS Cartography and GIS Guest Lecture Series presentation, "A Study of Homicide in Washington, DC using Desktop GIS," by Dr. Dewitt Davis, University of the District of Columbia. Abstract: The research is a longitudinal study which examines homicides at the street, census

tract, ward, and quadrant levels. These are the geographic units utilized by the different agencies in the DC City Government when making political and economic decisions. The study also attempts to establish if there is cause and effect in relation to relevant 1990 US census socioeconomic data and Washington DC Metropolitan Police total homicides data for several year periods from 1990 to present. [Editor: Dr. Davis taught geography at Ohio State University (1973-1980), University of Arkansas Pine Bluff (1980-1984), and University of District of Columbia (1984-present). Since arriving at the University of the District of Columbia, he chaired the Department of Geography four years, was acting assistant dean for one year, and is currently the coordinator of the Geography Program in the Department of Urban Affairs and Geography]

4. Genetic Speaker Series, CDC: The Office of Genetics and Disease Prevention, in collaboration with the Office of the Associate Director for Science and HRMO announces a distinguished speaker series on ethical, legal and social issues in genetics and public health. "Genetics, Public Health and the Law" is the topic of a December 3 presentation (1:00-2:00 p.m., Auditorium A, Clifton Rd) by Ellen Wright Clayton, J.D., M.D., Associate Professor of Law, Vanderbilt University. Dr. Wright Clayton is a member of the National Advisory Council for the National Human Genome Research Institute and has been writing about legal and ethical issues posed by developments in genetics for the last 20 years. Dr. Wright Clayton will begin her presentation by discussing the constitutional foundations of the government's power to protect the health of its citizens, focusing on the particular issues raised by public health efforts in genetics. She will then discuss the limits on this authority as well as governmental actions that might give rise to liability.

#### B. Technical News

5. From **Bob Wentworth**, EPA: I'm part of a team working to increase public access to environmental data via the EPA's Envirofacts data access system. I will share your feedback with the team. We are interested in knowing more of the responses from the

public about their reactions to our new home page and if they may have any issues with the following: loading of the new page, colors, are they getting any kind of error messages, either from Netscape or from Envirofacts, or otherwise, and how our new functionality works for them; how intuitive is it, e.g., the 'mouse over' capability. This will be valuable information for us to have. The more diverse our target audience for these tests, the better. Please feel free to pass along this request for feedback! We're hoping to develop a data access system that will be as useful as possible to the public. Let us know how we can improve! The URL for Envirofacts is: [http://www.epa.gov/enviro/html/ef\\_home.html](http://www.epa.gov/enviro/html/ef_home.html). My e-mail address for response is: [Wentworth@merc.rx.uga.edu](mailto:Wentworth@merc.rx.uga.edu).

6. From **Loren Hall**, EPA (through [ppgis-scope@igc.org](mailto:ppgis-scope@igc.org)): Here are a few suggestions for places where you can use on-line mapping and find out more about applying GIS in the environmental justice context: **On-line Mapping**: Pollution Mapping Projects and Toxics Databases (collection of links) at <http://www.envirolink.org/issues/pollution-map/index.html>; EPA's Maps on Demand (not real time, but can produce large and detailed maps with point sources, population characteristics, drinking water facilities, etc, but also be aware that large maps in some formats need lots of PC memory to display) at <http://www.epa.gov/enviro/html/mod/index.html>; Surf your watershed (interactive display and retrieval of data by watershed - portions require a Java-capable web browser) at <http://www.epa.gov/surf/>; Demographic Data Viewer (national interactive Census data thematic mapping) at <http://plue.sedac.ciesin.org/plue/ddviewer/>; West Virginia state interactive mapping site (includes a Java applet that you will need a recent web browser to use) <http://poca.osmre.gov/form/frame-1.html>.

**Data Sources**: Government Information Sharing Project: Oregon State University - collection of national and state-specific population, housing, economic data, etc. - suitable for download and mapping) at <http://govinfo.kerr.orst.edu/>; Egrets online guide to geographic information about East St. Louis (mostly a map library of products created by the East

St. Louis Action Research Project) at <http://www.imlab.uiuc.edu/egrets/>; Guide To On-line And Mostly Free U.S. Geospatial and Attribute Data at <http://www.cast.uark.edu/local/hunt/index.html>.

7. From **John Gardenier**, NCHS: On December 15-17 at NIH, there will be a Conference on Scientific and Technical Data Exchange and Integration. It is sponsored by the U. S. National Committee on CODATA, which represents U.S. interests at the international Committee on Data for Science and Technology (CODATA), an Interdisciplinary Committee of the International Council of Scientific Unions. Additional information is available on-line at <http://www.cisti.nrc.ca/programs/codata/welcome.html>, including the full program of this and other related conferences.

Within the conference, I have been invited as a Panelist at the Breakout Session on Bioinformatics, Tuesday 12/16, 1:15-4:15. [Panel description and participant list at the end.] Having only a 15 minute window, I would still like to mention a range of the topics most crucial to DHHS components, to biostatisticians, and to the statistics profession as a whole. If you would be willing to provide a paragraph or two per topic, I would be happy to "get it on the record" at the conference - with your name and affiliation, of course. Whether or not you can do so personally, you might (also) refer this request to someone else whose input you would like to see included.

Apart from whatever else you may suggest, the topics I am currently hoping to touch upon briefly are: (1) Concerns about potential adverse impacts on data sharing from proposed treaties and legislation regarding new forms of intellectual property protection for databases. Emphasis on the AAAS and American Statistical Association (ASA) positions. (2) Concerns about statistical ethics and competence in biomedical research and reporting - with special emphasis on problems of advocacy, bias, and incomplete/misleading reporting of experiments and analyses. Activities of the ASA Committee on Professional Ethics (John Bailar). (3) Concerns about the criticality of global access to sound data-gathering,

analysis, and reporting methods. Mention of at least the CDC EPI-INFO epidemiological computer system as an example of a globally distributed, inexpensive, flexible, and useful product promoting and protecting public health (Andy Dean). (4) Concerns about the explosion of data coming out of molecular biology and genomic research, its non-statistical and quasi-statistical properties requiring new mathematical and computer science approaches and the dissemination of both the new knowledge and the tools to take advantage of it (Genome project; ACM?). (5) Concerns about data policy and standards and their impact on the comparability and intelligibility of data from various national and international data systems (Various sources?). (6) Capabilities and limitations of geographic information systems (GIS) in both analysis of data and in its presentation to policy makers, journalists, and the public at large (Charles Croner, others?).

Each topic, of course, is worthy of/has merited entire conferences so there will be no attempt to get into technical detail. For interested attendees, it would be useful to provide a handout of (non-exhaustive, but interesting) bibliographic references. Anything you would be willing and able to contribute would be greatly appreciated and acknowledged. Here is the Chair's description of the core questions for the panelists: The breakout session on Bioinformatics will include presentations by a panel of speakers who represent research, government and academic organizations. The panelists will be asked to address the issues of a) identifying areas in which data exchange and integration in bioinformatics are important, including exchange and integration of biological data and that of other sciences. Some possible examples are the interactions between molecular biology information and taxonomy, ecological risk assessment, drug design, epidemiology, biodiversity studies, sustainable agriculture, emission of greenhouse gases, etc. b) addressing examples of successful or planned efforts to foster serious and significant cooperation among scientific and engineering disciplines and different types of organizations in science and technology; as well as identifying factors which foster such cooperation and

factors which impede it.

More specifically, the core questions for the panelists are: 1) Define important and emerging issues involving bioinformatics in which multi-disciplinary data exchange and integration are or will be necessary. Give brief examples of the types of data that are involved. 2) List the driving forces (scientific, information technology, applications, policy, organizational, etc.) that will make this data exchange and integration happen. Identify barriers (such as scientific, information technology, policy, organizational, etc.) that might impede this process. 3) Given that most, if not all, Federal funding agencies have biological data components to their programs, what are the roles of the federal funding agencies? Inter-agency groups? Professional and technical societies in the biological and other sciences? Standards organizations? Industry? Scientists themselves? Others? 4) What actions can be taken to overcome identified and potential barriers? How can CODATA help? 5) What is needed to educate the communities involved in the varied areas of biological sciences and associated technologies about existing computer science approaches and other information technology developments applicable to the issues of data access and dissemination? 6. What is the best scenario for five years from now? What else is needed to get there?

The participants and references are: I. Introduction: Micah I. Krichevsky, Chair, Bionomics International & George Mason U.; II. Brief Presentations by Panelists Lois Blaine, Rapporteur, ATCC, Peter Buneman, U. Penn., John Gardenier, CDC, Chris Overton, U. Penn., Mark Segal, EPA; III. Panel Discussion (with session participants); IV. Conclusions. References: 1) The study on "Bits of Power: Issues in Global Access to Scientific Data" was recently published. It investigated the barriers and other issues in the international transfers of scientific data with the goal of improving access to scientific data and services internationally. The study contains recommendations in the area of bioinformatics bearing on the discussion (<http://www.nas.edu/cpsma/bits2.htm>); 2) The Committee on a Pilot Study for Database Interfaces analyzed the data management

problems associated with interfacing different types of environmental data, especially small-scale ecological data with large-scale geophysical data. The Committee's report, "Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data", was published in 1994. The study also contains pertinent examples and recommendations germane to the discussion (<http://www.nap.edu/readingroom/enter2.cgi?0309050820.html>)

### C. Internet News

8. USGS Launches New Electronic Clearinghouse for Biological Data: Dennis B. Fenn, Chief Biologist, today announced that the U. S. Geological Survey's Biological Resources Division has initiated a new Internet-based clearinghouse that can be used to search for and locate existing sources of biological data and information from a variety of sources. This clearinghouse is part of the National Biological Information Infrastructure (NBII), which is a cooperative effort led by the USGS to increase access to biological data and information maintained by a variety of Federal and State government agencies, universities, museums, libraries, and private organizations.

Through the NBII Clearinghouse <<http://www.nbii.gov/clearinghouse.html>>, Internet users can search through an assortment of standardized descriptions of different biological databases or information products to identify those that meet their particular requirements. These descriptions (metadata) concisely convey such things as subject matter; how, when, where, and by whom the data were collected; whom to contact for more information; and how to access the database or information product. The NBII Clearinghouse includes metadata descriptions of biological databases and information products developed and maintained by USGS scientists, as well as data and information developed and maintained by other NBII participants, including Federal and State government agencies, universities, and private organizations. The NBII Clearinghouse also functions as a part of the National Spatial Data Infrastructure (NSDI) Clearinghouse <<http://www.fgdc.gov/clearinghouse/index.html>>, as many of the

biological databases described in the NBII Clearinghouse employ geospatial references.

"We are very pleased to be offering this new service to the public, resource managers and scientists, and anyone else interested in locating existing sources of biological data and information," said Fenn. "This is a great opportunity not only for us to help get the results of USGS biological science out to those who can use the data and information, but also for us to provide a tool that our partners and cooperators in NBII can use to help share their own data and information."

Users can search through the NBII Clearinghouse (much as they would use a card catalog in a library) using a variety of criteria, such as the name of the investigator or author who collected the data or produced the information, subject-matter keywords, and spatial coordinates for the location of the study/project. Special biological search criteria, including the ability to search for data or information relating to a particular species or other taxonomic group, are also provided. Metadata descriptions in the NBII Clearinghouse are developed according to the NBII's biological metadata standard <<http://www.nbii.gov/current.status.html>>, which also serves as a biological "enhancement" or "profile" of the Federal Geographic Data Committee's Geospatial Metadata Content Standard. The mission of the USGS/BRD is to work with others to provide the scientific understanding and technologies needed to support the sound management and conservation of the Nation's biological resources. [Source: Duncan Morrow, NBS-MIB; Contact: Contact: Anne Frondorf 703/648-4205]

9. From H. Gyde Lund, European Forest Institute, Finland (through [fwim-l@listserv.vt.edu](mailto:fwim-l@listserv.vt.edu)): On 22 October 97, U.S. Sec. of Interior, Bruce Babbitt, approved the Federal Geographic Data Committee's (FGCC) Vegetation Information and Classification Standard after being endorsed by the FGDC Steering Committee. This will now be the standard vegetation classification system for use by U.S. Federal Agencies and their cooperators. The standard has been under development since 1992 and has undergone public

reviews. It may be viewed at <<http://www.nbs.gov/fgdc.veg/>>. The final format of the standard will change slightly from what is presented at this site when it is officially published, but the content is essentially the same. In addition to the public review, the FGDC vegetation standard was closely coordinated with the development of the FAO Land Cover Classification System (see website <<http://www.fao.org/waicent/faoinfo/agricult/AGL/AGLS/FGDCFAO.HTM>> to provide international links and to work towards the development of a truly global vegetation and land cover classification system. For information on the development of other U.S. geospatial standards, see the FGDC home page at <<http://www.fgdc.gov>> [Contact: Gyde at email [gyde.lund@efi.joensuu.fi](mailto:gyde.lund@efi.joensuu.fi) or web site <http://www.efi.joensuu.fi>]

10. From **Lois Dean**, HUD (through [ppgis-scope@igc.org](mailto:ppgis-scope@igc.org)) Announcement of Proposal Deadline concerning Competition for the 1998 National Brownfields Assessment Demonstration Pilots; Copies of this document are available via the Internet on the EPA Brownfields Home Page at <http://www.epa.gov/brownfields/applicat.htm>. Summary: The United States Environmental Protection Agency (EPA) will begin to accept proposals for the National Brownfields Assessment Pilots on October 9, 1997. The brownfields assessment pilots (each funded up to \$200,000 over two years) test cleanup and redevelopment planning models, direct special efforts toward removing regulatory barriers without sacrificing protectiveness, and facilitate coordinated environmental cleanup and redevelopment efforts at the federal, state, and local levels. EPA expects to select approximately 100 additional National brownfields assessment pilots by May 1998.

Applications will be accepted on a "rolling submissions" schedule. The deadlines for new applications for the 1998 assessment pilots are December 15, 1997, and March 23, 1998. Applications postmarked after December 15, 1997, will be considered in the second round of competition. Previously unsuccessful applicants are advised that they must revise and resubmit their applications. The National brownfields assessment pilots are

administered on a competitive basis. To ensure a fair selection process, evaluation panels consisting of EPA Regional and Headquarters staff and other federal agency representatives will assess how well the proposals meet the selection criteria outlined in the newly revised application guidelines document entitled "The Brownfields Economic Redevelopment Initiative: Proposal Guidelines for Brownfields Assessment Demonstration Pilots" (October 1997). This action is effective as of October 9, 1997, and expires on March 23, 1998. All proposals must be postmarked or sent to EPA via registered or tracked mail by the expiration dates cited above. The application guidelines document can be obtained by calling the Superfund Hotline at the following numbers: Washington, DC Metro Area at 703-412-9810, Outside Washington, DC Metro at 1-800-424-9346, TDD for the Hearing Impaired at 1-800-553-7672. Copies of this document are available via the Internet at: <http://www.epa.gov/brownfields/applicat.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)

☞ From **Christopher Hartwell**, Harvard University: We're looking to see if there's a GIS database out there on worldwide disease distributions, including (but not limited to ) malaria and other tropical diseases, for as far back as the stats allow. Our research team is compiling a database on important indicators for the world, as how these indicators may help or hinder economic growth. If anyone knows of GIS sites that would hold this data, it would be appreciated if they could contact me [Contact Chris at Harvard Institute for International Development, e-mail [chartwel@hiid.harvard.edu](mailto:chartwel@hiid.harvard.edu) or voice (617) 496-5783]

☞ October 27 "request for assistance" to GIS Users: **Brian McCarthy**, Division of Reproductive Health, NCCDPHP just called to say he is leaving in a few days for Jordan and has a recently approved GIS project regarding disease outbreaks in Palestinian refugee camps. Brian will use MapInfo software and a

GPS enabling link. If you know of ANY existing sources of digital maps of Jordan (especially), Syria, Lebanon, and Gaza and West Bank, please contact Brian by cob Wednesday or before by voice (770) 488-5229 or e-mail: BJM1@CDC.GOV. Thank you in advance for any help you can provide Brian. Editor

1. Response from **Morris Maslia**, ATSDR: In response to Chuck Croner's e-mail about the need for digital maps of Jordan, Syria, etc., a colleague of mine from the USGS just recently returned from the area where the USGS is being asked to be the "objective partner" for looking at the water resources of the area for the Peace Process. He indicated he had visited Israel, Jordan, the west bank area, etc. Knowing the USGS, they may have "ginned up" some quicky digital maps of all the countries of the Middle East area. Try contacting John Clarke at the USGS (Ga. District, Water Resources Division). His telephone and e-mail are as follows: John Clarke, voice (770) 903-9100 and e-mail jclarke@usgs.gov. Hope this helps and best regards.

2. Response from **Gerry Rushton**, U. of Iowa: I am copying your message to one of my colleagues, Rex Honey, who spent a year in Jordan a few years ago and had extensive contacts with their basemap people. I don't recall him ever talking about the availability of digital maps, though, so I don't know how much he may be able to help. Perhaps he can give a good contact in Amman. I'm sure that will be helpful. I have also received the following note from **Art Getis**, San Diego State University: "Thanks for the information on the Palestinian health project. If my son-in-law were still there (he now lives with my daughter and two children about one mile from us in Tierrasanta), he would likely be the liaison for the health care study. Actually, Israel land management officials have excellent maps of the region (great detail). There is a planning group in Gaza that also has good maps."

3. Response from **Grant Thrall**, U. of Florida: You might check out Maptitude version 4.0 from Caliper Corporation (Newton Mass). Caliper is just now shipping version 4. I understand version 4.0 has GPS software interlink capability; I have not seen version 4.0 myself. Again, check out Caliper [for

country files]. I understand that they have upgraded their international boundary files, which I believe comes with Maptitude.

4. Response from **Dabo Brantley**, NCCDPHP: We don't use MapInfo but if you can translate from ArcView GIS, I will scan our disk for the shape files requested below. [Dabo sent ArcView files to Brian. Editor]

5. Response from **Ron Bijeau**, Department of Defense (referred by **Miller Dayton**, NCEH): I am responding to a note from Charles Croner requesting GIS data on your behalf. I work for the Department of Defense in Atlanta, GA, and have a personal and professional interest in seeing that data produced with DoD resources are made available to other Federal Agencies (i.e., CDC). As you may know the primary source for DoD geospatial data is the National Imagery and Mapping Agency (NIMA). Their charter is to produce, manage and disseminate worldwide digital geospatial data to DoD users and other authorized Federal Agencies. You may be familiar with a NIMA product called Vector Map 0 (VMAP0). It is the smallest scale component (scale of one to one million) in a series of worldwide multi-layered GIS databases. Obviously at this small scale its application is limited to regional or broad area analyses. It is distributed by NIMA in Vector Product Format (VPF) or commercially as Digital Chart of the World in ARC/INFO format. (I've not tried these data using MapInfo yet, but I'm certain it can be done.) Other types of geospatial data (i.e., satellite imagery, raster digitized maps, digital terrain elevation models, etc.) are available depending on your requirement.

The NIMA office that services non-DoD customers can be contacted at (703) 264-3012. I believe Ms. Cindy Burns is the Chief of the Federal Agencies Customer Support Team. I don't know if they can provide you with the data you need prior to your departure, however, I personally think CDC could benefit greatly from these valuable resources in the mid to long term. Your project sounds very interesting and I trust you'll find the data you need to make it a complete success. If there is anything I can do to help you further, please feel free to call me at (404) 363-5473 or e-mail me at bijeur@forscom.army.mil.

6. Response from **Jamie Purvis**, ATSDR: Contact Bill Banks at the number below for information requested in earlier e-mail from Chuck Croner. From Bill Banks [response to Jamie]: Thanks for the information Jamie. I believe that I can get all of the material that CDC desires regarding these types of Maps. Have some one call me at (301) 865-4506 or at (717) 593-5266. I can get the Maps and Overnight mail them to you at CDC. No problem!

7. Response from **Fred Broome**, US Census Bureau: I suggest you contact the Central Statistics Office in Jordan. The Census Bureau installed a GIS system there last year to assist their census activities. Also, they have a Central Geographical Office (not certain of the name) that has quite a large amount of digital data. Maybe the U.S. Embassy can help. They (the AID representatives there) were our contacts.

Editor: Thanks to all for the timely response to Brian's situation. He was very appreciative of this assistance. I want to also acknowledge other comments forwarded to Brian from **Jonathan Mayer**, Harvard U., **John Kelmelis**, US Geological Survey and **Nina Lam**, Louisiana State University.

☞ From **Richard Hoskins**, WA State Department of Health: We are trying to develop a organizational model for organizing GIS in the WA State Dept. of Health. I would appreciate very much any information whether an org chart, document you have, or just a few lines of e-mail which describes how GIS is organized and funded in your organization. The issues are central vs decentralized, program funding vs indirect funding, and where it is located in the organization, that is, in Information Services, off on its own, part of another unit or group, etc. Also I would be interested in future plans, etc., staffing concerns, budget worries, so forth. If you are not involved with GIS organizational issues I would very much appreciate knowing the e-mail or phone number of the person who is. I guess I could do a survey, but something informal will likely serve my need. If you prefer, send me a phone number and I will call you for a brief discussion. Thanks in advance, Richard E. Hoskins, PhD MPH, WA State Public Health Geographer, GIS and Spatial Epidemiology

Unit, Office of Epidemiology at e-mail REH0303@hub.doh.wa.gov or voice (360) 705-6050.

1. Response from **Tom Richards**, PHPPPO: I think your question is excellent. I do not know the answer. The most recent GIS state survey report that I am aware of was in 1992, and so would now be pretty much out of date. Warnecke L. State Geographic Information Activities Compendium. Lexington, KY. Council of State Governments. 1992. I have heard your thought about the need to conduct a state health department GIS-related survey echoed in several conversations over the past six months. Examples include: 1) Dr. Mark Oberle, Information Network for Public Health Officials, CDC, voice (770) 488-2427, has been thinking about doing a survey and/or developing a GIS special focus within the INPHO states; the INPHO group and NACCHO conducted a survey about a year ago on electronic computer equipment in state and local health departments, but this survey did not address GIS capacity; and 2) Dr. Bill Henriques, ATSDR, phone 404-639-6088, has been thinking about doing a survey of state GIS needs. Several other groups also might be interested and/or thinking about doing a GIS related survey. Examples include: the Council of State Governments (potential contact: Dr. Keon Chi, phone: 606-244-8247; the Association of State and Territorial Health Officials (potential contact: Jacklyn Bryan, phone 202-371-9090, ext 224); and the National Association for Public Health Statistics and Information Systems (www.naphsis.com).

If you wanted to do a "quick" informal survey, one additional option that you might consider would be to contact Dr. Littleton Fowler, President of the Association for State and Territorial Local Health Liaison Officials (ASTLHLO) (phone: 405-579-2261; e-mail: Littf@aol.com ). Dr. Fowler has e-mail connections with liaisons in about 30 states, and (depending on how the question was worded and/or the number and types of questions that you have) the liaisons might be able to provide some quick informal insights. In turn, this raises another question: Should each state health department have a GIS coordinator/contact person -- with e-mail capabilities -- so your type of question might be answered more directly?

And -- does such a list already exist -- perhaps through Chuck Croner and/or the Federal Geographic Data Committee ([www.fdgc.gov.html](http://www.fdgc.gov.html)) ? I hope this helps some. Thomas B. Richards, Public Health Practice Program Office, CDC, at e-mail [tbr1@cdc.gov](mailto:tbr1@cdc.gov) or voice (770) 488-2544.

2. Response from **Chet Moore**, NCID, CDC: I am forwarding your request to several GIS specialists who may be able to offer more extensive advice. My comments relate only to our local program, as I have no experience with other programs within CDC or HHS. Also, I am an entomologist/ecologist with a strong interest in GIS, rather than a geographer/GIS specialist, so my experience is limited.

The Division of Vector-Borne Infectious Diseases (DVBID), NCID, is a small field station, so most of our activities are "stand-alone" in relation to services available in Atlanta. This is true of our GIS program, which is now nearly two years old. We have attempted to build a fairly centralized facility that is shared by the DVBID staff. This makes the most sense, given the cost of hardware, software, and data. There is no separate, formal "GIS group" or section, but the equipment is located in one room. We do make an effort to coordinate funding requests and related issues between the branches. We do not have any full-time, dedicated GIS staff at this point, although we would like to move in that direction. At the moment, we have three or four staff with serious interest in applying GIS and related methods to vector-borne disease ecology and epidemiology. To assist us in our work, we draw heavily on the expertise of faculty at Colorado State University and a USDA laboratory that has a well-developed GIS program.

We have obtained funding from a variety of sources: disaster-related activities, Emerging Infections funds, and the regular Division budget. At least in the startup phase of a program, I suspect that central funding will be important. As the program(s) grow, decentralized funding may become more appropriate. As I have never worked with a group as large as the Washington Dept. of Health, I don't have a good feel for the issues of organization. I suppose the most important considerations are that it be accessible to all staff with a need for services, and that the GIS

program should not drift off into its own world.

The thing that seems to me to be most important for successful program development is that there should be a firm, long-term commitment at the highest level. P. A. Burrough, in his introductory text on GIS, makes the comment that "It is simply not sufficient for an organization to purchase a computer and some software and to hire or retrain one or two enthusiastic individuals and then expect instant success." This brief statement could well serve as a motto for any developing program. Good luck in developing your program. If you have additional questions, please give me a call or send e-mail. Chester G. Moore, Ecology/Epidemiology Section, ADB, Division of Vector-Borne Infectious Diseases, NCID, CDC Ft. Collins, at e-mail [cgm2@cdc.gov](mailto:cgm2@cdc.gov) or voice (970) 221-6423.

#### IV. Special Reports

(Submissions are open to all)

❖ "The National Imagery and Mapping Agency Global Geospatial Information Infrastructure: A Model and Potential Source for Global Geospatial Information and Services Support"

**Ron Bijeau**, Department of Defense

The National Imagery and Mapping Agency (NIMA) [formerly the Defense Mapping Agency] is responsible for the collection, production, management and dissemination of geospatial information to Department of Defense (DoD) customers worldwide. The National Performance Review and some amazing advances in technology have resulted in some very exciting changes in how NIMA will perform this mission in the future. In order to accomplish its mission, NIMA has recognized the need to transition from an "industrial age" production to an "information age" services-based organization. One of the most exciting aspect of this transition has been the apparent broadening of its mission to potentially include providing geospatial information and services support to non-DoD customers. The overarching concept for accomplishing this monumental task is the Global Geospatial Information Infrastructure (GGII).

Millions of dollars are spent every year to provide geospatial information and services support to

meet the President's national security goals. Although there is no doubt they are required, these data should and could potentially be made available to non-DoD agencies to more effectively utilize this valuable resource. The Centers for Disease Control and Prevention (CDC) could benefit from this resource. For the purposes of this discussion let us assume CDC, and its affiliated agencies, centers, and offices, has a requirement for geospatial data worldwide. In some cases the requirement goes down to the census block or street address level.

The NIMA has been working in concert with the Open GIS Consortium to develop the GGII. The GGII is a concept for providing authorized users access to vast repositories of geospatial information via existing and future communications networks. It is conceivable that CDC centers or individual project managers, in the US and abroad, could "plug into" the GGII using existing and future information technologies. There are customer support teams at NIMA to coordinate such issues. Centers and individual project managers in the US could access geospatial data repositories from the US component of the GGII, the National Spatial Data Infrastructure (NSDI). Due to limitations in NIMA's mission to produce geospatial data of the US, coordination with other government agencies and commercial sources would be required. A liaison or coordinating office would be required to develop and establish such a program. Perhaps such an office already exists within CDC. This office should possess the following knowledge:

- NIMA organization and policies: NIMA has limited resources, therefore, a knowledge of what programs and services exist and who to speak to is very important. This knowledge will get CDC's foot in the door.
- NIMA international cooperative production agreements and treaties: There is a myriad of international programs that could either help or hinder CDC's efforts to acquire global geospatial data resources. Knowledge of how to effectively leverage these agreements and treaties would be required.
- Security classification issues: In general there is a move to declassify DoD data sources where

appropriate. Knowledge of these processes and procedures is very important for gaining access to recently declassified geospatial data sources.

- NIMA GGII concepts: NIMA has just released the first draft of the GGII Master Plan. This plan is the "blue print" for providing geospatial data directly to the user. This user may be a center, division, office or individual in the field with a laptop computer and a GPS receiver. Knowledge of this plan is very important in building an infrastructure for exploiting the vast resources of the GGII.

- USGS organization and policies: For similar reasons, a knowledge of the USGS organization and policies will ensure successful access to, and employment of, NSDI resources. Knowledge of USGS international and cooperative production programs is important for gaining access to these resources.

- Meta data: A thorough understanding of national standards for meta data and its value in the successful implementation of a geospatial data management program is very important.

- Standards and specifications: A commitment to international, national, commercial and defacto standards and specifications is very important in developing a seamless interface to the GGII.

- Developing, designing and implementing computer networks: Information technology and the effective use of wide area networks (i.e., INTERNET) are the keys to a successful program. Knowledge of what resources exist, those planned for in the future and the best tools to exploit them is critical.

- Geospatial data requirement process: Every project starts with a clear understanding of the problem to be solved. A clear understanding of the importance of clearly defined requirements and how they should be articulated to agencies with the potential to satisfy them is very important.

- GIS concepts and technology: A thorough knowledge of GIS concepts and its application to complex problems is very important.

- Contract management: The general trend in government is to contract out many functions that are not explicitly the purview of government. Therefore, a knowledge of government contracting and the development and management of contracts is very

important.

Data production and management are considered by many to be the most costly element in the development of an effective geospatial data program. It would be in keeping with the NPR and may be well worth the expenditure of resources to explore this potential resource. With further study it may be seen that an office established for the purpose of gaining access to valuable DoD and other government-owned geospatial data could potentially pay for itself many times over. [Contact: Ron Bijeau at e-mail [bijeur@forscom.army.mil](mailto:bijeur@forscom.army.mil) or voice (404) 363-5473]

### V. Public Health GIS Literature

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

GIS research award to **Grant Thrall**: Grant Thrall, Professor of Geography, University of Florida, president of Spatial Decisions and Analysis, and member of *Geo Info Systems'* Editorial Advisory Board received first place for co-authored research in a national competition for a new market analysis model for senior health care and senior housing. Grant received the award at the October annual meetings of the National Investment Conference for the Senior Living and Long-Term Care Industries (Annapolis Maryland). "The Development of a Senior Housing Expert Decision System: A Theoretical Model and Systems Framework" (*NIC Review*, volume 5, 1998, 3-16) applies to the senior housing and long term care industries the technology of GIS, gravity models, and a variation of the seven-step procedure for site selection introduced in the *Geo Info Systems'* November 1997 "ShopTalk." Editor [Grant can be reached at e-mail [thrall@afn.org](mailto:thrall@afn.org) or <http://www.afn.org/~thrall/>]

From **Roger Friedman**, NIP/CDC (the following dialogue was picked up from the ai-geostats listserv): Question- I have seen reference to a modeling method referred to casually as "ASH." I know next to nothing and have only heard mention of this method as a faster substitute to kriging with fewer distortions at the borders....Response from Gerald Whittaker- The Averaged Shifted Histogram (ASH) is a kernel density

estimation technique introduced by David Scott, Department of Statistics, Rice University. You may also see references to weighted averaging of rounded points (WARP), by Scott and Haerdle. The algorithm proceeds in two steps: first the data is binned to a rectangular array, then kernel density estimation or nonparametric regression is applied to the binned data. The ASH is discussed in detail in "Multivariate Density Estimation" by Scott. The ASH density estimation code is available from STATLIB. The code is a set FORTRAN routines called by functions in S-PLUS. At the Economic Research Service, we have found the ASH to be much faster, and generally more robust than other local regression techniques. A typical surface will cover the conterminous U.S. with a large number of features. Locfit, Loess and kriging do not do well in application to this data. An additional problem is that the data is from stratified, complex design sample surveys. There are no versions of Loess or Kriging that can deal with stratification, while the ASH deals with this in a simple way. Some results: 535,000 observations of population at the block group level from the 1990 census took approximately 15 minutes to bin in three dimensions (population, latitude, longitude) to a 600 x 400 grid on a pentium 90. The nonparametric regression then took about 5 min. There is a description of this work in Scott and Whittaker, "Multivariate Applications of the ASH in Regression," *Communications in Statistics*, v.25, 2551-2530 (1996). [Contact: Gerald Whittaker, Economic Research Service, USDA, at e-mail [gerryw2@econ.ag.gov](mailto:gerryw2@econ.ag.gov)]

From **Mike Mungiole**, NCHS (mortality data smoothing work): This study was conducted as initial research to determine an appropriate method for smoothing mortality data maps of the contiguous U.S. and the results of this method were effectively used in the recently published Atlas of United States Mortality. Past work has shown that the median-based smoother called head-banging has been a good method to eliminate spikes and retain edges in two-dimensional data (Hansen, *IEEE Trans. Geoscience & Remote Sensing*, 1991). We were interested in adding data reliability to this algorithm and extended head-banging

to include weights which were based on the variance of the data to be smoothed. When applied to both actual and simulated datasets, as expected, spikes were generally removed while ridges and clusters of high rates near the U.S. borders were maintained by both the unweighted and weighted algorithms. Adding weights, however, had a substantial effect on resulting data patterns in that they strongly influenced whether an observed spike was retained or smoothed away. In an effort to consider the appropriate degree of smoothing, we also manipulated the values of the headbanging parameters and showed how they influenced the resulting smoothed maps. In summary, this algorithm retained edges along with spikes for reliable data while unreliable isolated data points (spikes) were smoothed toward adjacent, more reliable values. For more detailed information, see "Application of a Weighted Headbanging Algorithm to Mortality Data Maps" (by M. Mungiole, L.W. Pickle, K.H. Simonson, & A.A. White) pp. 45-49, in the ASA 1996 Proceedings of the Section on Statistical Graphics.

**Gerald Rushton and Marc Armstrong**, Department of Geography, The University of Iowa: Improving Public Health Through Geographical Information Systems: An Instructional Guide to Major Concepts and Their Implementation, VERSION 2.0 (CD-ROM), November, 1997. The CD-ROM contains learner-centered instructional activities. It also contains software for analyzing the geographic pattern of disease events. The materials on this CD-ROM show the key steps in performing a detailed analysis of health data using GIS:

- acquiring digital road maps of local areas from public domain sources or from private vendors;
- acquiring software to match addresses stored in health files to the digital map;
- making tests of statistical significance for geographical patterns of diseases;
- relating the geographic patterns of incidence to socio-economic data from sources such as the U.S. Census;
- evaluating the geographic pattern of health facilities in relation to need; and

- evaluating alternative locations for appropriate health services development.

Gerry and Marc explain that their interpretation of GIS differs from the popular "desktop mapping" concept of GIS which is common in many state and federal programs. "We deal with health, environment and socio-economic data at many geographic scales of analysis, starting with the individual entity. The CD-ROM offers instruction in the matching of health information by addresses to digital maps. After this has been successfully accomplished, software available on the CD-ROM as freeware (DMAP) can be used to prepare input for GIS software to make isoline maps showing areas of high and low disease incidence rates. Alternatively, individually geo-coded health data can be aggregated to any defined geography, including census entities or other areas of interest--for example, the pollution footprint of a point source pollutant. We teach the analysis of disease distributions as continuous spatial distributions that are tested for statistical significance. Programs in DMAP are available for this purpose." [Editor: Copies of the CD-ROM "Improving Public Health Through GIS" may be ordered by mail from Department of Geography, 316J.H., The University of Iowa, Iowa City, IA 52242 or by FAX at (319) 335-2725. The cost, including postage and handling is \$20 per copy (\$10 for orders of 10 or more)]

## **VI. Related Census, DHHS and Other Federal Developments**

**A. Excerpts from the June 3-4, 1997 meetings**, San Francisco, California, "Perspectives on Privacy, Confidentiality, Data Standards and Medical/Clinical Coding and Classification Issues in Implementation of Administrative Simplification Provisions of P.L. 104-191," of the Public Health Service, National Committee on Vital and Health Statistics-EXECUTIVE SUMMARY: The National Committee on Vital and Health Statistics convened a two-day hearing on June 3-4, 1997 on perspectives on privacy, confidentiality, data standards and medical/clinical coding and classification issues in implementation of the administrative simplification provisions of P.L. 104-191. The hearing, held in San Francisco,

California, was the seventh NCVHS hearing on Kennedy/Kassebaum (K2). It was the first to deal with the full range of issues related to the legislation. The hearing had panels on insurers, health plans and providers; public health and research; public hospitals, community health centers and academic medical centers; advocacy; integrated health systems; employers; and state health data. NCVHS Chair Dr. Don Detmer said the testimony and discussion would contribute to the recommendations the Committee is mandated to make to the Secretary.

#### **PANEL ON PUBLIC HEALTH AND RESEARCH**

Ms. McCall [Laguna Research Associates] focused her remarks on the uses of standardized data to improve the health delivery system. She encouraged the Committee to use its position to explore and promote integrated standards that apply to all forms and settings of health care. Standardization must be broad-based and include standards for long-term care insurance policies. Dr. Phibbs [Center for Health Care Evaluation] focused on the potential gains of linking data from different sources. He recommended that the standards include a set of identifying information rather than just an ID number.

Dr. Flores [Sonoma City Department of Health Services] stressed the fundamental importance of maintaining the trust of all those for whom public health is responsible, something that depends on their confidence in the privacy and security of their medical information. He described the dire consequences of Proposition 187, with both documented and undocumented immigrants avoiding essential health care for fear of jeopardizing their status in the U.S. He recommended local oversight to control access by researchers and others, so that local authorities know who is using the information.

Much of the discussion with this panel focused on ways to protect confidentiality and still permit research, and on the need to make a stronger case for research to the public and Congress. There was a long and inconclusive discussion as to what meaningful distinctions can be made between research and law enforcement in terms of their use of health records.

Panelists agreed that greater scrutiny is required for data uses that focus on individuals. Mr. Gellman pointed out that researchers need to clarify the black and white areas related to confidentiality, and then deal carefully and creatively with the many gray areas.

Much of the discussion period focused on confidentiality issues and their relation to research and public health priorities. The panel and committee explored Dr. Luft's [UCSF] ideas about "approximate value" and other ways of protecting confidentiality while allowing research to go forward. In a discussion of the difficulties of getting data from health plans, it was pointed out that plans are improving their data collection and this trend will probably continue. Asked about ICIDH, Dr. LaPlante [UCSF Disability Statistics Center] repeated his point about the importance of data on functional performance, but he said the ICIDH may not be the best standard for this purpose. He and Dr. Iezzoni discussed the tradeoffs between small area research on disabilities and rare conditions to document underservice, on the one hand, and the threats to privacy this research represents, on the other. Mr. Scanlon suggested that a certificate of confidentiality from the DHHS Secretary might add further protections for researchers from subpoena.

#### **ADVOCACY PANEL**

Because of discrimination against gay and lesbian patients and people with HIV and AIDS by society in general and by the health care system, the consequences of information's getting into the wrong hands are very serious. Ms. Plumb [Gay and Lesbian Medical Association] stressed the need to give top priority to privacy and confidentiality. Her organization recommends that the Committee recognize that only universal access to health care and strong anti-discrimination laws will protect against discrimination, that it request strong measures to protect patient information, and that it call for strong penalties and sanctions for inappropriately divulging information.

Ms. Hansen [AIDS Legal Referral Panel of the San Francisco Bay Area] urged that electronic data collection and transfer of information not be done unless/until the strongest possible protections are in place, in view of past abuses and the evidence of

discrimination against people with HIV/AIDS. The AIDS Legal Referral Panel is concerned that HIPAA seriously threatens patients' confidentiality in that the administrative simplification provision emphasizes the goal of reducing the cost of health care and gives privacy concerns secondary consideration, at best. She urged that constraints be imposed on the implementation of HIPAA until a comprehensive federal privacy statute is enacted.

### **PANEL ON PROVIDERS**

Ms. Forbis [American Association for Medical Transcription] drew attention to the burgeoning of a huge, unregulated industry to transcribe dictated medical records, with serious implications for the confidentiality of those records. She called the industry attention to the fact that some transcription services are becoming data repositories and looking at how to market the information they control. AAMT recommends that information release forms include information about where documentation is stored, who controls it, and the patient's right to review it. Quality assurance standards for dictation and transcription process should be developed, and ASTM's standard on security and confidentiality of transcribed health records should be adopted.

Dr. Simons [U.S. Public Policy Committee of the Association for Computing] encouraged the Committee to call on ACM for disinterested technical assistance. She discussed the current threats to confidentiality, and called for aggressive Committee leadership in this area. She strongly criticized the use of the Social Security Number as a unique identifier. Much of the discussion with this panel focused on Ms. Forbis' alert about medical transcription and the general need for new solutions that take into account the global nature of information exchange. Dr. Detmer observed that the Committee and the government are being challenged to be unusually proactive in their approach to policy planning.

### **PANEL ON INSURERS, HEALTH PLANS AND PROVIDERS**

Turning to confidentiality, Mr. Gellman asked for comments on proposed legislation. Mr. Matejka

[California Medical Billing Association] asserted that the Bennett Bill is very restrictive, and he appealed for clarity in the guidelines about who is entitled access to data on patients. Ms. Franks noted that California already has strict patient confidentiality laws. Panelists commented on the law's provision that information can be disclosed for billing purposes without patient authorization.

Asked to comment on the National Provider Identifier, which Mr. Scanlon noted they all seem to support, the panelists noted that the longer the number is, the more room for error and that it will be good to replace the current "sloppy" systems. Concerns were expressed about possible misuse or misinterpretation in connecting providers to mortality data.

### **PANEL ON PUBLIC HEALTH AND RESEARCH**

Dr. Phibbs (Palo Alto V.A.; Stanford University; speaking for himself only) focused on the potential gains of linking data from different sources, something that will be facilitated by unique identifiers and data standardization. He recommended that the standards include a set of identifying information rather than just an ID number. To demonstrate the benefits of data linkage, he described his recent analysis of neonatal mortality and hospital patient volume, published in JAMA. Using several data sources increased the predictive accuracy of the model and yielded substantive changes in the results. He noted that it will be important to link the emerging clinical data sets to existing data, to remove the effects of selection bias, among other things. As another example of data linkage issues, he discussed the benefits of linking data in the V.A. system with those from care outside it, as veterans receive care in both places. Finally, he noted that increased data linkage also increases the risk to patient confidentiality, and he called for reasonable and appropriate steps to protect confidentiality without denying data access to qualified researchers.

Dr. Flores (Sonoma County Public Health Director and President, Health Officers Association of California) appeared on behalf of the Health Officers Association. He stressed the fundamental importance of maintaining trust with all for whom public health is

responsible, which depends on their confidence in the privacy and security of their medical information. In particular, the concerns of California's immigrant population about confidentiality and security have been severely exacerbated by Proposition 187, leading to precipitous drops in clinic attendance and the failure to get care for serious conditions. This puts both individual and societal health at risk. Dr. Flores also discussed the concerns around disclosure of HIV status. He noted the potential advantages to public health from linking information systems, benefitting immigrants as well, but also the considerable risks through intrusion by the Immigration and Naturalization Service, insurance companies, apartment owners, and others who would use medical records against people. He recommended local oversight to control access by researchers and others, so that local authorities know who is using the information. Localities also need access to information derived from shared information systems. The goal is to use health information for "systematized improvements in health care" and to avoid their use for "systematized entrapment." In conclusion, he stressed the need to assure immigrants that their health care information will not be used against them, and he urged the Committee to "first, do no harm" with its data sets.

### **Discussion**

Much of this discussion focused on ways to protect confidentiality and still permit research, and on the need to make a stronger case for research to the public and Congress. The panelists said that some, but not all, research can be done with nonidentifiable data. Dr. Flores said public health's use of encrypted identifiers for AIDS information has no disadvantages for epidemiology but prevents contact tracing. Special access may be necessary for public health. Mr. Gellman noted that some legislative proposals require patient consent before their records can be used for research. Given that every user group asserts their special entitlement to access, it is incumbent on researchers to make the case for research more effectively than has been done in the past. Failing that, research is at risk. Several panelists pointed out that there are already limits on access to data for research

uses, but they agreed that more public education is needed. Dr. Detmer noted the paucity of research on the nature and extent of abuses and problems in this area.

Mr. Gellman stimulated a long and inconclusive discussion with a query as to what meaningful distinctions can be made between research and law enforcement in terms of their use of health records. It was posited that the INS and law enforcement are interested in individuals, while researchers generally are not, but this distinction was disputed. Nevertheless, it was agreed that greater scrutiny is required for data uses that focus on individuals. Ms. Coltin pointed out that a single unique identifier that could be encrypted would obviate the need for additional data, but Dr. Phipps stressed that a dataset is necessary to definitively identify an individual. Responding to a question, he said it is theoretically possible, but not currently realistic, to have a reliable unique identifier. Mr. Gellman pointed out that any access to a person's medical record is a violation of that person's privacy, whether or not it leads to a decision regarding that individual. He also observed that researchers need to clarify the black and white areas related to confidentiality and then deal carefully and creatively with the many gray areas, including this one.

The panelists were supportive of the IRB model for health services researchers around their access to individual data. Asked about new provisions that require the patient's authorization for research use of medical records, Dr. Phibbs and other panelists said it is "a disaster."

### **PANEL ON PUBLIC HEALTH AND RESEARCH**

Dr. Newcomer stressed the need to track the service utilization of people in group housing, including both elderly and nonelderly disabled people. Currently, although as many people live in group housing as in nursing homes, the industry is virtually invisible in national data systems. It also is not clearly defined or well regulated in state and local ordinances, threatening an erosion of the quality of care. Until adequate information systems are in place, hospital discharge abstracts can be used to track marker

conditions. Another possible source is Medicare claims, to track diagnoses. Neither of these data systems is adequate, however. One alternative is to rely on survey systems, but these are national and do not yield useful community-level estimates.

Definitions are a key problem because the term "group housing" encompasses a wide range of settings and purposes. To simplify matters, states use a simpler definition based on having five or more unrelated persons in the same home. Many group homes are not included in the National Health Interview Survey and other surveys. Dr. Newcomer provided the Committee with a critique of 75 surveys, all of which have fundamental problems. He recommended putting and maintaining better information on hospital discharge abstracts and claims data as the first step toward better information in this area, and also that NCVHS or another body look at how to improve the sample frames used in the NHIS and other national surveys with respect to group housing. In addition, the disability information in the American Housing Survey and the Census must be improved.

Dr. Abbott (California Department of Health Services, appearing as an individual) noted that California has been developing approaches to many of the issues in the National Committee's mandate. He stressed the importance of standards to public health and research, with identifiers playing a key role in following individuals longitudinally and across programs. He then described his state's project to develop an approach to a unique identifier, and the evolution of a core data set for this purpose with five core and seven confirmatory data elements. The system is premised on the patient's voluntary cooperation, and uses stable data items the client can remember. The Department of Health Services has a goal of full implementation by June 1998, and it hopes other state agencies and departments will also adopt the system. They did not choose the SSN because it is not reliable and can too easily be linked with other information that can be damaging to the individual. Other stakeholders such as schools were involved in the process.

Dr. Abbott echoed Dr. Luft's recommendation that federal standards be viewed as minimums on

which states and localities can build. Regarding confidentiality, he noted the need to determine whether data are confidential in their own right or when combined with other data. California uses a kind of IRB to review requests for data by researchers and others. He described its process and criteria, noting the importance of public tolerance of the use of health information and thus the need to establish and maintain trust. He urged the Committee to maintain the balance between confidentiality and data access rather than being too restrictive.

#### **PANEL ON HEALTH PLANS AND EMPLOYERS**

The PBGH [Dr. David Hopkins, Pacific Business Group on Health] is one of the most active employer coalitions in the country. Because two-thirds of all the covered lives in the PBGH network are enrolled in HMOs, PBGH formed a negotiating alliance in 1994. It has done a variety of quality studies, leading to efforts to improve health care information systems. Its data initiative developed a three-part vision involving computer-based patient records, open-architecture systems, and built-in real-time feedback mechanisms. The Business Group also has involved a core group of managed care stakeholders in a commitment to build a data infrastructure, with a short-term focus on adopting unique patient and provider identifiers and uniform data standards. PBGH is exercising its leadership in getting employers to use ANSI enrollment standards.

Dr. Hopkins stressed that for all these initiatives, the private sector is dependent on the federal government to make key decisions. He offered the following recommendations: 1) Adopt the SSN as the core element of the unique individual health identifier. Confidentiality concerns can be met through encryption and other security measures. 2) Do not establish confidentiality provisions that will inhibit the construction of integrated databases for legitimate research and evaluation. Patient privacy can be protected by restricting the use of identifiable information. 3) Ensure that provider group identifiers fit the realities of today's marketplace, in which California may have the most complex set of arrangements. Provider group identifiers need to

indicate each provider's contractual relationship to the patient at the time of the encounter, thus connecting him/her to a particular medical group or IPA. 4) Make every effort to meet the mandated deadlines, because other stakeholders cannot act until the HHS Secretary has determined the standards to be used, and 5) After the standards are developed, publicize them widely.

#### **PANEL ON PROVIDERS [Discussion]**

Mr. Gellman questioned Mr. Schinderle about the assertion in his written testimony that the health care industry has "an excellent track record" in protecting privacy. He noted that the only published investigations in North America (Denver, 1976 and Canada, 1980) showed widespread abuses and lack of control. Mr. Schinderle acknowledged there are no independent studies to confirm his feeling that the industry takes adequate care of patient records. Mr. Kassis repeated his recommendation for accreditation to improve practices and bolster public confidence. Both Mr. Gellman and Ms. Frawley expressed support for the idea. Ms. Frawley noted that the National Research Council found that organizations have no data to validate their claim to be safeguarding confidentiality. The industry needs to develop standards in this area and hold itself accountable. Dr. Mor suggested that JCAHO, or something like it, might be an appropriate mechanism.

Asked about incentives, Mr. Schinderle said he had tax incentives in mind, and/or other ways of compensating for the up-front cost of conversion. He added that the costs are offset by savings in a ratio of about 4 to 1. Dr. Oliva said the vendors she has talked to welcome standards and expect they will cut costs. Mr. Kassis cautioned that incentives could motivate vendors to cut corners in order to meet deadlines, and Dr. Mor observed that they penalize those who do the re-engineering on their own. Mr. Schinderle said the vendor community has been unwilling to respond to Californian managed care needs because the other 49 states don't need the software yet; now the hope is that the legislation creates the critical mass to move the vendors to change their systems to make electronic commerce possible.

Dr. Detmer pointed out the need for a national

strategy for the health component of the national information infrastructure, including the computer-based record, help facilitating its dispersion, and financial incentives as well as privacy, confidentiality and security provisions and data dictionaries. These factors coalesce as part of a public/private national strategy. He thanked the panelists and organizers, and invited those wishing to make public comment.

**B. Excerpts from the June 24-25, 1997 meetings, Washington D.C. Public Health Service, National Committee on Vital and Health Statistics:** Dr. Detmer reviewed the Committee's work on standardization mandated by the Kennedy-Kassebaum ("K2") legislation, as it has prepared for forthcoming recommendations to the Secretary. The Subcommittee on Health Data Needs, Standards and Security heard from a total of 134 witnesses in eight days of public hearings, including two in San Francisco, and at full NCVHS meetings. Dr. Detmer added that the Executive Subcommittee developed an executive work plan at the San Francisco meeting. Two additional days of hearings on security standards are scheduled for August, with attention to the NRC report. The Subcommittee on Privacy and Confidentiality heard from 42 witnesses at six days of public hearings plus an additional 40 at the San Francisco hearings on cross-cutting issues. Dr. Detmer commented that the effort to get out of Washington, D.C. and communicate with people in the West was very worthwhile.

#### **UPDATE FROM THE DEPARTMENT AND HHS DATA COUNCIL**

Mr. Scanlon reminded the Committee of the six major themes in the Council's work plan: health data standards, health information privacy, serving as the focal point for work with NCVHS, coordinating the Department's positions on ad hoc issues, developing a data collection strategy for the Department, and health information applications of the national information infrastructure (NII). He focused his remarks on the last two of these.

Developing and overseeing an HHS data collection strategy is a challenge because of the Department's size and diversity. The Council is

working on a multi-year plan for FY 1997 through 2001, covering survey content, periodicity and budget planning as well as doing a data "cross-cut" to identify gaps and redundancies. The survey integration plan, on which the Committee has been briefed several times, is part of this broader effort. Its major feature is using the National Health Interview Survey as the hub for other population surveys. NHANES will go into the field in 1998; the Medical Expenditure Panel Survey will be in the field most of the time. Other components of the data collection strategy are the HHS research planning initiative, strategies for state level data, and work on special populations and race/ethnicity data. An inventory of surveys and survey plans will be available on the HHS website. HHS is also reconceptualizing the area of data on providers, resources and capacities, with reference to the changing nature and structure of the health care system.

The research planning initiative began in HHS last year, associated closely with welfare reform, the chief question being whether the nation has the capacity to monitor the changes occurring through devolution and reform, particularly at the state level. Developing strategies for state level data is a key effort in the Department -- for example, for a state and local area integrated telephone survey system. The Data Council also has a workgroup on race and ethnicity data that has been reviewing all the previous recommendations. It is heading the Department's review of OMB's recommendations on Directive 15.

Regarding the NII, Mr. Scanlon noted that the Council is charged by the Vice President with leading interagency efforts on data standards, privacy, telemedicine, and consumer information. It formed an interagency working group on telemedicine which is looking at regulatory barriers, co-sponsoring a demonstration project, and developing an inventory of telemedicine efforts. In the area of consumer health information, HHS has developed a project called HealthFinder to help consumers get reliable health information over the Internet.

Asked about the evaluation of telemedicine, Mr. Scanlon said the Data Council's broad focus is to develop a methodology that can identify the boundary

and effectiveness of this technology and to compare its outcomes to other methods. Dr. Detmer observed that the country lacks, and could benefit from, a robust and comprehensive strategy for the health applications of the NII, which he suggested the Committee can help develop. He noted that the context is global, and the U.S. must see its role in terms of partnerships.

#### **HHS IMPLEMENTATION OF P.L. 104-191, PRIVACY COMPONENT**

The Secretary's recommendations for the protection of health privacy are being drafted and will be presented to Congress by August 21. The National Committee's recommendations will be fed into that process. In addition, the Department has commissioned a report on privacy and health research, which is available on the Data Council's website.

#### **DATA STANDARDS COMPONENT**

On June 11, the Data Council heard from the implementation teams that are developing regulations on the K2 standards. A public meeting will be held on July 9 at which the teams' co-chairs will present the information and elicit feedback about the Department's positions. The draft regulations will be ready by October, and the final rules in February, 1998. Dr. Braithwaite expressed appreciation for the Committee's hearings, which have provided important input into the Department's recommendations.

The crux of those recommendations is that the ANSI X12-N version 3070 standards be used for all but a few transactions in the claims area. Pharmacy claims will use NCPDP standards and dental claims, the 837. HCFA's National Provider Identifier and National Payer ID will be recommended as the proposed standards. In the coding area, the Department will recommend the continued use of ICD-9-CM, adding that systems should be ready to move to the ICD-10-CM by 2001. For procedures, the Department recommends the continued use of the ICD-9 Vol. 3 and CPT-4 and HCPCS, but with the goal of moving to an as yet undetermined single procedure coding system by 2002 or 2003. Recommended security standards will be policy (not technical) standards. Implementation guides will be available free on the Internet for all the

standards.

Discussion: Turning to the unique provider identifier, Dr. Detmer noted that the Department recommends that an 8-digit alpha-numeric identifier be assigned to all providers, with identifying information. It is proposed that NCVHS endorse this proposal and recommend that HHS publish it for public comment.

### **STATE-BASED STANDARDS AND PRIVACY INITIATIVES**

Elliot Stone (Massachusetts Health Data Consortium): All the relationships among stakeholders have been crafted by the MA Health Data Consortium, and they all have a vested interest in seeing that the national standards are implemented expeditiously. The Affiliated Networks do not advocate building centralized regional data repositories, but instead favor standardizing components of private data bases held by individual providers. A comprehensive health data system will be achievable once national standards are recommended by the Secretary. Walter Suarez (Minnesota Health Data Institute): Regarding privacy and confidentiality, he noted that Minnesota has stringent data privacy legislation and argued for a national legislative framework that states can build upon.

Other Discussion: Mr. Gellman asked the panelists to elaborate on their comments on federal privacy legislation and the issue of preemption versus allowing more stringent state laws to prevail. Dr. Suarez spoke in support of a more stringent federal standard rather than a lower common denominator, so that states are not inclined to push it higher. Mr. Rubin said that despite the obvious tradeoffs, he supports preemptive federal law that does not allow state-level customization because so much health care is delivered on a regional basis. Dr. Suarez noted that states can act independently in ways such as moving faster, or implementing more transactions, than federal law requires. Mr. Stone urged the Committee to consult with Massachusetts' Attorney General and his staff, a good resource, and noted that the most important issue has to do with employer access to information.

Returning to the subject of a single national privacy law versus multiple ones, Dr. McDonald asserted that it is contradictory to support standards and argue for individual state privacy laws. Mr. Rubin agreed, noting that many Washington health plans are becoming regional or national. The other panelists said they saw this trend in their states as well, along with the facts that most physicians belong to more than one plan and patients are expected to move from plan to plan. Mr. Rubin added that people are also starting to see the merits of standardization in the security area.

### **DISCUSSION OF PRIVACY RECOMMENDATIONS:**

Mr. Gellman presented the draft recommendations prepared by the Subcommittee on Privacy and Confidentiality, explaining that the document emphasizes the urgent need for federal legislation. It also takes positions on subjects about which the Subcommittee is in agreement (e.g., liberal access for research and public health) and identifies the areas where it is not (e.g., preemption and disclosures with written authorizations).

### **STATUS REPORT ON OMB DIRECTIVE 15 REVISION**

Dr. Detmer welcomed Katherine Wallman, Chief Statistician at OMB. In view of the fact that OMB had not yet released the recommendations of its interagency committee, she provided the Committee some background and a general idea of what they could expect. The administration has had standards for collecting race/ethnicity data since 1977. Particularly since the 1990 census, the standards have been criticized as not in tune with changes in the demographic composition of the population. In 1993, OMB undertook a comprehensive review of the standards in cooperation with other agencies that use and produce data, to culminate in recommendations for any modifications to be incorporated beginning with the next decennial census. Decisions must be made by fall 1997.

The agency has made an effort to go out to diverse constituencies and elicit their concerns and

suggestions, and also to ensure that extensive research was conducted. There were four public hearings and an extensive Federal Register notice with comment period in summer 1994, resulting in testimony from 100 people and some 800 letters. OMB published another major notice a year later that provided a synopsis of the issues raised and laid out research plans. The research involved three major studies: the CPS supplement, the National Content Survey, and the Race and Ethnicity Targeted Test. There also have been focused studies in specific areas such as health, education and civil rights.

The research results and public comments have been studied by the interagency committee, which has just issued a report and recommendations. This will be the subject of the Federal Register announcement to be released in early July, followed by a 60 day comment period. On the basis of these comments and further analysis, OMB plans to release its recommendations by mid-October. Ms. Wallman stressed that the standard is intended as a minimum set of categories for cross-government use; more detailed categories can be used as needed in specific contexts.

#### **RECOMMENDATIONS ON PRIVACY AND CONFIDENTIALITY**

Mr. Gellman presented the revised draft of the recommendations, prepared by the Privacy Subcommittee in response to the previous day's discussion. Its major points remain the same: that the nation is in a health privacy crisis, that privacy protection has eroded over the last 20 years, and that the issue urgently needs attention. It asks for legislation by the end of this Congress. He and Ms. Frawley described the changes in detail; in some, they strengthen some recommendations (e.g., on discrimination) and clarify various positions in the areas of disagreement. Dr. Amaro spoke in support of Dr. Harding's concerns about special protections, noting that members of less advantaged groups are at the greatest risk of loss of privacy as well as at high

risk for such sensitive conditions as substance abuse, STDs and HIV. She urged that the Committee offer language to move toward solutions. Mr. Gellman responded that despite the widespread awareness of this problem, no solution has been found. Several members observed that here the Committee is struggling with the result of social inequities, including the lack of universal health care. Dr. Detmer expressed agreement with these concerns but stressed the need to move forward with the overall recommendations and not let them founder on this issue. He noted that the full Committee, the Privacy Subcommittee, and those working on security, as well as parallel processes in the Department, will continue to struggle with these concerns. Non-discrimination policies are critical, given society's decision to limit access to basic health care. Dr. Detmer expressed agreement with these concerns but stressed the need to move forward with the overall recommendations and not let them founder on this issue. He noted that the full Committee, the Privacy Subcommittee, and those working on security, as well as parallel processes in the Department, will continue to struggle with these concerns. Non-discrimination policies are critical, given society's decision to limit access to basic health care.

**Net Site(s) of Interest for this Edition:** You may want to visit a very well organized instructional presentation on infectious disease epidemiology and related topics at <http://www.sph.umich.edu/~jkoopman/802Web/>. Start with Chap1.htm. The course was prepared by Jim Koopman MD MPH, Dept. of Epidemiology SPH-1, University of Michigan School of Public Health, 109 Observatory St., Ann Arbor, MI 48109. [Jim may be reached at e-mail [jkoopman@sph.umich.edu](mailto:jkoopman@sph.umich.edu) or voice (313) 763-5629]

**Final Thought(s)**

These are exciting times for GIS and public health. There are many opportunities to collaborate and work together on a wide range of public health issues involving GIS science. I am pleased to spotlight one such opportunity with the following communique received from GIS colleague **Susan Perlin**: “Dr. Susan Perlin is leading an effort at the Environmental Protection Agency’s National Center for Environmental Assessment to conduct a study on children’s respiratory health using the National Health and Nutrition Examination Survey (NHANES-3) data. As part of the study, we want to use GIS technology to overlay environmental data with the NHANES-3 data. We would like to discuss this type of study with researchers who have experience with using GIS to analyze health data in conjunction with environmental exposure, or exposure surrogate, data. We would be particularly interested in hearing from individuals who are involved in this type of study and are using NHANES data. We would be interested in possible collaboration, too. [Contact: Susan Perlin at e-mail [perlin.susan@epamail.epa.gov](mailto:perlin.susan@epamail.epa.gov) or voice 202-260-5877]

**Charles M. Croner**, Ph.D., Editor, ***PUBLIC HEALTH GIS NEWS AND INFORMATION***, Office of Research and Methodology, National Center for Health Statistics <[cmc2@cdc.gov](mailto:cmc2@cdc.gov)>. Copyright Notice: This report is in the public domain but its contents are not to be altered or changed without prior written approval of the editor.

**August 17-20, 1998 “GIS in Public Health” Conference...Plan Ahead Now**