PUBLIC HEALTH GIS NEWS AND INFORMATION

July 1999 (No. 29)

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

I. Public Health GIS (and related) Events

SPECIAL CDC/ATSDR GIS LECTURES: Tracking and Responding to Human Health Issues from Hurricane Mitch (Central America) with GIS, by James R. Jancaitis, Chief, Office of Systems and Technology, U S Geological Survey, **July 7**, 1999, 2:15-3:15 P.M., at the NCHS Auditorium, Hyattsville, MD (Envision available to offsite locations); and, The Use of a Spatial Statistics Package for GIS Analysis in Public Health Applications by Ned Levine, Ned Levine and Associates, Annandale, VA, **September 16**, 1999, 2:00-3:15 P.M., at the NCHS Auditorium, Hyattsville, MD (Envision is available to offsite CDC/ATSDR locations for these lectures). Abstracts are included in this edition. These talks are sponsored by the CDC/ATSDR Behavioral and Social Science Working Group, the CDC Statistical Advisory Group, and the NCHS Cartography and GIS Guest Lecture Series. [Contact: Chuck Croner at e-mail cmc2@cdc. gov]

****

L The 24th Annual Conference of the Association of Public Data Users, October 24-27, 1999, Alexandria, Virginia [See: http://www.apdu.org]

K Symposium on the Use and Potential of Spatial Data, Consortium for Geographic Information, November 10-12, 1999, Los Angeles, CA [See: www.drivenbydata. org]

L International Symposium on Digital Earth (ISDE), Chinese Academy of Sciences (CAS), November 29 - December 2,1999, Beijing, China [Contact: Chongjun Yang at e-mail cjyang@digitalearth.net.cn or http:// www.digitalearth.net.cn/de99.htm]

K The 11th Annual Colloquium of the Spatial Information Research Centre, "Our Safe LivingEnvironment: The Changing Face of Spatial Systems," December 13-15, 1999, University of Otago, Dunedin, New Zealand [See: http://divcom.otago.ac.nz/ sirc/ default.html]

L International Conference on the Analysis and Interpretation of Disease Clusters and Ecological Studies, jointly organized by the Department of Epidemiology and Public Health at Imperial College, and the Royal Statistical Society, December 16-18, 1999, London [Contact: Jon Wakefield, Imperial College School of Medicine at voice 0171 594 3318]

II. News from GIS USERS

(Please communicate directly with colleagues on any issues)

A. General News (and Training Opportunities)

1. From **Bill Henriques**, ATSDR: The GIS in Public Health conference held August 1998 in San Diego will be held next in 2001. This coming year we will begin to identify the location and get prepared for the next conference. We are currently working on the 1998 conference proceedings which should be on the web by late Summer 1999. [Contact: Bill, GIS Coordinator, at voice (404) 639-6088 or e-mail wdh2@cdc.gov]

2. From **Wayne Johnson**, NCHS CDC: MathSoft is offering S-PLUS Fundamentals, a four day training course, in Atlanta, Georgia, August 2-5, 1999. This class is designed for not only for those new to S-PLUS, but also for those who have some previous experience. Beginners can take this course to get up and running in S-PLUS, while experienced users can "fill in the gaps" of their knowledge. The course will expose users to

essential S-PLUS functions. This course covers the basics of S-PLUS for Windows with an emphasis on using S-PLUS to analyze data. It will cover the fundamentals of the Graphical User Interface (GUI) and the syntax of the S-PLUS command line. It will explore the tools for data analysis, including powerful and flexible graphics functions, as well as several of the most common statistical procedures. Upon completion of this course, every participant will have a strong foundation for becoming a successful S-PLUS user. [Contact: Cathie Lynch at voice 800-569-0123 ext. 246 or e-mail at cathie@statsci.com]; in addition, MathSoft announced a new release of S-PLUS for Windows-S-PLUS 2000. S-PLUS 2000 offers new and updated statistical techniques, powerful interactive visualization features, more data manipulation capabilities, and all new project management tools. The new release also delivers numerous ease-of-use enhancements and an updated Microsoft Office look and feel [see http://www.mathsoft.com/splus/splsprod/ splus2000pro.htm]

3. From **Greg Elmes**, West Virginia University: UCGIS Congressional Reception an Outstanding Triumph! A Governor, a representative, dozens of congressional staff, hundreds of federal and state officials and other VIPs crowded the National Geographic Society Tuesday evening to participate in the 1999 Congressional Reception. The evenings events and displays were an unmitigated success and represent a high point in the ever more impressive organizational outcomes of UCGIS and its members. I want to recognize the efforts of the 23 member Universities of UCGIS who voluntarily contributed their time and energy to provide the displays and presentations at the 1999 Congressional Reception, and were the backbone of the reception's success. The visual impact was outstanding and the work spoke very loudly about our contribution to livable communities. Let's build on this remarkable achievement as we enter our Summer Assembly later this month, and continue to build towards greater participation in next year's Congressional outreach. We should and can feel extremely proud of our consortium's visibility and accomplishments.

In following up on the reception, I encourage you to take every opportunity to talk to your legislative representatives at every level to spell out exactly why geographic information science is essential for our many communities. Persevere in you efforts to involve your representatives. Thank them for participating, remind them that we'll be inviting them again in 2000, and make sure they have the 'talking points' package. Dare I encourage you to become more involved in the democratic process? Explain what you are doing in your research and teaching, and why the kinds of research you are undertaking and are capable of accomplishing is crucial to meet our nation's geographic information science needs. I wish to extend my thanks to some very special people for Tuesday's success. The 1999 National Geodata Forum provided space for us in the program, and FGDC staff, Kathy Covert and John Moeller accommodated us very well. Thank you. Mr. Julian Haines, Dr. Harlan Onsrud worked very hard on all our behalf, and especially Mrs. Suzy Jampoler for her tireless efforts in making the evening a success.

Finally, please join me in taking any opportunity you have to thank Senator Collins of Maine, and our other honorary co-hosts, Senator Boxer of California and Representative Gutknect for Minnesota and the other members of government present at the UCGIS Congressional Reception, including Governor Geringer of Wyoming and Representative Kanjorski of Pennsylvania, in taking time out from their busy schedules to show their support for our initiatives. If your legislative representatives sent staff, please find a means to thank them by way of a follow-up. This is a process NOT an event! Again, thanks to all who made it work. [Contact: Greg, President UCGIS, at e-mail gelmes@wvu.edu]

4. From **Tom Usselman**, National Academy of Sciences: A new report on Distributed Geolibraries: Spatial Information Resources was released on Monday, June 7. The report was produced by a Panel of the Mapping Science Committee of the National Research Council and was based on a workshop held last summer. The Panel was chaired by Michael Goodchild and included Prue Adler, Barbara Buttenfield, Robert Kahn, Annette Krygiel, and Harlan Onsrud. From the Executive Summary: A distributed geolibrary is a vision for the future. It would permit users to quickly and easily obtain all existing information available about a place that is relevant to a defined need. It is modeled on the operations of a traditional library, updated to a digital networked world, and focused on something that has

never been possible in the traditional library: the supply of information in response to a geographically defined need. It would integrate the resources of the Internet and the World Wide Web into a simple mechanism for searching and retrieving information relevant to a wide range of problems, including natural disasters, emergencies, community planning, and environmental quality.

A geolibrary is a digital library filled with geoinformation--information associated with a distinct area or footprint on the Earth's surface--and for which the primary search mechanism is place. The development of distributed geolibraries and their services and functions should be a critical step in the evolution of the Digital Earth program espoused by Vice President Gore. Distributed geolibraries can also form a major cog in data access and distribution for the National Spatial Data Infrastructure. The report is online at http://www.nap.edu/html/geolibraries/. Limited copies of the report (119 pp.) are available from my office; once this complimentary supply is depleted, the report can be purchased from the National Academy Press http://books.nap.edu/catalog/ 9460.html). [Contact: Tom at e-mail usselman@nas. edu]

5. From **Deserene Worsley**, The National Capital Planning Commission: The Washington Geographic Information System (WGIS) Consortium of the National Capital Planning Commission (NCPC) is developing a metropolitan GIS database for coordinated planning and exchange. Focus group exercises were conducted with member agency representatives in May and will help define the basis of the WGIS Strategic Plan. Copies of the results will be posted on our web page at www.ncpc.gov. The next GIS application that uses WGIS Consortium data is being sponsored by the U.S. Attorney for the District of Columbia. Mr. Robert Crump, GIS Specialist for INDUS Corporation will discuss the Photomapper application. This application was developed by a WGIS partner and is just one of many examples of how the membership is currently using our orthophotography. [Contact: Deserene at voice (202) 482-7235 or e-mail des@ncpc.gov]

6. From **Bob Rugg**, Virginia Commonwealth University: The UCGIS Research Management Committee announces a new Request for Proposals under our agreement with the Federal Geographic Data Committee. The topic of the RFP is "Exploratory Studies of the NSGIC/FGDC Framework Survey". Deadline for proposals is July 30. Details can be found on the RMC website now at http://www.spatial.maine.edu/~max/ ucgis_rpc.html [Source: Bob Rugg, Chair, UCGIS Research Management Committee, at e-mail rugg@vcu.edu]

7. From **Duane Marble**, The Ohio State University (The UCGIS GIScience Initiative): Following action by the University Consortium for Geographic Information Science's (UCGIS) Education Committee and its Assembly in 1998, a Working Group has been established to develop a model, multi–path curriculum for GIScience. This Model Curriculum will identify several paths corresponding to those required by students who desire to become involved with GIScience at various levels and will also identify specific learning modules, levels of knowledge/competence, and the general prerequisites to be associated with each path. The Working Group is composed of academics from different disciplines and institutions as well industry representatives. A limited number of other representatives will be added to the Working Group as the need develops. The construction of the model curriculum is taking place in phases and, at critical points in the process, drafts of results will be made available for general public comment thru the UCGIS web site at www.ucgis.org. We anticipate that the final report from the Working Group will be available in about thirty months. [Contact: Duane, Chair, UCGIS Working Group, at the Center for Mapping, voice (614) 292-3409 or e-mail marble.1@osu.edu]

8. From **Ronald Abeles**, NIH: The NIH Behavioral and Social Sciences Research Coordinating Committee (BSSR-CC) maintains an electronic mailing list (listserv) for announcing activities of interest and for exchanging questions and information among NIH behavioral and social scientists. To receive announcements of the Behavioral and Social Sciences Interest Group, see http://www.nih.gov/sigs/bssrig/ [Contact: Ron, National Institute on Ageing, at e-mail AbelesR@gw.nia.nih.gov]

B. Technical and Research News

9. From **Daniel Exeter**, University of Auckland

(ArcView 3D Analyst):I have made extensive use of the 3D Analyst extension of ArcView to view disease data effectively. I am all for it. If you have point observation data, then use Spatial Analyst to create a density surface, rather than being constrained to the census tracts. If you want to see some of the examples of what I have done, visit the web survey that I did last year as part of my MA research. It shows you different visualization techniques that can be used to visualize data. The survey takes about 30 minutes to do, but you can just skim through the methods. There is a number of different ways of representing data as surfaces shown on the surface mapping page. The page is located at our site http://www.geog.auckland.ac.nz/cgi-bin/dan. At this stage, the only online version of these techniques is in the form of the survey. My thesis is available in PDF Format at ftp://ftp.geog.auckland.ac.nz/pub/outgoing/dxthesis.zip. [Contact: Dan at e-mail d.exeter@auckland.ac.nz]

10. From **David Moriarity**, NCCDPHP CDC: An interesting article on GIS in the April 26, 1999 edition of the LA Times- Mapping Finds Its Way to the Mainstream (http://www.latimes.com/home/business/cutting/lat_map990426.htm)-it includes news of Microsoft's soon-to-be-released "MapPoint 2000" "which, for just $109, includes a detailed map of the entire country along with census data about household size, income, ethnicity and even age. It's designed so users can easily find any address in the country, plop a map into a Microsoft Word document or transfer data to a map from a spreadsheet." [Contact: Dave at e-mail: dgm1@cdc.gov]

11. From **Gus Kubica**, Ogeta: Forum for Spatial Data Interoperability, previously Atlanta Metropolitan Forum for Open Spatial Information Systems (Originally Fulton County Open GIS Forum). Presentation by Bill Iler, Senior Director, Orbimage, Dulles, VA entitled “Satellite Imagery to Assist Future Orthophoto Data for Ogeta.” Location- Board Room of the Atlanta Regional Commission, 200 Northcreek, 3715 Northside Parkway, Atlanta, GA 30327-2809, September 27, 1999, 10:00-12:00. [Contact: Gus at e-mail gus.kubica @ogeta.com or for a map to the Atlanta Regional Commission please visit www. ogeta.com/forum.htm]

12. From **Paul Van Zuyle**, University of California Santa Barbara: I have an interest in whether geostatistical methods developed for mineral exploration/geology could be applied in a public health context to predict where cases of disease will occur. I plan to start with tuberculosis in Santa Barbara as a test case. Potentially, the model also could be applied to other diseases.

Background: The "Weights of Evidence" (WofE) technique is an application of Bayes Rule for calculating posterior probabilities (Good, 1950). Its use in medicine has been described in some detail by Speigelhalter and Knill-Jones (1984). A geographic application for mineral exploration was described by Bonham-Carter, et al. (1988), and generalized in a book (1994). The technic involves using a matrix of predictors for geographic areas (in this case, geologic properties) along with a set of training sites (here, gold mines) to predict the likelihood of gold occurrence in different regions of a map. Gary Raines of the USGS office in Reno presented a colloquium at UCSB where he described an ArcView Extension that implements this technique (Kemp, 1999). A detailed discussion can be viewed at: http://ntserv.gis.nrcan.gc.ca/wofe/.

It occurred to me while listening to his presentation that the same technique could be applied to the spatial prediction of disease occurrence. The analogy that came to mind was finding tuberculosis cases using only census data and previous occurrences. Unlike most other statistical techniques that have received attention recently, such as Kernel filtering described by Rushton, or the spatial scan described by Kulldorf, this is not a test of significance. The nature of the application is such that low probabilities can be reliably calculated, and expert knowledge can be added in specific cases.

The potential value in such a technique is for prospective, rather than retrospective analysis of disease occurrence. While the predictive power of available evidence might be low, in the absence of better information a little guidance is always better than none in the application of limited resources. While the payoff structure is different for early intervention or prevention of a disease case than for the discovery of a gold mine, the total benefit may be comparable.

My intent is to experiment with this technique on at least two datasets. The first is address-level tuberculosis incidence in Ventura County, California. As the GIS analyst for the Health Dept., I have access to

multiple years of CMR records. One potential design is to use 1997 cases along with census block group data to predict 1998 cases, and measure the utility of the result. The second dataset I have considered is the Snow Cholera map of London from 1854. My attraction is simply based on the extensive analysis this data has already received from John Snow, Andrew Cliff, Edward Tufte, Waldo Tobler and others.

The point of all this, however, is to test the WofE technique as implemented by Kemp and others to find out if it is useful for epidemiological investigation. The important properties are as follows: it is a probability measure rather than a significance test; it is already implemented in a GIS and thus could be quickly adopted if it proves useful; and it is designed to use both universally available data (for our purposes, Census data and disease incidence) as well as expert knowledge.

References: Bonham-Carter, G. F., et al. 1988. Integration of Geological Datasets for Gold Exploration in Nova Scotia. Photogrammetric Engineering and Remote Sensing Vol 54, No. 11, November, pp. 1585-1592; Bonham-Carter, G. F. 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. New York:Pergamon; Good, Isidore Jacob. 1950. Probability and the weighing of evidence. London:C. Griffin; Kemp, L.D., Bonham-Carter, G.F. and Raines, G.L., 1999. Arc-WofE: ArcView extension for weights of evidence mapping. http://gis.nrcan.gc.ca/software/arcview/wofe; Snow, J. 1936. Snow on Cholera: A reprint of two papers. New York:Oxford University Press, at site http://www.ncgia.ucsb.edu:80/pubs/snow/snow.html; and Spiegelhalter, D. J., Knill-Jones, R. P. 1984. Statistical and Knowledge-based Approaches to Clinical Decision-support Systems, with an Application in Gastroenterology. J. R. Statist.Soc. A. 147:35-77. [Contact: Paul at voice (805) 893-8652 or e-mail van zuyle@geog.ucsb.edu]

C. Internet News

13. From **Jimmie Givens**, NCHS CDC (forwarded-The FY2000 Blue Book is available on the web!): The FY2000 Supplement to the President's Budget: Information Technology Frontiers for a New Millennium is now available at http://www.ccic.gov/ pubs/blue00/. A highlight of the President's FY 2000 R&D budget is the proposed Information Technology for the Twenty-first Century (IT2) initiative. This research initiative proposes $366 million in increased investments to help advance the knowledge base in fundamental information science and to train the next generation of researchers who will sustain the Information Revolution well into the 21st Century. Building on the Government's previous accomplishments and existing investments in High Performance Computing and Communications (HPCC), including the Next Generation Internet (NGI), and the Department of Energy's (DOE's) Accelerated Strategic Computing Initiative (ASCI), the initiative will extend some existing research and provide opportunities to address new, complementary research topics in three key areas: 1) Long term information technology research that will lead to fundamental advances in computing and communications 2) Advanced computing infrastructure as a tool to facilitate scientific and engineering discoveries of national interest, and 3) Research on the economic and social implications of the Information Revolution, and efforts to help train additional IT workers at our universities. [Contact: Kay Howell, Director, National Coordination Office for Computing, Information, and Communications at voice (703) 306-4722 or email howell@ccic.gov]

14. From **David Mark**, University of Buffalo: As many of you know, the House Committee on Government Reform's Subcommittee on Government Management, Information, and Technology held hearings on "Geographical Information Systems Policies and Programs" on Wednesday, June 9, 1999, in conjunction with the National GeoData Forum. The testimony given at that hearing is now on the House web site: http://www.house.gov/reform/gmit/hearings/ testimony/990609h.htm. I believe that this hearing, which put GIS issues on the record in Congress, is yet another important step for our field. [Editor: I agree with David- please see the "Web Site of Interest" in this edition which expands on this notice; Contact: David, UCGIS Past President, at e-mail dmark@geog.buffalo.edu]

15. Editor: The CDC Washington (CDC/W) Office provides up-to-date, science-based information for public health policy and legislation. A quick CDC resource for Congressional staff and other Washington, DC-based organizations, CDC/W provides information about public health and CDC programs, links Hill staff with CDC experts, arranges briefings with CDC officials,

and provides background on a broad range of public health issues. The office also ensures that CDC staff receive information, advice, and guidance regarding Congressional activities that affect CDC. Address: Hubert H. Humphrey Building, 200 Independence Avenue, SW, Room 746G, Washington, DC 20201.

Resources and Publications- Listed below are examples of some of the information that CDC can provide to Congressional offices: Health Statistics, including vital statistics, morbidity and mortality data, disease rates and trends; State Health Profiles, booklets that include highlights of health status, prevention efforts, and CDC funding for each State; Health, United States, an annual report on the health of the Nation that includes state-specific data; MMWR (*Morbidity and Mortality Weekly Report*), for news on selected health issues and cases of selected diseases; Fact Sheets, on current and emerging health and safety issues; Subject-Matter Experts, for briefings and special or local events; and, International Travel Information, such as health precautions and vaccination requirements. [Contact: CDC/W at voice (202) 690-8598 or see http://www.cdc.gov/od/wash/ or]

D. CDC/ATSDR- Atlanta GIS News

16. Editor: I am informed by **Donna Stroup**, EPO, that the June 15, 1999 presentation "The Use of GIS to Guide Childhood Lead Poisoning Prevention Activities," by CDC/ATSDR colleagues **Dori Reissman** (Speaker), NCEH, **Virginia Lee** (Discussant), ATSDR and **Andy Dannenberg** (Moderator), EPO/DAPHT, was the first example of GIS used in CDC's Epidemiology (or Epi) Grand Rounds. As I watched from Hyattsville, I was reminded of Tipper Gore's challenge to all of us soon after she arrived in Washington, D.C., "There is no environmental problem that poses a more immediate threat to America's children than lead poisoning" (*The Washington Post*, May 5, 1993).

The purpose of Epi Grand Rounds is that CDC scientists, EIS Officers, and medical professionals need to be knowledgeable of current problems and issues facing public health today. The EPI Grand Rounds provides that forum for presentation and discussion of current epidemiologic investigations and new epidemiologic methods used, so that CDC can remain at the forefront of current public health issues. Congratulations to Dori, Virginia and Andy for making GIS history at Epi Grand Rounds. [For further EPO and Epi Grand Rounds program information visit http://www.cdc.gov/epo/whatsnew.htm; and, GIS presentation contact: Dori at dvs7@cdc.gov]]

17. From **Janet Heitgerd**, ATSDR: A meeting of the Atlanta CDC/ATSDR GIS Users Group is set for Tuesday, September 7, 1999 at Corporate Square, Building 11 from 1:00 -4:30 p.m. The agenda will include the following: demonstration of the Consequence Assessment Tool Set software by SAIC; discussion of trends in GIS and Public Health by ESRI; and information on the CDC/IRMO GIS initiative. At this meeting, we will also discuss the background and structure of the users group, invite participation by those attending, and allow plenty of time to address questions and concerns. If you have any questions feel free to call Jerry Curtis (NCEH) at 770-488-7262 or Janet Heitgerd (ATSDR) at 404-639-0602. More information will be forthcoming in the September edition of *Public Health GIS News and Information*.

E. CDC/ATSDR- Ft. Collins, Hyattsville, Morgantown and Others GIS News

18. Editor: "An Interactive Tool for Mapping NCHS U.S. Mortality Data," by CDC colleagues **Michael Mungiole** and **Sherry Weitzen** (Speakers), and **Jimmie Givens** (Moderator), NCHS, was presented on June 23, 1999 at Hyattsville. Abstract: Few opportunities have existed to examine the U. S. mortality database in an interactive mapping environment. To this end a mapping tool is being developed which facilitates examination of this geo-referenced (e.g., county) database. In this demonstration a graphical user interface for the analysis of mortality data will be presented. Functionality that exists within the ArcView GIS and mapping software will be shown as it is linked to the interface. The mortality mapping point and click interface permits users to create, visually display and print (age-adjusted, smoothed age-specific, and significance) mortality maps similar to ones available in the NCHS publication *Atlas of United States Mortality*. Mortality rate maps are aggregated to the health service area (HSA) level. In addition, a limited set of correlate variables that may be categorized as demographic, socioeconomic, education, or risk factor-related are incorporated and can be used when generating maps

using this analysis tool. One can examine geographic patterns among the various mortality causes and correlate variables. These variables are aggregated to the county, HSA, or state level. The ability to view mortality and correlate variable maps simultaneously allows researchers and analysts to develop hypotheses related to one or more of the 18 leading causes of death. [Contact: Mike at mim4@cdc.gov]

19. Editor: The NIOSH *Atlas of Respiratory Disease Mortality, United States: 1982-1993*, NIOSH publication 98-157, is now on the Internet. The URL is www.cdc.gov/niosh/98-157pd.html. If anyone is interested in a printed copy please contact: FAX (513) 533-8573, or Automated Router System (1-800-356-4674), or voice outside the U.S. 1-513-533-8471, or e-mail Pubstaft@cdc.gov. [Source: Author **Jay Kim**, NIOSH CDC at e-mail jhk0@cdc.gov]

III. GIS Outreach

(Editor: All requests for Public Health GIS User Group assistance are welcome; 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)

F From **Barbara Tempalski**, University of Washington: I am trying to take an inventory of agencies that utilize GIS as a tool for disease surveillance purposes. I want to assess what type of framework these agencies have in place for use of surveillance information for developing disease control polices and doing policy evaluation. My theory is, many use GIS as a tool for disease surveillance but do not have a framework in place for developing disease control polices and policy evaluation based on data coming out of the surveillance program. Any information is considered useful. [Contact Barbara at the Department of Geography, voice (206) 543-6388 or e-mail bjtemp@u.washington.edu]

? From **Barbara Rice**, National Immunization Program, CDC: Do you know a mapping program that also does animation? Early responses: 1. **Lee DeCola**, USGS: Barbara, my limited experience with map animation is some Mathematica images I put on the web using a GIF animation program, to demonstrate the growth of the Baltimore Washington region. See my homepage at http://home.earthlink.net/~ldecola/; a person who might have some ideas is Janet Tilley here, who is creating historical datasets of urban areas. Please keep me posted on what you find, as the subject interests us! and 2. **Daniel Exeter**, University of Auckland: I used an animated GIF in my web survey that I did as part of my thesis. To construct this, I was making the maps in ArcView, and exporting the View as a Postscript or bitmap file. I then used image alchemy to convert the 50ish images to GIFs. However, I did not like the idea of no user control. We are currently working on a census area-based animation. To do this, I made the field names sequential (Month1.. monthN), I used Avenue again (ArcView scripting language), and loaded a saved legend for each time period. I create an A5 layout, and then add text, the legend, the frame, the sponsor's logos and a time line, coded to season. Each map is created as a Windows Bitmap and then all images are converted to an AVI file. This means that basically anybody running Windows95/98/NT can view the result. They can also have the option of stopping and starting at leisure, and can go back to particular frames too. I know that the loading of a saved legend is tedious, but I am working on that. [Contact Barbara at voice (404) 639-8236 or e-mail ber2@cdc.gov]

? From **Anita Farel**, UNC School of Public Health: I am the Principal Investigator for a distance training course funded by HRSA [Health Resources and Services Administration]. The course currently has participants from state and local Title V programs. Through a series of six modules (EpiInfo, cost analysis, qualitative data collection, social inequities in health, quantitative data, and GIS),the course addresses the need to improve the capacity for data collection and use among public health professionals. For the GIS module, we are attempting to give the participants resources in their states for follow up or future development of GIS interests or capacity. I am trying to find out whether anyone has a listing of state GIS contacts, particularly those in public health. [Contact Anita at Department of Maternal and Child Health, voice (919) 966-5983 or e-mail Anita_Farel@unc.edu]

F From **Bruce Brown**, Visiting Fellow, NCID CDC,

Office of Global Health: I am currently working with Dr LeDuc in NCID CDC on developing GIS surveillance systems. We are interested particularly in South America. As I am just beginning to familiarize myself with GIS I was wondering if you could suggest any resources or contacts to discuss our interests. [Contact: Bruce at bqb5@cdc.gov]

? **Charity Tan**, Department of Health, Manila, Philippines (through health-GIS@who.ch): We are trying to develop a standard coding system for sectoral indicators (e.g. lands, utilities and socio-economic) to be used in GIS applications. While other sectors have already standardized codes for their indicators, we (in the social sector) are having problems in standardizing socio-economic indicators. Can it be done? Has any country tried it successfully? [Contact: Charity at e-mail chat@doh.gov.ph]

IV. Special Reports

(Submissions are open to all)

Inaugural issue of *Environmental Health Educators Quarterly* (EHEQ): EHEQ is an e-mail message about resources for students, educators, and the public on environmental health, cancer prevention and bioscience. This electronic publication is sponsored by the Center for Research on Environmental Disease, an arm of the University of Texas M. D. Anderson Cancer Center, Science Park, Research Division, Smithville, Texas 78957. The Center is located near Austin, Texas, and is funded by the National Institute of Environmental Health Sciences, a component of the federal government. For a free subscription send an e-mail message to coep@sprd1.mdacc.tmc.edu. This issue's news contains: 1) Glossary for the Human Genome Project (Secondary Education), 2) NIEHS Kids' Site (Elementary Education), 3) Molecular Biology Program (Secondary Education), and 4) Cancer Clusters (Secondary Education).

Glossary for the Human Genome Project: There is a newly established glossary of genetic terms established online for anyone wishing to learn definitions in the field. The Universal Resource Locator (URL) is: http://www.nhgri.nih.gov/DIR/VIP/ Glossary/. The web site has the capability to convey the definitions in three ways: text based narrative, an audio file with a voice explaining the term and a visual image of the chemical structure or similar diagram of the term. The web site is sponsored by the National Human Genome Research Institute, an arm of the NIH and the federal government.

Cancer Clusters -- The New Yorker Magazine has an article in the February 8, 1999 issue (p. 34-37), that is a good education tool and reference on why cancer clusters are so difficult to prove. The article titled The "Cancer-Cluster Myth" by Atul Gawande talks about the science of environmentally caused cancer groupings and explains some of the reasons behind the fact that so few suspected clusters turn out to be proven cases of exposure to environmental carcinogens. Other disease clusters that are often shown to HAVE a causal relationship to disease- such as infectious disease clusters-seem to increase the pressure to prove similar relationships between local environmental causes and cancer. One of the primary issues in suspected cancer clusters is the fact that they are often cancer cases from several of the over 100 different cancers. These different cancers will usually have a mix of known and unknown causes. The article also mentions the "multiple hit" theory that many cancer researchers have come to accept.

The Center for Research on Environmental Disease has over 40 scientists studying the interaction among genetic makeup of the individual, environmental exposures and age. The scientists are in three locations: The University of Texas at Austin--College of Pharmacy; the M.D. Anderson Cancer Center in Houston and at the Science Park Research Division of M. D. Anderson near Smithville, TX. The resource suggestions that will make up this and future newsletters will be classified as oriented toward primary, secondary education or both. In some cases the resources cited will be cross-referenced to the Texas Essential Knowledge and Skills activities. We welcome comments and feedback. [Source and contact: Don Cook, University of Texas M. D. Anderson Cancer Center at voice (512) 237-9404 or e-mail cred@sprd1.mdacc.tmc.edu]

****

The January-February 1999 issue of CDC's journal, *Emerging Infectious Diseases* (EID), is now available at http://www.cdc.gov/eid. An abbreviated table of contents is listed below. For more information about receiving and contributing to CDC's journal, send an e-mail to the EID Help mailbox (eidhelp@cdc.gov).

Emerging Viral Diseases in Australia; Economic Impact of *S. aureus* in NY Hospitals; Antibiotic Resistance in Developing Countries; *C. jejuni*-An Emerging Foodborne Pathogen; Comparative Genomics and Host Resistance; Cyclospora; Monoclonal Antibodies and Mucosal Transmission of Diseases; Dual and Recombinant Infections and HIV in Brazil; Diversity and Distribution of Hantaviruses in N. America; Climatic and Environmental Patterns Associated with HPS; Rationale, Potential, and Methods of Hantavirus Reservoir Studies; Long-Term Hantavirus Persistence in Arizona Rodents; Longitudinal Study of SNV in Rodents, SE Arizona; Spatial and Temporal Patterns in Rodent Populations; Natural History of SNV in Western Colorado; A Synthesis of Hantavirus Reservoir Studies; Proficiency of Laboratories to Detect VRE; *S. aureus* with Reduced Susceptibility to Vancomycin; Candida dubliniensis Candidemia; Household Transmission of *S.pneumoniae*; Zoonotic Diseases in Immunocompromised Persons; *M. penetrans* Bacteremia and Primary Antiphospholipid Syndrome; Infectious Diarrhea in Tourists at a Resort Hotel; and, Midcourse Assessment of HPS.

V. GIS and Related Presentations and Literature

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

NCHS Cartography and GIS Guest Lecture Series

Tracking and Responding to Human Health Issues from Hurricane Mitch (Central America) with GIS: To be presented by James R. Jancaitis, Chief, Office of Systems and Technology, U S Geological Survey, **July 7**, 1999, 2:15-3:15 P.M., at the **NCHS Auditorium**, Hyattsville, MD (Envision available to offsite locations). ABSTRACT: Hurricane Mitch was, from an economic perspective, **the most destructive hurricane in the history of the western hemisphere**. When the Honduran government requested mapping and technical aid from the U.S. government, quickly followed by similar requests from the White House Reconstruction Task Force (WHRTC), the U. S. Geological Survey built an on-going research and development effort to address these needs in the USGS Center for Integration of Natural Disaster Information, CINDI. CINDI is a research laboratory focusing on advanced disaster information integration and delivery technology. An important part of the CINDI implementation strategy is the use of a public/private partnership with support from and involvement of multiple government and private organizations. Government partners include DOI, NOAA, and DoD COE. Private sector partners include ESRI, LizardTech, Microsoft, Hewlet Packard, and Motorola. The WHRTC includes the Department of State, USAID, HUD, DoD, and DoA. The CINDI partnership has been asked and is focused on finding, integrating and making available key map and map related scientific data important to disaster relief and reconstruction activities. Human health data and factors such as water quality were a key data element.

Federal agencies, in partnership with the private sector, are providing affected Central American countries a wide range of tools and technical assistance to aid in the reconstruction process. This partnership, coordinated by the U. S. Geological Survey, has provided numerous products and technical service and assistance items including: an atlas, conventional and satellite image maps, aerial photography (from the Air Force Open Skies program), geographic information systems CD-ROM data bases, computer software, advanced computer systems, technical assistance, and training. USGS scientists are providing basic mapping data, data integration and distribution, as well as transfer of geographic, hydrologic, geologic, biologic and information science. Human health has been one of the top concerns since the advent of this disaster. Readers are encouraged to visit the web site cindi.usgs.gov where an ftp site allows downloading of CD-ROM data- login ftp with the name "anonymous" and your e-mail address as the password. In addition, users can make maps from the "Interactive Disaster Atlas of Central America" link on the CINDI web page through the generous donation by ESRI of ArcView IMS software. [Contact: Jim at voice (703) 648-4505 0r e-mail jrjancaitis@usgs.gov]

The Use of a Spatial Statistics Package for GIS Analysis in Public Health Applications: To be presented by Ned Levine, Ned Levine and Associates, Annandale, VA, **September 16**, 1999, 2:00-3:15 P.M., at the **NCHS Auditorium**, Hyattsville, MD (Envision available to offsite locations). Abstract: *CrimeStat* is a spatial statistics package that was developed for the National Institute of Justice (NIJ) to aid in crime mapping applications. The program is a public domain stand-alone Windows NT/9x program that can interface with

most desktop GIS packages. It can read 'dbf' files and can write various graphical objects to ArcView, MapInfo, Atlas*GIS, Surfer, and Spatial Analyst. It has a collection of statistical tools for the analysis of point/incident location (i.e., having single X/Y coordinates) and includes a range of diagnostic spatial statistics. Some of the routines are specific to crime analysis, but many could be applied to epidemiological and public health problems. The program will be illustrated with motor vehicle accident data and its applicability to related public health topics will be discussed. The program began as a series of Unix routines developed to analyze motor vehicle crashes in Hawaii. NIJ got interested in the routines and wanted development of a version for crime analysis. In Hawaii, Dr. Levine developed a spatial analytic framework for analyzing crashes over an entire metropolitan area. The routines were developed to describe different crash distributions. He then took the logic farther and began spatial modeling as a way for identifying areas with higher-than-expected crashes, and for predicting future crash levels. He also will talk about the spatial modeling tools. The final version of CrimeStat and the documentation file will be available by September for download from the web. [Contact:Ned@NedLevine.com or http://www.ned levine.com/]

Snow's Cholera Map

The significance of Snow's famous cholera map is that, by closing the Broad Street pump by removing its handle, Dr. Snow stopped a major cholera epidemic, and thus demonstrated that cholera is a water borne disease. This was not previously understood. The map is the most famous and classical example in the field of medical cartography. For more details an informative source is: A. Cliff & P. Haggett, 1988, Atlas of Disease Distributions, Blackwell, Oxford, ISBN 0-631-13149-3.

In 1992, as part of the development work for an NCGIA technical report, Rusty Dodson of NCGIA Santa Barbara, digitized details from Snow's map reproduced in: "Snow on Cholera: being a reprint of two papers by John Snow, M.D., together with a Biographical Memoir by B.W. Richardson, M.D. and an Introduction by Wade Hampton Frost, M.D.", London, Oxford University Press, 1936. The scale of the source map is approx. 1:2000. Coordinate units are meters. The data in these files consists of: the relevant 1854 London streets ("streets"); the location of 578 deaths from cholera ("deaths"); and, the position of 13 water pumps ("pumps").

Each coordinate point in the file "deaths" specifies the address of a person who died from cholera. When many points are associated with a single street address, they are "stacked" in a line away from the street so that they are more easily visualized. This is how they are displayed on John Snow's original map. The dates of the deaths are not recorded. The data files were created for a student exercise included in NCGIA Technical Report 93-5: Teaching Introductory Geographical Data Analysis with GIS: A Laboratory Guide for an Integrated Spacestat/Idrisi Environment, edited by Rusty Dodson, preface by Luc Anselin. Death coordinate order was randomized and Thiessen polygon boundaries added by Waldo Tobler.

A program to display this data (cholera.exe) was written by Waldo Tobler on a 486 class PC, with a VGA display. It may not work quite correctly on other systems, but can easily be modified. [Hint: if you are using Windows 3 or Windows 95, start the program inside a DOS window.] It is also possible to change the colors, or to use dots instead of lines for the street map, or to slow the (random) time interval between the display of the individual deaths, etc. See the source code in "cholera.bas". In order to get printed copies of the map(s) it is necessary to use a graphics screen capture program. The coordinate data are also suitable for analytical investigations; e.g., bivariate density estimation, etc.

The program in this data set (cholera.exe) makes a map using John Snow's 1854 Cholera data. It will draw a street map, display the location of the deaths, followed by the location of the wells. Run the program by just typing cholera. The program and the data files must all be in the same directory. When you've seen the map press "enter" again, and again, ...[Source: Prof. Waldo Tobler, National Center for Geographic Information and Analysis Geography Department, University of California Santa Barbara at e-mail tobler@geog.ucsb.ed; also, see http://www.ncgia.ucsb.edu:80/pubs/snow/snow .html to download files]

VI. Related Census, DHHS and Other Federal Developments

The Biomedical Information Science and Technology Initiative, prepared by the Working Group on Biomedical Computing (Executive Summary), Advisory Committee to the Director, National Institutes of Health, June 3, 1999. **CHARGE TO THE WORKING GROUP ON BIOMEDICAL COMPUTING**: The biomedical community is increasingly taking advantage of the power of computing, both to manage and analyze data, and to model biological processes. The working group should investigate the needs of NIH-supported investigators for computing resources, including hardware, software, networking, algorithms, and training. It should take into account efforts to create a national information infrastructure, and look at working with other agencies (particularly NSF and DOE) to ensure that the research needs of the NIH-funded community are met. It should also investigate the impediments biologists face in utilizing high-end computing, such as a paucity of researchers with cross-disciplinary skills. The panel should consider both today's unmet needs and the growing requirements over the next five years (a reasonable horizon for extrapolating the advances in the rapidly changing fields of computing and computational biology). The result of deliberations should be a report to the NIH Director, which will be presented to the Advisory Committee to the Director. The report should include recommendations for NIH actions to support the growing needs of NIH-funded investigators for biomedical computing.

EXECUTIVE SUMMARY

In science and technology in the latter half of the 20th century, two fields have stood out for their speed of progress and their effect on society: biomedicine and computation. The charge of this Working Group is to assess the challenges and opportunities presented to the National Institutes of Health by the convergence of those two disciplines. The principal obstacle impeding effective health care is lack of new knowledge, and the principal mission of the NIH is to overcome this obstacle. At this point the impact of computer technology is so extensive it is no longer possible to think about that mission without computers.

Increasingly, researchers spend less time in their "wet labs" gathering data and more time on computation. As a consequence, more researchers find themselves working in teams to harness the new technologies. A broad segment of the biomedical research community perceives a shortfall of suitably educated people who are competent to support those teams. The problem is not just a shortage of computationally sophisticated associates, however. What is needed is a higher level of competence in mathematics and computer science among biologists themselves. While that trend will surely come of its own, it is in the interest of the NIH to accelerate the process. Digital methodologies- not just digital technology- are the hallmark of tomorrow's biomedicine. The NIH therefore must find ways to discover, encourage, train, and support the new kinds of scientists needed for tomorrow's science.

To make optimal use of information technology, biomedical researchers need, first of all, the expertise to marry information technology to biology in a productive way. New hardware and software will be needed, together with deepened support and collaboration from experts in allied fields. Inevitably, those needs will grow as biology moves increasingly from a bench-based to a computer-based science, as models replace some experiments and complement others, as lone researchers are supplemented by interdisciplinary teams. The overarching need is for an intellectual fusion of biomedicine and information technology.

Invariably, scientists learn best by doing rigorous science. Indeed, the NIH mission is to do science, including teaching and learning. Socially meritorious goals of improving human health and preventing, detecting, diagnosing, and treating disease and disability are achieved most effectively when pursued within the overall context of rigorous science. This report and its recommendations focus, therefore, on science - both for its insights and as a path toward building an educated interdisciplinary workforce. The centerpiece of our recommendations is the proposal to inaugurate National Programs of Excellence in Biomedical Computing. It is in the context of those National Programs that the best opportunities can be created for doing and learning at the interfaces among biology, mathematics, and computation. With such new and innovative programs in place, scientists will absorb biomedical computing in due course, while supporting the mission of the NIH.

Recommendation #1: The NIH should establish between five and twenty National Programs of Excellence in Biomedical Computing devoted to all facets of this emerging discipline, from the basic

research to the tools to do the work. It is the expectation that those National Programs will play a major role in educating biomedical-computation researchers. National Programs of Excellence in Biomedical Computing would advance research in particular areas of biomedicine, focusing on those in which computation is becoming increasingly essential. They would be funded in part through a new program, and in part through research grants from one or more of the Institutes that make up the NIH. The academic or research institutions at which the National Programs would be housed would be expected to contribute to the programs - and teaching would be an essential contribution.

National Programs could range in size. At a modest level, three to five researchers in complementary disciplines might receive $1.5 million a year to undertake the exploration of a single problem. Larger National Programs might bring together several problems and several technologies, perhaps in association with more than one institution or Institute, for up to $8 million a year. The NIH will determine the number and scope based on the applications and the grant process. One important goal of the National Programs will be to develop and integrate the use of computational tools to meet the important challenges of biomedical research. These Programs are in keeping with the conclusions of the President's Information Technology Advisory Committee (PITAC) report in that it focuses on basic information technology research in the pursuit of insight into the issues facing biomedical research. Concurrently, the National Programs will create homes for interdisciplinary teams, and those teams will establish nurturing environments for exploration and education. In establishing National Programs, the NIH will send a powerful message, both in academe and within the NIH community itself, about the importance of computation and the value of interdisciplinary research.

Strong action by the NIH is required because the existing biomedical research and teaching structures of the universities and research institutions of this country inadequately value interdisciplinary efforts generally, and computation in particular. Few grant programs and fewer academic departments foster the kind of interdisciplinary work required to meet biomedical challenges, let alone educate students about them. National Programs specifically would include formal and informal instruction from the undergraduate through post-graduate levels, and incorporate a range of opportunities for scholars and researchers to participate.

Recommendation #2: To make the growing body of biological data available in a form suitable for study and use, the NIH should establish a new program directed toward the principles and practice of information storage, curation, analysis, and retrieval (ISCAR). The information that biomedical researchers are amassing in profuse quantities today-from the Human Genome Project, clinical trials, statistics, population genetics, and imaging and visualization research - creates enormous digital repositories of information. The scale of those databases swamps all the information collected before. They encompass multigigabyte, multivariate functions of three dimensions, of wavelength, and of time. The infrastructure needed to make them available is phenomenal: A single biomedical laboratory could produce up to 100 terabytes of information a year - about the same as the information in one million encyclopedias. In order to be useful, the data must be indexed and stored, and the challenges for data analysis and abstraction are formidable.

The creation and development of advanced databases and database technologies (methods for storing, retrieving, and analyzing biomedical data) is becoming more important in all biomedical fields. The emerging technology of bioinformatics helps researchers gather and standardize data from basic research and computer modeling, and combine and manipulate databases to tease out the knowledge they contain. The goal is a system of interoperable databases that will make available the fruits of the increased productivity enabled by computation.

That is particularly true in clinical research: As more information from clinical trials becomes available, the need for standardization and interoperability of clinical databases will increase dramatically. Coordinating knowledge gained from clinical trial data with new insights from genetic research could appreciably advance knowledge about the treatment of disease. A system of interoperable databases would allow clinical researchers to track any finding back to its basic science roots; conversely, a research scientist might track forward to postulate from hypotheses through potential applications based on innovative uses of existing data. As the amount of data grows, the tools to compare and manipulate the data become more

important. These tools form software bridges between databases that will allow researchers to link disparate information sources.

The NIH has been a leader in establishing databases of valuable information and making them available for study. Now it must organize and expand database resources internally and externally. Currently the agency uses contracts, grants, and cooperative agreements in bioinformatics, but no program focuses specifically on database development. Both the collection of the information, and the creation of the tools for storage, management, and access are increasingly important. Therefore, the NIH needs a program that will rally new and important bioinformatics efforts and build this vital part of the biomedical infrastructure.

Recommendation #3: The NIH should provide additional resources and incentives for basic research (through R01 grants) to provide adequate support for those who are inventing, refining, and applying the tools of biomedical computing. Biomedical scientists know best what they need, and they often need to take advantage of computational opportunities. However, in evaluating research grants and programs, reviewers and staff sometimes have been reluctant to provide support for computation and computational infrastructure at the level required. The computational infrastructure, of course, includes not only the hardware but also the people with the expertise to make good use of the hardware. It is time for the NIH to recognize the importance of both the tools and those who build them. In order to do that, the NIH needs to ensure that R01 grants may be used for biomedical computation. That is particularly important for grants that support environments rich in teaching potential as well as research excellence. Researchers who work with students should have the resources that will allow them to set an example of the use of biomedical computing. As with any special emphasis or targeted funding, evaluation at three years is recommended.

Recommendation #4: The NIH should foster a scalable national computer infrastructure. To assure that biomedical researchers can gain access to the computing resources they need beyond their desktops, the NIH should provide financial resources to increase computing capacity, both local and remote. The purpose of this recommendation is to establish a balanced infrastructure for all computational needs.

Biology is becoming increasingly complex and computation is becoming increasingly sophisticated. Today's biomedical computing needs resources that go beyond desktop computers to local clusters of processors, to mid-level facilities, and to the most powerful computers at national centers. Many biomedical researchers cannot do their work on their desktop computers alone. They need varying amounts of computing power at different times, and those resources should be made available. The infrastructure must be better balanced for a dynamic range of computational needs.

Powerful computers alone are not enough. The entire support system must be in place. Even researchers who can do their work on small clusters need access to the expertise to set up and manage those clusters, and need support from programmers who can write or adapt the necessary software. As the computing-power needs increase, so do the support needs.

The NIH should support facilities with mid-level computers where new algorithms and applications can be developed specifically for biological problems. The biomedical expertise at those facilities would support researchers seeking to adapt and apply the best computer technology to their work. For some applications, mid-level facilities could offer smaller versions of scalable systems that exist at the national supercomputer centers. Researchers might use those resources to test and develop code or design before moving to national supercomputer centers, or - in appropriate cases - to do their work on more powerful computational resources than they have in their laboratories. Mid-level facilities could be created through National Programs that focus on supercomputing science, or the resources could be made available through cooperative agreements with existing extramural centers as well as at intramural centers.

NIH scientists have long taken advantage of the national supercomputer centers run by the National Science Foundation and the Department of Energy for high-level computing. Because the number of biomedical researchers who can profit from using those facilities is increasing, the NIH should take a strong leadership position and help support the national supercomputer centers. Such NIH support would provide a welcome opportunity for a partnership

between NSF and the NIH as the future of science unfolds in the 21st century.

CONCLUSION: The NIH can make a powerful contribution to the development of tomorrow's biomedical research community by increasing efforts to promote and support computational biology today. With the appropriate support in place, interdisciplinary research teams will coalesce for National Programs of Excellence in Biomedical Computing and ISCAR efforts. The natural byproduct of their emphasis on biomedical research will be a new generation of researchers who are skillful with computing, and who will have helped to create the computational tools they need to meet tomorrow's challenges.

As biomedical research becomes more computationally intensive, the Biomedical Information Science and Technology Initiative (BISTI) is essential if the NIH is to fulfill its mandate. This Initiative will be the means by which new techniques are developed, new knowledge is discovered, new research communities are created, and new ideas are disseminated to the institutions and people who can use them to solve the mysteries of life and health. [Editor: Thanks to **Charlie Rothwell** and **Jimmie Givens**, NCHS, for bringing this to our attention. It has applicability to all agencies. The complete report can be found at site http://www.nih.gov/welcome/director/ 060399.htm]

****

NCI Awards Contract for Geographic Information System To Support Breast Cancer Research on Long Island: The National Cancer Institute (NCI) announced today the award of a contract to AverStar, Inc., of Vienna, Va., to develop and implement a prototype geographic information system (GIS) for breast cancer studies as part of the Long Island Breast Cancer Study Project (LIBCSP). The contract award is for $4,872,309 for Phase 1 (two years) and Phase 2 (three option years). Phase I is to develop and deliver the GIS system, and Phase 2 is for system maintenance and data expansion to respond to research needs. "The Long Island geographic information system provides the opportunity to apply a powerful emerging technology to the study of environmental causes of breast cancer," said G. Iris Obrams, M.D., Ph.D., associate director of NCI's Epidemiology and Genetics Research Program (EGRP), Division of Cancer Control and Population Sciences (DCCPS). "This prototype GIS will be the first such system developed to study relationships between environmental exposures and breast cancer, and will provide researchers a new tool with which to conduct their investigations."

GISs are powerful computer systems that can store, manipulate, analyze, and display the spatial (geographic location) relationships between dissimilar data types. They provide a tool to study potential relationships between the location of breast cancer cases and sources of air or water pollution. The systems have been in development for more than 20 years, but only with recent advances in computer hardware, software, and increased availability of geo-coded health and environmental data (data tagged to location on the earth's surface) have they come into use in public health protection.

The LIBCSP is a multistudy effort to investigate whether environmental factors are responsible for breast cancer in Suffolk, Nassau, and Schoharie counties, N.Y., and in Tolland County, Conn. The investigation began in 1993 under Public Law 103-43 and is funded and coordinated by the NCI, in collaboration with the National Institute of Environmental Health Sciences (NIEHS). The Public Law includes development of a GIS for Long Island.

The GIS data layers will include geographic data for general mapping purposes and demographic data. Data on health care facilities, health care surveys, breast cancer, and the environment will also be included. The environmental data will include information on contaminated drinking water; sources of indoor and ambient air pollution, including emissions from aircraft; electromagnetic fields; pesticides and other toxic chemicals; hazardous and municipal waste; and radiation. The system will rely chiefly on existing databases obtained from federal, state, and local governments, and private sources, with emphasis placed on high-quality data. The Long Island community will also be asked to provide descriptive information about the environment and history of the area.

"The geographic information system for Long Island will be modular, flexible, and expandable so that it can be adapted to research needs. As additional exposure data become available, they can be added to allow researchers to explore important exposure-disease relationships," said Ellen Heineman, Ph.D., NCI project

officer for the contract. "The quality of the data is crucial so that investigators and the public can have confidence in research findings that are based on the system."

Although some of the data to be included in the GIS are publicly available, other data are confidential or proprietary, such as medical records. As a result, various levels of access to the GIS will be established to safeguard data while maximizing the system's usefulness as a research tool. This will ensure that individual records remain confidential. A Web site will be available through which the public and researchers will be able to obtain information on the progress of the GIS's development and summary information about the databases.

Use and access to the system will be guided by an oversight committee. The membership will be made up of members of the established Ad Hoc Advisory Committee to the LIBCSP, which includes scientists and community representatives; as well as federal, state, and local government specialists; external consultants; and additional representation from the Long Island community. The committee will advise NCI on policies and procedures, review the system's operations, and review and approve research proposals from investigators to use the GIS.

Questions and Answers on NCI's Contract Award for Geographic Information System To Support Breast Cancer Research on Long Island:

1. What is the Long Island Breast Cancer Study Project (LIBCSP)? The LIBCSP is a multistudy investigation of possible environmental causes of breast cancer in Suffolk, Nassau, and Schoharie counties, N.Y., and in Tolland County, Conn. The investigation began in 1993 under Public Law 103-43 and is funded and coordinated by the National Cancer Institute (NCI), in collaboration with the National Institute of Environmental Health Sciences (NIEHS). The LIBCSP is part of the overall research approach of the National Institutes of Health (NIH), of which NCI and NIEHS are a part, to investigate the causes of breast cancer and find ways to prevent it.

2. What does Public Law 103-43 say? The Public Law enacted in 1993 directed that a study be conducted of "potential environmental and other risks contributing to the incidence of breast cancer" in Nassau, Suffolk, and Schoharie counties, N.Y., and in Tolland County, Conn. Further, it directed that the study on Long Island "should include the use of a geographic system to evaluate the current and past exposure of individuals, including direct monitoring and cumulative estimates of exposure to (1) contaminated drinking water; (2) sources of indoor and ambient air pollution, including emissions from aircraft; (3) electromagnetic fields; (4) pesticides and other toxic chemicals; (5) hazardous and municipal waste; and (6) such other factors as the director [of NCI] determines to be appropriate."

3. What constitutes the Long Island geographic information system (GIS) for breast cancer studies? The GIS will consist of hardware, software, and data tables. The computer hardware will be located in AverStar's offices in Vienna, Va., and when the system becomes operational, authorized users will be able to access the public data tables from offices or laboratories. Commercially available software programs developed for GIS purposes will be used initially. The databases will be obtained from federal, state, and local governments, and private sources.

4. What types of data will be in the GIS? Four types of databases will be included: *Geographic framework data, such as Nassau and Suffolk county base maps from the respective county governments, road maps from New York State, and U.S. Postal Service ZIP codes files; *Demographic data, such as data from the U.S. Bureau of the Census, Medicare data from the U.S. Health Care Financing Administration (HCFA), and data from the National Health and Nutrition Examination Surveys (NHANES) of the U.S. National Center for Health Statistics (NCHS); *Health outcome and health care data, including breast cancer incidence and mortality, hospital discharge, and health facilities data from New York State; *Environmental data, such as on land use, land cover, and railroads from the U.S. Geological Survey (USGS) and Nassau and Suffolk counties; traffic from New York State; water use and potential sources of water pollution from USGS, Nassau and Suffolk counties, New York State, and the U.S. Environmental Protection Agency (EPA); air pollution monitoring and potential point sources of air, water, or soil contamination from Nassau and Suffolk counties, New York State, and EPA; chemical usage and release, and hazardous/toxic waste sites from sources including EPA, New York State, and the U.S. Department of

Agriculture (USDA); radiation from EPA; and power lines from New York State. Data from EPA's biological monitoring of human exposure from human milk and adipose tissue surveys also are expected to be included; and, *Other data, such as meteorological and climatological data from New York State, weather charts for the National Oceanic and Atmospheric Administration (NOAA), satellite image maps for the U.S. Department of the Interior (DOI), and topographic data from the U.S. Department of Commerce (DOC).

5. How will individuals access the GIS? Levels of access will be established to protect confidential data, such as medical records and other data that could identify individuals, and in accordance with data sharing requirements of the owners, or custodians, of data. Authorized users will be able to access the GIS from their laboratories or offices. A public Web site will provide reports on the progress in developing the system, summary information about its contents, links to publicly available datasets that are included in the system, and in time, findings from studies that use the GIS. The public Web site will be accessible from a home or library computer.

6. Will the community have input into the development of the GIS? The contractor will conduct two town meetings on Long Island to obtain descriptive information about the environment and its history, and input on creating and prioritizing information included in the GIS. An oversight committee for the GIS will be established that will include consumer representation. In addition, the LIBCSP has a 15-member Ad Hoc Advisory Committee, which oversees the overall Project including the GIS; the committee includes five community members. During the GIS planning phase, public workshops were held on Long Island to gain community input.

7. What is the time frame for development of the GIS? The prototype GIS will take two years to develop. The system is not expected to be ready for research purposes, except on a pilot basis, until after this time. The science of developing a GIS to support studies on relationships between environmental exposures and cancer is in its infancy, and the implementation will pose many challenges. This GIS for Long Island is the first one to be developed for studies on relationships between environment and breast cancer.

8. How are GISs used? GISs have many purposes. They are used to geographically display data for community and economic development, emergency management, national defense, and land and natural resource development, to name just a few uses. Nassau and Suffolk counties each have GISs to support their government operations. In public health, applications for GIS are more recent as geo-coded health data and environmental exposure data increasingly become available, and new and easier-to-use GIS software is developed. The U.S. Agency for Toxic Substances Disease Registry (ATSDR) uses GISs to monitor the health of persons living near hazardous waste sites, and to identify areas of potential concern resulting from accidental release of chemicals in the environment. The Centers for Disease Control and Prevention (CDC) uses GIS for disease surveillance, and EPA uses it to support risk assessment, environmental justice analysis, and ecological assessments.

9. What steps were involved in preparation of the Request for Proposals (RFP) for the GIS? A feasibility study was first conducted that indicated that it was possible to develop the GIS for Long Island. Then a working group of federal and non-federal experts in GIS and epidemiology assisted NCI to develop the technical Request for Proposals (RFP), including the identification of databases that would be important to include. The working group held public workshops on Long Island in order to obtain input from the community on development of the proposal, and also met with investigators to obtain their views. Community participation was included in all aspects of the contract evaluation and selection process where federal government procurement regulations permit.

10. What studies are included in the Long Island Breast Cancer Study Project (LIBCSP)? The LIBCSP consists of more than 10 studies, and include human population (epidemiologic) studies, the establishment of a family breast and ovarian cancer registry, and laboratory research on mechanisms of action and susceptibility in development of breast cancer. Most of the studies are being conducted, or were conducted, by investigators at research institutions in the Northeast. At least $22 million is being spent for these research activities from 1993 through 2000 by NCI in collaboration with NIEHS.

The major projects are: * Columbia Case-Control Study. This centerpiece study of the LIBCSP is a

population-based, case-control study that focuses on determining whether DDT, commonly used as a pesticide, and polycylic aromatic hydrocarbons (PAH), a ubiquitous pollutant caused by incomplete combustion of various chemicals including diesel fuel and cigarette smoke, are associated with increased risk for breast cancer on Long Island (Nassau and Suffolk counties).

All women diagnosed with breast cancer on Long Island during a one-year period were invited to participate in the study. About 1,500 women who had been diagnosed with the cancer (cases), and 1,500 women who did not have the disease (controls) enrolled. Each study participant was asked to complete an in-person interview and provide a blood and urine sample. In addition, a random sample of study participants who had resided in their homes for at least 15 years was asked to permit collection of house dust, tap water, and yard soil samples (home study). About 340 cases and 340 controls participated in this component of the study. Recruitment of study participants is completed, and data analysis is beginning. Findings are expected to be available in the year 2000. The principal investigator is Marilie Gammon, Ph.D., of the Joseph L. Mailman School of Public Health, Columbia University, New York City.

*Electromagnetic Fields (EMF) and Breast Cancer. A population-based, case-control study on Long Island is being conducted to determine if EMFs are associated with increased risk for breast cancer. The study population is a subgroup of the population participating in the Columbia case-control study and who have lived in their current residences for at least 15 years. About 600 cases and 600 controls are participating in the study. The women were interviewed about their EMF exposure, and their homes were visited to take EMF measurements, including spot and 24-hour measurements, ground current measurements, and assessments of the external power lines. No published study to date has included actual measurements of EMFs inside the homes of women who have breast cancer and healthy women. Findings are expected to be available in the year 2000. The principal investigator is M. Cristina Leske, M.D., M.P.H., of the University Medical Center at Stony Brook, Stony Brook.

*Epidemiology of Breast Cancer on Long Island. A hospital-based, case-control study is investigating breast cancer risk in relation to levels of organochlorine compounds such as polycyclic biphenyls (PCB) and chlorinated pesticides, such as DDT, on Long Island and Schoharie County, N.Y. Levels of these compounds are being analyzed in adipose tissue and blood serum. The study includes patients who were seen between 1994 and 1996 at Long Island Jewish Medical Center, New Hyde Park; North Shore University Hospital, Manhasset; and Bassett HealthCare, Inc., Cooperstown, N.Y., which serves Schoharie County. Breast tissue samples taken from about 400 women who had breast cancer surgery, and fatty tissue from about 700 women who had other surgeries were analyzed. Findings from these analyses are expected by late 1999. The study participants from the Long Island hospitals are now being followed to attempt to determine if survival or recurrence of breast cancer is related to body burden of organochlorines. A questionnaire is being administered to identify lifestyle changes and other factors that may have changed since they were diagnosed with breast cancer, and new blood samples are being taken to determine changes in serum levels over time. Steven Stellman, Ph.D., of the American Health Foundation, New York City, is the principal investigator.

*Organochlorines and Risk of Breast Cancer. This exploratory case-control study is investigating the relationship between exposure to organochlorine compounds and risk for breast cancer in Tolland County, Conn. About 150 women diagnosed with breast cancer (cases) between 1994 and 1997, and 150 women who have not had breast cancer are enrolled (controls). The study participants provided a blood sample and answered a questionnaire about environmental exposures, medical and diet history, alcohol usage, smoking, and lifetime residences. Findings are expected by late 1999. Tongzhang Zheng, M.D., Sc.D., of Yale University, New Haven, Conn., is the principal investigator.

*Reducing Barriers to Use of Breast Cancer Screening. This study focuses on women who are known to underuse breast cancer screening. It is investigating whether telephone counseling for these women, either with or without an accompanying educational session for their physicians, will result in increased use of breast cancer screening. The researchers identified over 3,400 women on Long Island who had not received mammograms in the past two years and the two years prior to that. The intervention phase then tested the

effectiveness of telephone counseling and physician education in reducing underuse of mammography. Implementation of the interventions and the post-intervention surveys of the women and physicians are now completed. The pre-and post-intervention comparative data analysis is underway to evaluate the efficacy of the interventions. Findings are expected in the year 2000. Dorothy Lane, M.D., M.P.H., of the State University of New York, Stony Brook, is the principal investigator.

*New York Metropolitan Registry for Breast and Ovarian Cancer Families. The Registry is recruiting families who have a history of breast and/or ovarian cancer. Participants are asked to contribute information, blood, and urine samples that can be used for studies on the causes of these disease. More than 800 families are now enrolled. The Registry is one of six international registry sites funded by NCI to provide researchers a source of information and specimens to speed their investigations into the causes of breast cancer. Rubie Senie, Ph.D., of the Joseph L. Mailman School of Public Health, Columbia University, New York City, is the principal investigator. New York metropolitan area residents are welcome to call 1-888-METRO-08 to inquire about enrollment. On Long Island, individuals may enroll through Stony Brook University Hospital and Medical Center, 1-800-867-3561.

Published reports on LIBCSP thus far have been about the research methods used in the studies and on findings from some basic research. Eliot Rosen, M.D., Ph.D., of Long Island Jewish Medical Center, New York City, evaluated how a growth factor called scatter factor may regulate growth of breast cancer. Using tissue samples from Long Island women, he found that levels of scatter factor are higher in invasive breast cancers than in non-invasive cancers. It also causes human breast cancer cells to move faster and to be more invasive in cell cultures. Further, the growth factor induces breast cells to produce an enzyme that degrades tissue, thus facilitating tumor invasion, and stimulates formation of new blood vessels, which is essential for tumor growth and spread. ("Scatter Factor Protein Levels in Human Breast Cancers," American Journal of Pathology 146 (5):1707-1712, November 1996.) Michael Wigler, Ph.D., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., examined genetic changes in breast tumor tissue from patients on Long Island using a technique called representational difference analysis. Certain point mutations in genes may be linked to environmental exposures, and certain mutations may be characteristic of specific exposures. ("Comparative genomic analysis of tumors: Detection of DNA losses and amplification," Proceedings of the National Academy of Sciences 92:151-155, January 1995.)

[Contacts: See NCI's new LIBCSP website at http://www-dccps.ims.nci.nih.gov/EGRP/index.html; for response to inquiries, please call the NCI Press Office at voice (301) 496-6641; for more news and information about cancer, visit NCI's Web site for patients, public and the mass media at http://rex.nci.nih.gov; questions about the overall LIBCSP may be directed to: Linda Anderson, Director of Communications, LIBCSP, Epidemiology and Genetics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, at voice (301) 496-9600 or e-mail la30e@nih.gov]

Web Site of Interest for this Edition

The Pennsylvania GIS Consortium

The Consortium is one of only six National Spatial Data Infrastructure Community Demonstration Projects in the United States. The NSDI is defined as "the technologies, policies, and people necessary to promote sharing of geospatial data throughout all levels of government, the private and non-profit sectors, and the academic community."The PAGIS Consortium, created in 1998, is a partnership between the Pennsylvania Marketing and Planning (MAP) Center at King's College and The Center for Geographic Information Sciences at Wilkes University. Both institutions are located in the city Wilkes-Barre, PA and cooperate in a number of higher education organizations. The MAP Center maintains this web site and is responsible for creating and maintaining the web site for the American Heritage River Initiative's Upper Susquehanna- Lackawanna (US-L) Watershed. Consortium activities also include visualizing GIS data and making the US-L Watershed Master Plan available to the public in electronic form. The Wilkes University Center for Geographic Information Sciences is a high-end GIS/GPS research center. It creates the technical and educational components that are part of the National Spatial Data Infrastructure (NSDI) Community Demonstration Project and directs the collection and

creation of GIS data for the American Heritage River Initiative's Upper Susquehanna-Lackawanna Watershed Master Plan. An open linkhas been established to EPA's Enviromapper where users can define the geographic area they would like to show using Enviromapper. PAGIS, in cooperation with the EPA has created this page to demonstrate to the public that GIS data from the Federal level is directly available in a form that is immediately useful. An example of this can be seen as part of the Upper Susquehanna/Lackawanna Demo Project at this web-site: http://www.mapcenter.org/pagis/pagis.html.

[Editor: I also want to call attention to June 9, 1999 testimony by Tom Sweet, Pennsylvania GIS Consortium, "Geographical Information Systems Policies and Programs," and others, before the House Subcommittee on Government Management, Information, and Technology, Committee on Government Reform which is available at http://www.house.gov/reform/gmit/hearings/testimony/990609ts.htm; For example, in Suzanne Hall's (Assistant County Executive, Wayne County, Michigan) testimony she states "In 1997, Wayne County invested $14 million from its own general tax funds to build a parcel-based map with two foot (plus or minus) accuracy level. This compares to the United States Geographic Survey (USGS) map, which has an accuracy level of 35-40 feet (plus or minus). All Wayne County departments as well as all the local municipalities under our jurisdiction will use this accurate parcel map to build application layers needed to better manage their daily business operations. Our goal is to incorporate GIS into all of Wayne County's operations. We estimate that the total cost of building all necessary Wayne County GIS applications will exceed $60 million in the next five years.

Wayne County is proceeding to build a seamless, accurate GIS map. The digitizing of all aerial photography with triangulation, planimetric and street centerlines will be completed in September of this year. By December 2000, we will complete the conversion of 900,000 parcels. Concurrent with building the base map, we are proceeding to build applications in various departments. To date, we have completed applications for emergency management, brownfields, road mobile mapping, and numerous environmental cleanup areas."]

Final Thought(s)

Special Tribute: Lee De Cola and Cynthia Warrick

We all need to recognize the extraordinary efforts of Lee De Cola, Research Physical Scientist and Geographer, U.S. Geological Survey (USGS) at Reston, VA, and Cynthia Warrick, Assistant Professor, School of Pharmacy, Howard University, Washington, D.C., for their tireless dedication to bring mainstream GIS to America's **Historical Black Colleges and Universities** (HBCUs). Since 1996, Lee and Cynthia have teamed together to conduct their HBCU Summer Faculty GIS Workshop to plant the seeds of using geospatial data, technology, and analysis in minority institutions throughout the U.S. They do this by conducting summer GIS workshops for a dozen or more minority faculty scientists each year. HBCU collaboration was actually formalized by the USGS in the 1950s and the workshops began in the early 1980's. The first GIS instruction, in 1984, drew representatives from Grambling State University (LA) and Jackson State University (MS). In all, more than 200 minority scientists have received GIS summer training in this program. Nationally there are 118 HBCUs (including the Virgin Islands).

This year was no exception as 12 faculty participants (students) from ten HBCUs that had not participated before received training at the U.S. Fish and Wildlife Service's National Conservation Training Center in Shepardstown, WVA. The venue for the workshop has varied in time from the USGS, Reston, VA to NASA's Stennis Space Flight Center to North Carolina Central University to Howard University in 1997, the same year financial support was initiated from the Centers for Disease Control and Prevention, Bureau of Land Management, and National Imagery and Mapping Agency (NIMA).

For the past several years I have had the privilege to make presentations on public health GIS developments at the workshops. The enthusiasm for GIS is always at a high level. Every student with whom I have spoken has exclaimed "Now I know what GIS can do and I can't wait to use it back in the department and community!" Through

the generosity of ESRI's Jack Dangermond (onsite sponsor for the last six years), the students receive expert GIS fundamentals and ArcView instruction (this year from old acquaintance Walt Rennick), a copy of ArcView and other GIS-related instructional materials to take back to their campus. In addition the students learn the fundamentals of global positioning systems (GPS) and receive several days of GPS fieldwork (provided by a GPS expert from NIMA). I felt honored to exchange ideas with students such as Obadiah Njue, Wiley College (TX), Alexandrine Randriamahefa, Oakwood College (AL), Robert Copeland, Jr., Howard University (D.C.), Hussain Al-Fadhli, Tougaloo College (MS) and Casandra Walker, Morehouse School of Medicine (GA). Other schools represented included Coppin State College (MD), Hampton University (VA), and South Carolina State University (SC).

This program is having far reaching effects on minority institutions in addition to the rewards of personal technology empowerment and networking. Many institutions are expanding their programs related to the areas of geographic information science and community involvement. New courses and community development grant programs are emerging. For example, Savannah State (GA) will introduce GIS in coursework to examine teenage pregnancy among African Americans in Chatham County and study vectors of Dirofilaria immitis and Eastern Equine Encephalitis virus endemic in the county. Fayetteville State University (NC) has introduced GIS in its collaborative EPA Brownfields redevelopment grant program with the city.

Bridges between local public health departments and six minority health professions schools have been recently funded (North American Association for Environmental Education -Urban Collaboratives Grant Program) as an extension of the HBCU Summer Faculty GIS Workshop. These are designed to use GIS to assess potential environmental health risks of local communities. Participating institutions include Howard University College of Medicine, Morehouse School of Medicine, Meharry Medical College (TN), Xavier School of Pharmacy (LA), Drew School of Medicine (CA), and the University of Texas Health Science Center (TX).

Perhaps one of the most important benefits to society of the HBCU Summer Faculty GIS Workshop is that it connects with many HBCUs that tend to be small, rural and financially less endowed. The workshop brings mainstream GIS to those institutions that otherwise might miss GIS technology. It will have a significant effect (cumulative) over time to establish greater parity (capability) among all institutions of higher learning in GIS. One day, when we no longer as a nation experience under-representation of minorities in scientific fields, we can proudly point to the HBCU Summer Faculty GIS Workshop as having played a role in this transition. The HBCU Summer Faculty GIS Workshop is a reality thanks to the vision and dedicated leadership of Lee De Cola and Cynthia Warrick.

Charles M. Croner, Ph.D., Editor, ***PUBLIC HEALTH GIS NEWS AND INFORMATION***, Office of Research and Methodology, National Center for Health Statistics, at e-mail ccroner@cdc.gov. Although this report is in the public domain, the content should not be changed.

Special Tribute to Lee De Cola and Cynthia Warrick-HBCU GIS Summer Faculty Workshops