26. Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., and Rising, G. R., eds. (1973). *Statistics by Example*, 4 vols. Addison-Wesley, Reading, MA.
27. Mosteller, F. and Tukey, J. W. (1977). *Data Analysis and Regression*. Addison-Wesley, Reading, MA.
28. Mosteller, F. and Wallace, D. L. (1984). *Applied Bayesian and Classical Inference*. Springer, New York.
29. Rubin, D. B. (1984). *Ann. Statist.*, **12**, 1151–1172.
30. Shortliffe, E. H. (1976). *Computer-based Medical Consultations, MYCIN*. Elsevier, New York.
31. Siegmund, D. (1985). *Sequential Analysis*. Springer, New York.
32. Smith, A. F. M. (1984). *J. R. Statist. Soc. Ser. A*, **147**, 245–259.
33. Spiegelhalter, D. J. and Knill-Jones, R. P. (1984). *J. R. Statist. Soc. Ser. A*, **147**, 35–77.
34. Stigler, S. M. (1986). *The History of Statistics*, Harvard University Press, Cambridge, MA.
35. Tanur, J., Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., Rising, G. R., and Lehmann, E. L., eds. (1978). *Statistics: A Guide to the Unknown*, 2nd ed. Wadsworth, Belmont, CA.
36. Wolfowitz, J. (1943). *Ann. Math. Statist.*, **14**, 280–288.

E. L. LEHMANN

STATISTICS AT CDC

The Centers for Disease Control and Prevention (CDC) is recognized as the lead U.S. Federal agency for protecting the health and safety of people. It accomplishes this mission by promoting health and quality of life with programs preventing and controlling disease, injury, and disability. In this mission CDC is joined by the Agency for Toxic Substances and Disease Registry (ATSDR), which is charged with evaluating the human health effects of exposure to hazardous substances. Both CDC and ATSDR are components of the U.S. Department of Health and Human Services (DHHS). In 2004, CDC undertook an ambitious strategic planning process called the *Futures Initiative* to better meet the challenges to public health in the 21st century. Modernizing CDC will enhance its impact on emerging health problems, support its capacity to respond to public health emergencies, and more effectively serve the American public. In this transition, CDC decided to align its priorities and investments under two overarching health protection goals:

- *Health promotion and prevention of disease, injury, and disability:* All people, especially those at higher risk due to health disparities, are to achieve their optimal lifespan with the best possible quality of health in every stage of life.
- *Preparedness:* People in all communities are to be protected from infectious, occupational, environmental, and terrorist threats.

To accomplish these goals, the newly reorganized CDC will be comprised of four new Coordinating Centers (Coordinating Center for Infectious Diseases, Coordinating Center for Health Promotion, Coordinating Center for Public Health Information and Services, and the Coordinating Center for Environmental Health and Injury Prevention), three Offices (Office of Global Health, Office of Terrorism, Preparedness, and Emergency Response, and Office of the Director), and the National Institute for Occupational Safety and Health with responsibilities for: birth defects and developmental disabilities; chronic disease prevention and health promotion; environmental health; health statistics; infectious diseases; human immunodeficiency virus (HIV) infection, sexually transmitted diseases, and tuberculosis prevention; injury prevention and control; vaccine-preventable disease; occupational safety and health; public health surveillance; and community practice of public health. CDC/ATSDR employs more than 8,500 people in locations throughout the United States and over 50 foreign countries, with headquarters in Atlanta, Georgia. CDC/ATSDR employs approximately 330 mathematical and health statisticians.

The CDC began in 1946 as the Communicable Disease Center and successor to the Office of Malaria Control in War Areas, part of the U.S. Public Health Service. In 1951

it established the Epidemic Intelligence Service (EIS), which is a training program in applied epidemiology. EIS officers investigate all types of epidemics, including chronic disease and injuries. Publication of the *Morbidity and Mortality Weekly Report (MMWR)*, providing weekly data on deaths and morbidity in each state, was transferred to CDC in 1961. CDC has been renamed several times to reflect its growth and dynamic mission, becoming the current Centers for Disease Control and Prevention in 1992. ATSDR was created in 1980 by congressional mandate to implement the health-related sections of laws protecting the public from hazardous wastes and environmental spills of hazardous substances. Thus, ATSDR concentrates its efforts at Superfund sites and works closely with the Environmental Protection Agency as well as with CDC. CDC performs many of the administrative functions for ATSDR, and the Director of CDC also serves as the Administrator of ATSDR. In 1987, the National Center for Health Statistics (NCHS) became a part of CDC. NCHS is a Federal statistical agency with broad responsibilities to monitor the health of the nation. In addition to conducting a data collection program that encompasses vital statistics, interview surveys, examination surveys and provider surveys, NCHS prepares the annual report, "Health, United States" [10] submitted by the DHHS Secretary to the President and Congress to present a comprehensive profile of health in America and track key indicators and trends. NCHS is also responsible for advancing the field of health statistics through research into statistical and analytical methods.

PUBLIC HEALTH DATA AND STATISTICS AT CDC/ATSDR

Since the inception of the agency, an important function of CDC has been compilation, analysis, and interpretation of statistical information to guide actions and policies to improve health. Sources of data include vital statistics records, medical records, personal interviews, telephone and mail surveys, physical examinations, and laboratory testing. Surveillance data has been used to characterize the magnitude and distribution of illness and injury, to track health trends, and for development of standard curves such as growth charts. Beyond the development of appropriate program study designs and analytic methodologies, statisticians have played roles in the development of public health data collection systems and software to analyze collected data.

STATISTICS AND RESEARCH AT CDC/ATSDR

The integration of statistics and analytic techniques into public health research is a critical asset to the agency, and has resulted in important applications in various disciplines such as epidemiology, economics, and the behavioral and social sciences. Examples include: the use of statistical methods to demonstrate the relationship between removing lead from gasoline and decreased blood lead levels in children and adults [1]; economic determinations contributing to folic acid supplementation of foods to decrease birth defects [7]; behavioral science methods leading to the development of strategies for the prevention of HIV/AIDS [6]; quantitative epidemiologic analyses leading to an understanding of the relationships between radon and lung cancer in coal miners [5], and evaluations of the effectiveness of using back belts in reducing back injury claims and back pain [22]. Other areas of continuing statistical contribution include survey planning and analytic methodology, data collection systems, detection algorithms and scan statistics to document health trends and identify emerging health issues, and model development to project disease incidence and injury or numbers of cases prevented through treatment and public health measures during an outbreak. For example, new methodologies have been developed to make it possible to compare population characteristics across data collection programs and over time when there is a change in data collection methods [14] and to quantify disparities in health and health care [8]. Methodological work has also been done on how to deal with high levels of nonresponse on central variables such as income [15]. The National Laboratory for Collaborative Research in Cognition and Survey Measurement applies cognitive

methods to questionnaire design research and the testing of data collection instruments to improve data quality [11] The reliance on statistics for policy and programmatic use and the growing number and diversity of users has required on-going research and innovative approaches to protect the confidentiality of respondents and data sources while offering the widest possible access to data [3].

Recent CDC activities presenting new analytic challenges include counterterrorism and emerging infectious diseases. CDC statistical programs have contributed to development of methodologies for syndromic surveillance, techniques useful during anthrax investigations and remediation, and approaches to the civilian smallpox vaccination effort. Such techniques have also been useful for emerging infectious diseases such as severe acute respiratory syndrome (SARS) and in extending analytic capabilities for chronic diseases and in developing approaches to national health report cards (i.e., health assessments or health profiles).

The anthrax investigations of September-December 2001 spurred development of multiple analytic techniques. Maps linking analytic sampling activity with analytic results were developed to better understand the spread and deposition of spore-containing particles. Numbers of cases reported were used to estimate parameters of mathematical epidemiological models, which in turn simulated the potential size and duration of the outbreak and the effect of preventative measures [2]. Analyses of environmental sampling information have included comparison of analytic sampling techniques, tests of sensitivity and specificity of various presumptive and confirmatory tests for *Bacillus anthracis*, and use of statistical methods such as survival analysis and dose-response modeling to quantify critical exposure levels [17].

Stochastic simulation has been used to optimize patient flow-through in clinics dispensing oral antibiotics following a bioterrorism attack [21]. Aberration detection in public health data represents another area of statistical contribution. For example, the Smallpox Preparedness and Response Activity (SPRA) at CDC receive vaccination and adverse event data from several sources. These sources employ both active and passive data collection and provide registry, contraindication, and adverse events information. These disparate data are joined into analyzable elements needed for decision making.

THE CDC/ATSDR STATISTICAL ADVISORY GROUP

Coordination of statistical activities across CDC and ATSDR resulted in formal recognition of the statistical community in 1989. In that year, at the request of the CDC Office of the Director, a coordinating group for analytic methods known as the CDC/ATSDR Statistical Advisory Group (SAG) was created. The SAG was established to act in an advisory capacity to the Office of the Director to facilitate and address statistical issues, problems, and opportunities which influenced the quality and integrity of science at CDC, and to increase communication across organizational components.

Some activities of the SAG further illustrate the breadth of statistical activity across CDC/ATSDR. Biennial symposia have been held since 1989 on topics of interest to the public health community (clustering methods [20], surveillance [12], evaluation of intervention and prevention strategies [13], use of multi-source data [23], small area statistics [19], public health decision making [4], emerging statistical issues in public health [16], and study design and data structures [9]). Each year the SAG recognizes outstanding statistical papers published during the previous year with the CDC/ATSDR Statistical Science Awards. Recent winners included manuscripts on spatiotemporal modeling [18] and marginal analyses of clustered data [24]. SAG is responsible for advanced statistical/epidemiological training and maintains a listserv and internet site (*http://www.cdc.gov/sag*).

Other statistical activities include participation in statistical/protocol review and leadership in the development, procurement, and installation of statistical software available for use by the CDC/ATSDR community. The SAG has provided review and advice on complex statistical and broad scientific issues

such as validation of the statistical design of the Vietnam Experience Study of the health of Vietnam veterans, and co-developed an evaluation of recruitment and retention policies at CDC/ATSDR, for example. Other special requests, such as for development of training materials or requests for interagency collaboration/representation, are also often handled through the SAG. Since 1990 the SAG has sponsored an exhibit booth highlighting statistical activities at CDC/ATSDR that has been displayed at the Joint Statistical Meetings and other conferences for informational and recruiting purposes.

FUTURE DIRECTIONS FOR STATISTICS AT CDC/ATSDR

The critical role of statistics toward accomplishing the mission of CDC/ATSDR will become even more apparent as the agency begins to align its activities around its overarching goals of health promotion and preparedness. The assessment of burden, effectiveness of interventions, cost considerations, and evaluation frameworks will all require rigorous attention to methods of data collection, study design, and analytic technique. The role of statisticians to ensure that quantitative sciences will be most effectively utilized in research and analysis, and in meeting new challenges in CDC/ATSDR's evolving public health mission will require a reexamination of statistical skills and contributions. A multi-disciplinary approach to investigation of public health problems such as counterterrorism or obesity is already being realized. Efficient use of new technologies such as in informatics, web-based query systems, geographic information systems, and survey data collection methodologies will be key to continued valuable statistical input. Advances in the field of relational databases, for example, and its coupling with web-based technology have facilitated improvements in the efficiency of data collection and increases in size and completeness of data available for analysis. The BioSense program currently under development at CDC, uses information from both traditional and non-traditional data sources (such as emergency calls and over-the-counter drug sales) for syndromic surveillance. Such use of multi-source data and further development of record linkage techniques to extract maximal information from existing data sources will also require addressing privacy and confidentiality concerns, as well as appropriate methods of communication of important public health findings to the nation.

Acknowledgment

The authors wish to express their appreciation to Dr. Jennifer Madans for her help in preparing this description of statistical activities across CDC and ATSDR.

REFERENCES

1. Annest, J.L., Pirkle, J.L., Makuc, D., Neese, J.W., Bayse, D.D., and Kovar, M.G. (1983). Chronological trend in blood lead levels between 1976 and 1980. *N. Engl. J. Med.*, **308**, 1373–1377.
2. Brookmeyer, R.D. and Blades, N. (2003). Statistical models and bioterrorism: application to the U.S. anthrax outbreak. *J. Am. Stat. Assoc.*, **98**(464), 781–788.
3. Cox, L.H. (2002). Disclosure Risk for Tabular Economic Data. In: *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, Doyle, P., Lane, J., Theeuwes, J., Zayatz, L., eds. Elsevier, New York, pp. 167–183.
4. Falter, K.H., Betts, D.R., Rolka, D.B., Rolka, H.R., and Sieber, W.K., eds. (1999). Symposium on statistical bases for public health decision-making: from exploration to modelling. *Stat. Med.* **18**(23), 3159–3376.
5. Hornung, R.W. and Meinhardt, T.J. (1987). Quantitative risk assessment of lung cancer in U.S. Uranium miners. *Health Physics*, **52**(4), 417–430.
6. Kamb, M.L., Fishbein, M., Douglas, J.M., Rhodes, F., Rogers, J., Bolan, G., Zenilman, J., Hoxworth, A., Malotte, K., Iatesta, M., Kent, C., Lentz, A., Graziano, S., Byers, R., Peterman, T. (1998). Efficacy of risk-reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases. *J. Am. Med. Assn.*, **280**, 1161–1167.
7. Kelly, A.E., Haddix, A.C., Scanlon, K.S., Helmick, C.G., and Mulinare, J. (1996). Cost-effectiveness of strategies to prevent neural tube defects. In: *Cost-Effectiveness in Health and Medicine*, M.R. Gold, J.E. Siegel,

L.B. Russell, and M.C. Weinstein, eds. Oxford University Press, New York, pp. 313–348.

8. Keppel, K.G., Pearcy, J.N., Klein, R.J. (2004). *Measuring Progress in Health People 2010.* Healthy People 2010 Statistical Notes. No. 25, National Center for Health Statistics, Hyattsville, Maryland.
9. National Center for Health Statistics. (2004). *Health, United States, 2004, With Chartbook on Trends in the Health of Americans.* Hyattsville, Maryland.
10. National Center for Health Statistics. *Cognitive Methods.* Working Paper Series. *www.cdc.gov/nchs/products/pubs.*
11. Reynolds, G.H., McGee, D.L., and Stroup, D.F., eds. (1989). Symposium on statistics in surveillance. *Stat. Med.*, **8**(3), 251–400.
12. Reynolds, G.H., ed. (1993). Symposium on statistical methods for evaluation of intervention and prevention strategies, December 1990. *Stat. Med.* **12**(3/4), 191–414.
13. Schenker, N., Parker, J.D. (2003). From single-race reporting to multiple-race reporting: using imputation methods to bridge the transition. *Statistics in Medicine*, **22**, 1571–1587.
14. Schenker, N., Raghunathan, T.G., Chu, P., Makuc, D.M., Zhang, G., Cohen, A.J. (2004). *Multiple Imputation of Family Income and Personal Earnings in the National Health Interview Survey.* National Center for Health Statistics, Hyattsville, Maryland.
15. Sieber, W.K., Martinez, K.M., and Wacherman, B. (2002). Anthrax sampling and an investigative database, *Presentation 2nd Conference on Statistics in Counterterrorism*, May 29-30, Washington, D.C.
16. Smith, R.L., Klenikov, S., and Cox, L.H. (2003). Spatiotemporal modeling of PM2.5 data with missing values. *J. Geophys. Res.*, **108**, 11–1, 11–11.
17. Smith, S.J., ed. (1996). Symposium on small area statistics in public health: design, analysis, graphic and spatial methods. *Stat. Med.*, **15**(17/18), 1827–1986.
18. Steinberg, K., ed. (1990). National conference on clustering of health events, *Am. J. Epid.* **132**(1), S1–S202.
19. Washington, M.L. (2003). Optimizing patient flow-through in large-scale clinics for dispensing oral antibiotics following a bioterrorism attack using discrete-event computer simulation, *Presentation, 9th Biennial CDC/ATSDR Symposium on Statistical Methods: Study Design and Decision Making.* January 27-29, 2003, Atlanta, Georgia.
20. Wassell, J.T., Gardner, L.I., Landsittel, D.P., Johnston, J.J., and Johnston, J.M. (2000). A prospective study of back belts for prevention of back pain and injury. *JAMA*, **284**(21), 2727–2732.
21. Williamson, G.D., Massey, J.T., Shulman, H.B., Sieber, W.K., and Smith, S.J., eds. (1995). Symposium on quantitative methods for utilization of multi-source data in public health. *Stat. Med.*, **14**(5/6/7), 447–718.
22. Williamson, J.M., Satten, G.A., and Datta, S. (2003). Marginal analyses of clustered data when cluster size is informative. *Biometrics*, **59**, 36–42.

FURTHER READING

Centers for Disease Control and Prevention. Home page [Online]. *http://www.cdc.gov.* [2004, August 25].

Brookmeyer, R. and Stroup, D. F., eds. (2004). *Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance.* Oxford University Press, New York.

Cox, L. H., Copeland, J. R., Katzoff, M. J., Rolka, H. R., and Sieber, W. K. (2003). *White Paper: The Role of Statistical Science in Meeting CDC/ATSDR's Evolving Public Health Mission.*

Williamson, G. D., Snider, D. E., and Speers, M. A. (2000). Facilitating use of analytic methods at a Federal agency. *Am. J. Prev. Med.*, **19**(1S), 40–46.

WILLIAM K. SIEBER, JR.
DONNA F. STROUP
G. DAVID WILLIAMSON

STATISTICS CANADA

Statistics Canada is the nation's central statistical agency, mandated by the Statistics Act of 1971 to "collect, compile, analyse, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and condition of the people." The Act also made Statistics Canada responsible for collaborating with departments of government in the collection, compilation, and publication