LETTERS

DIET AND CARDIOVASCULAR DISEASE

I read with great interest "A Motivational Interviewing Intervention to Increase Fruit and Vegetable Intake Through Black Churches: Results of the Eat for Life Trial," by Resnicow and colleagues.[1] Modifying lifestyle risk factors has proven beneficial in preventing or controlling chronic diseases (cardiovascular disease and cancer) and decreasing mortality.[2,3] Resnicow et al. found a significant difference in fruit and vegetable intake among participants from urban Black churches who underwent 3 motivational phone interviews in addition to receiving culturally sensitive self-help materials and usual educational material, compared with participants who received the latter 2 (group 2) or just the latter (group 1).

Several additional points can be made. First, this study proves once again the importance of Black churches as a venue for recruiting and retaining African Americans from a wide range of age and socioeconomic groups. However, the investigators could have put their results into perspective by relating the amount of change to the degree of prevention of disease or risk factor

Letters to the editor will be reviewed and are published as space permits. By submitting a letter to the editor, the author gives permission for its publication in the Journal. Letters should not duplicate material being published or submitted elsewhere. Letters referring to a recent Journal article should be received within 3 months of the article's appearance. The editors reserve the right to edit and abridge letters and to publish responses.

Text is limited to 400 words and 10 references. Submit on-line at www.ajph.org for immediate Web posting, or at submit.ajph.org for later print publication. On-line responses are automatically considered for print publication.

TABLE 1—Relative Risk of Cardiovascular Outcomes in Relation to Fruit and Vegetable Intake in Selected Cohorts

Study (Setting)	Population	Follow-up	Outcome	No. Servings per Day	Adjusted Risk Ratio[a] (95% Confidence Interval)
Joshipura[4] (Nurses Health Study	84 251 women, 34–59 y	14 y	Coronary heart disease	Fruit and vegetables, 1	0.96 (0.94, 0.99)
and Health Professional	42 148 men, 40–75 y	8 y		Fruit, 1	0.94 (0.90, 0.98)
Follow-up Study)				Leafy vegetables, 1	0.77 (0.64, 0.93)
Liu[5] (Women's Health Study)	39 876 women, mean age	5 y	Cardiovascular disease[b]	Fruit, quintile median	
	(SD) 52 (6) to 55 (7) y			0.6	1.0
				1.3	0.71 (0.39, 1.26)
				1.9	0.67 (0.37, 1.23)
				2.6	0.79 (0.44, 1.42)
				3.8	0.57 (0.30, 1.09)
				Vegetables, quintile median	
				1.5	1.0
				2.5	0.63 (0.35, 1.14)
				3.4	0.59 (0.33, 1.09)
				4.5	0.96 (0.54, 1.61)
				6.8	0.45 (0.24, 0.89)
Joshipura[6] (Nurses Health Study)	75 596 women, 34–59 y	14 y	Ischemic stroke	Fruit and vegetables, 1	0.94 (0.90, 0.99)
	38 683 men, 40–75 y	8 y		Citrus fruit, 1	0.81 (0.68, 0.96)
				Leafy vegetables, 1	0.79 (0.62, 0.99)
				Cruciferous vegetables, 1	0.68 (0.49, 0.94)
			All stroke	Fruit or vegetables, 1	0.96 (0.93, 1.00)
Gillman[7] (Framingham Study)	Men, 45–65 y	20 y	All stroke, including transient ischemic attack	Fruit and vegetables, 3	0.77 (0.60, 0.98)

[a]Adjusted for cardiovascular risk factors (varied between studies): age; menopausal status; body mass index; smoking; exercise; alcohol use; energy intake; fat intake; multivitamin use; aspirin use; postmenopausal hormone use; personal history of diabetes, hypertension, or high cholesterol; family history of myocardial infarction; systolic blood pressure; glucose intolerance; and left ventricular hypertrophy.
[b]Nonfatal myocardial infarction, stroke, percutaneous transluminal coronary angioplasty, coronary artery bypass graft, or cardiovascular disease–related death.

modification. Indeed, the magnitude of effect, though statistically significant, may appear small.

During the 1-year follow-up, participants in group 3 consumed, on average, 0.7 more servings of fruit per day than participants in group 1 and 0.6 more servings per day than participants in group 2. The increase in vegetable intake was 0.5 and 0.4 servings for the same group comparisons. Interestingly, prospective studies do show a beneficial effect of a similar magnitude of combined fruit *and* vegetable intake on the incidence of coronary heart disease,[3] nonfatal myocardial infarction, stroke, coronary surgical procedures, and cardiovascular disease–related death (Table 1).[4–7]

Second, a 1-year follow-up, even if it is associated with positive outcomes, seems short in the disease course. Hence the methodological choices are either to perform longer studies or to monitor intermediate end points associated with hard end points, such as weight or subclinical disease.[8,9] It would also have been interesting to see whether the method had different effects depending on participants' initial health status. Finally, we need more work on how to recruit out-of-the mainstream participants who have different sets of additional contextual risk factors and who also deserve to have their risk factor profiles modified.

Dietary behavior modification will continue to be an important issue in the prevention and management of cardiovascular disease. Thus, this encouraging study calls for more investigations that involve (1) intensifying the intervention from "brief" to "full-blown" motivational study, as suggested by Resnicow et al., to assess whether the magnitude of change is greater; (2) adding some measurable intermediate end points that have been related to hard pathological end points; and (3) adding a time-series dimension with a longer follow-up period to determine the minimal maintenance period of behavioral change. This study illustrates that achieving even small differences can make a difference in cardiovascular outcomes for high-disparity populations. ■

Rebecca Din, MD, MPH, PhD

About the Author

Requests for reprints should be sent to Rebecca Din, MD, MPH, PhD, Morehouse School of Medicine, Social Epidemiology Research Division, 720 Westview Dr SW, Atlanta GA 30310-1495 (e-mail: rebecca_din@msm.edu).

References

1. Resnicow K, Jackson A, Wang T, et al. A motivational interviewing intervention to increase fruit and vegetable intake through black churches: results of the Eat for Life trial. *Am J Public Health.* 2001;91:1686–1693.

2. Ness AR, Powles JW. Fruit and vegetables, and cardiovascular disease: a review. *Int J Epidemiol.* 1997;26:1–13.

3. Cronin KA. Evaluating the impact of population changes in diet, physical activity, and weight status on population risk for colon cancer (United States). *Cancer Causes Control.* 2001;12:305–316.

4. Joshipura KJ, Hu FB, Manson JE, et al. The effect of fruit and vegetable intake on risk for coronary heart disease. *Ann Intern Med.* 2001;134:1106–1114.

5. Liu S, Manson JE, Lee IM, et al. Fruit and vegetable intake and risk of cardiovascular disease: the Women's Health Study. *Am J Clin Nutr.* 2000;72:922–928.

6. Joshipura KJ, Ascherio A, Manson JE, et al. Fruit and vegetable intake in relation to risk of ischemic stroke. *JAMA.* 1999;282:1233–1239.

7. Gillman MW, Cupples LA, Gagnon D, et al. Protective effect of fruits and vegetables on development of stroke in men. *JAMA.* 1995;273:1113–1117.

8. Khaw KT, Bingham S, Welch A, et al. Relation between plasma ascorbic acid and mortality in men and women in EPIC-Norfolk prospective study: a prospective population study. *Lancet.* 2001;357:657–663.

9. Dwyer JH, Navab M, Dwyer KM, et al. Oxygenated carotenoid lutein and progression of early atherosclerosis: the Los Angeles atherosclerosis study. *Circulation.* 2001;103:2922–2927.

RESNICOW RESPONDS

My coauthors and I appreciate Din's comments. An interesting concern she raised is that we could have better "put our results into perspective by relating the amount of clinical change to the degree of prevention of disease or risk factor modification." Din points out, on the basis of the findings of several prospective observational studies,[1–4] that the net relative increase of approximately 1 serving per day of fruits and vegetables in the intensive intervention group could result in a significant beneficial effect on disease outcomes.

Projecting potential morbidity or mortality effects from our intervention is complicated by several factors. First, as Din notes, our primary outcome was assessed only 1 year from baseline. The extent to which the behavioral changes observed would be maintained beyond this relatively short follow-up is unclear. Moreover, even if the change was fully maintained, it is likely that the impact on disease outcomes would be modest. Joshipura et al.[1] estimated that for 1 healthy middle-aged adult to avoid a coronary heart disease event, 1443 persons would have to increase their consumption of fruits and vegetables by 1 serving per day for 12 years. Finally, the studies cited are observational in nature, and the extent to which *changing* fruit and vegetable intake improves disease prognosis in healthy individuals is not well understood.[5]

With regard to her question as to whether the intervention may have had differential effects on persons with prior disease, we conducted new analyses and found that baseline disease status did not interact with the intervention.

Din raises several other important points, including the need for longer-term follow-up in health promotion intervention studies and the need to include physiological outcomes in such studies. She mentions the benefits of conducting research in Black churches and the importance of including the full socioeconomic spectrum of African Americans in public health research. We are currently conducting a study that focuses on changing diet and physical activity habits among African American adults recruited through Black churches.[6] ■

Ken Resnicow, PhD

About the Author

Requests for reprints should be sent to Ken Resnicow, PhD, Rollins School of Public Health, Emory University, 1520 Clifton Rd, Atlanta, GA 30322 (e-mail: kresnic@sph.emory.edu).

References

1. Joshipura KJ, Hu FB, Manson JE, et al. The effect of fruit and vegetable intake on risk for coronary heart disease. *Ann Intern Med.* 2001;134:1106–1114.

2. Liu S, Manson JE, Lee IM, et al. Fruit and vegetable intake and risk of cardiovascular disease: the Women's Health Study. *Am J Clin Nutr.* 2000;72:922–928.

3. Joshipura KJ, Ascherio A, Manson JE, et al. Fruit

and vegetable intake in relation to risk of ischemic stroke. *JAMA.* 1999;282:1233–1239.

4. Gillman MW, Cupples LA, Gagnon D, et al. Protective effect of fruits and vegetables on development of stroke in men. *JAMA.* 1995;273:1113–1117.

5. de Lorgeril M, Salen P. Modified Cretan Mediterranean diet in the prevention of coronary heart disease and cancer. *World Rev Nutr Diet.* 2000;87:1–23.

6. Resnicow K, DiIorio C, Blisset D, Braithwaite RL, Perisamy S, Rahotep S. Healthy body/healthy spirit: a church-based nutrition and physical activity intervention. *Health Educ Res.* In press.

INDUSTRY ATTACKS ON SCIENCE

I applaud Rosenstock and Lee for their article "Attacks on Science: The Risk to Evidence-Based Policy."[1] The authors describe tactics used by a variety of interests to undermine public health initiatives and suggest that the scientific community understand these threats and devise institutional responses to them. They illustrate the need for action by recounting the measures used by an industry group to delay the National Institute for Occupational Safety and Health and National Cancer Institute (NIOSH–NCI) epidemiological study of diesel exhaust. A lesser-known aspect of this story exposes the mining industry group's duplicity and intensifies the call for a public health response.

While the US Department of Health and Human Services was engaged in its legal battle with the Methane Awareness Resource Group (MARG), another cabinet-level agency was hearing a much different tale from this industry coalition. The US Department of Labor (DOL) was engaged in rulemaking to protect underground miners, the most heavily exposed workers, from diesel particulate matter.[2,3] In writing and at public hearings before DOL officials, MARG expressed its support for the NIOSH–NCI study. The group indicated that it was "participating cooperatively with government researchers"[4] and that it "endorsed the study."[5] At times, mining industry executives extolled the study's value, noting that it "has the potential to fill in many knowledge gaps."[6] MARG's statements to DOL officials suggested that the group eagerly awaited the study's results.

In reality, the opposite was true. The coalition worked to block the workplace standards by pressuring the Secretary of Labor to wait for completion of the study. Industry representatives argued that a delay in the new rules was necessary because the science was incomplete. One testified that the study would "offer definitive data on the actual mining population . . . not a biased view of various academic studies."[7]

The coalition, however, was actually engaged in inventive legal maneuvers against the Department of Health and Human Services to thwart the epidemiological research. Through its separate interactions with the 2 agencies (NIOSH and DOL), MARG successfully characterized itself as an active participant in and advocate for the study, at the same time that it was amassing a written record of opposition to the study. Ultimately, this was a new twist to a tried-and-true strategy: oppose protective regulations by arguing for better science, and obstruct the research that would enhance scientific understanding and improve evidence-based policy. Such blatant attacks on science will not be addressed by government agencies. The public health community must answer the call. ■

Celeste A. Monforton

About the Author

The author was with the Occupational Safety and Health Administration and the Mine Safety and Health Administration, US Department of Labor, Washington, DC, from 1991 to 2001.

Requests for reprints should be sent to Celeste A. Monforton, 14503 Four Chimney Dr, Centreville, VA 20120 (e-mail: cmonforton@yahoo.com)

References

1. Rosenstock L, Lee LJ. Attacks on science: the risk to evidence-based policy. *Am J Public Health.* 2002: 92:14–18.

2. Proposed rule on diesel particulate matter exposure of underground coal miners, US Department of Labor, Mine Safety and Health Administration. *Federal Register.* April 9, 1998;63:17492–17579.

3. Final risk assessment for diesel particulate matter, US Department of Labor, Mine Safety and Health Administration. *Federal Register.* January 19, 2001;66: 5560–5667.

4. *Public Hearing Before the Mine Safety and Health Administration* (November 17, 1998) (testimony of Rowdy Heiser, representing the FMC Corp and MARG Diesel Coalition).

5. *Public Hearing Before the Mine Safety and Health Administration* (May 11, 1999) (testimony of Kent Adamson, CIH, CSP, representing the Solvay Minerals Corp and MARG Diesel Coalition).

6. *Public Hearing Before the Mine Safety and Health Administration* (May 11, 1999) (testimony of Wes Ing, chairman of the National Mining Association's Metal and Nonmetal Mine Diesel Task Group).

7. *Public Hearing Before the Mine Safety and Health Administration* (May 11, 1999) (testimony of Christopher Pritchard, representing Tg Soda Ash Inc).

ROSENSTOCK RESPONDS

Monforton's letter underscores a main tenet of our article, namely, that the prime motivation for vested interests' undermining science is to thwart the policy implications that may follow. As director of the National Institute for Occupational Safety and Health during the time described, I can personally attest to the dual and seemingly contradictory strategy employed by a group of mine operators (Methane Awareness Resource Group). On the one hand, employ a battery of steps to block, in the Department of Health and Human Services, a sentinel diesel health effects study; on the other hand, attempt to block, in the Department of Labor, regulatory efforts to address diesel exposure. In both cases, the group cited too much scientific uncertainty and the need for more research.

The dual strategy was not a well kept secret but was quite visible, and, all the more disconcerting, remarkably effective. In fact, the very same congressional decisionmakers and their staffs—because the same committees had jurisdiction over both agencies—were lobbied on both sides of the issue (delay research and block regulation), often with success.

In addition to the other recommendations made in our article, this case study of diesel exhaust, which Monforton has aptly expanded on, demonstrates the need for a multifaceted response and reinforces one of our primary recommendations: "First, consider the context and the source of the attack. . . . [T]he economically and politically powerful can too easily compromise the use of good science." ■

Linda Rosenstock, MD, MPH

About the Authors

Linda Rosenstock is with the School of Public Health, University of California at Los Angeles.

Requests for reprints should be sent to Linda Rosenstock, MD, MPH, School of Public Health, University of California at Los Angeles, PO Box 951772, Los Angeles, CA 90095-1772 (e-mail: lindarosenstock@ph.ucla.edu).

TRAINING THE PUBLIC HEALTH EDUCATION WORKFORCE

The important information presented by Allegrante et al.[1] is timely, offering an initial framework to assist in the planning and development of training opportunities to enhance the effectiveness of the public health education workforce. Unfortunately, input from the practicing health educators targeted for training appears to be missing from the discussion.

Although the competencies identified by the consensus panel are important to consider, continuing education programs have been shown to be most effective when tailored to suit the specific needs of the professionals participating in the training.[2] Additionally, continuing education training differs from preservice training in that participants typically have on-the-job experience and have specific job-related tasks they are interested in improving. A panel may be able to address global issues related to health education job skill needs. However, for workforce training to be effective, emphasis must be placed on the wants and needs identified by currently employed health educators.

In addition to content, careful consideration must be given to identifying individuals and organizations responsible for developing and implementing training programs. An organized approach and the cooperation of many groups are important to the success of this type of effort. Workforce training comes from a less formalized training body than the university-based educational programs that teach entry-level and graduate competencies. Who will be responsible for addressing the continuing education needs of health educators? From where will the funding and motivation for delivering this type of program come? The success of health educators in addressing the changing needs of health education and health promotion in the 21st century relies heavily on the ability to provide a unified training focus.

Furthermore, consideration must be given to the multiple modes of delivery available for the instruction and facilitation of educational programs. In 1997, a report from the Public Health Functions Project[3] suggested that public health workforce training programs should maximize the use of evolving technologies such as distance learning. However, as Ehrmann cautions, "If you're headed in the wrong direction, technology won't help you get to the right place."[4(p21)] Little evidence-based research is available to draw from as health education workforce training programs are put into practice, and careful consideration must be given to the process to ensure movement in the right direction. ■

Jane Ellery, MA

About the Author

Jane Ellery is a doctoral candidate at the College of Public Health, University of South Florida, Tampa.

Requests for reprints should be sent to Jane Ellery, MA, University of South Florida College of Public Health, Department of Community and Family Health, 13201 Bruce B. Downs Blvd, Tampa, FL 33612-3805 (e-mail: jellery@hsc.usf.edu).

References

1. Allegrante JP, Moon RW, Auld ME, Gebbie KM. Continuing-education needs of the currently employed public health education workforce. *Am J Public Health.* 2001;91:1230–1234.

2. Swerissen H, Tilgner L. A workforce survey of health promotion education and training needs in the state of Victoria. *Aust N Z J Public Health.* 2000;24:407–412.

3. *The Public Health Workforce: An Agenda for the 21st Century.* Washington, DC: US Dept of Health and Human Services; 1997.

4. Ehrmann SC. Asking the right questions. *Change.* 1995;27:20–27.

ALLEGRANTE ET AL. RESPOND

We thank Ellery for her comments on our article. Nineteen of the 25 members of our consensus panel were currently-employed public health educators in local, state, and federal health agencies. We agree that the issue of responsibility for the continuing education needs of health educators is an important one. As we pointed out, it will require both cooperation and investments on the part of professional associations, universities, government, and foundations.

We certainly agree that continuing education will require "multiple modes of delivery" and emerging technologies to reach public health professionals with the new understandings we identified. Moreover, while the evidence base for health education has matured significantly in the last 30 years, Ellery is correct to point out that much work remains to be done.

The results of our consensus panel, along with others,[1,2] constitute only an important first step. We sought to bring attention to a critical issue in strengthening the public health infrastructure.[3] There has already been significant dialogue and action.

The Centers for Disease Control and Prevention initiated and continues to support the Strategic Plan for the Development of the Public Health Workforce, including the continuing education needs of all public health professionals. Representatives of health education have been actively involved since the inception of this plan. Several health education organizations jointly have developed a 15-month long Public Health Education Leadership Institute to enhance leadership capacity to perform public health essential services.[4]

The Society for Public Health Education and the Association for the Advancement of Health Education are also engaged in joint efforts to review accreditation of health education professional preparation programs to ensure that both entry-level and advanced practice address the competencies and skills that our panel found to be critical. Finally, the Coalition of National Health Education Organizations has catalyzed multisector involvement to address the continuing education needs of the profession.[5]

The quality assurance movement in public health will have an important impact on state and federal employment policies in hiring health educators. Despite the progress we have made in credentialing, many issues will continue to be debated in the coming decade, including the question of continuing education.[6,7] We would like nothing more than to see Ellery and others who understand the importance of maintaining the best possible public health workforce take up the work of looking beyond experts and content to better identify specific training needs and to devise innovative, responsive continuing education opportunities. ■

John P. Allegrante, PhD
Robert W. Moon, MPH
M. Elaine Auld, MPH, CHES
Kristine Gebbie, DrPH, RN

About the Authors

John P. Allegrante is with the National Center for Health Education, Teachers College, and the Division of Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY. At the time of the study, Robert W. Moon was with the Montana Department of Public Health and Human Services, Helena. M. Elaine Auld is with the Society for Public Health Education, Washington, DC. Kristine M. Gebbie is with the Center for Health Policy, Columbia University School of Nursing, New York, NY.

Requests for reprints should be sent to John P. Allegrante, PhD, National Center for Health Education, Teachers College, Columbia University, 525 West 120th Street, New York, NY 10027 (e-mail: jpa1@columbia.edu).

References

1. Gebbie KM, Hwang I. Preparing currently employed public health nurses for changes in the health system. *Am J Public Health.* 2000;90:716–721.

2. Boedigheimer SF, Gebbie KM. Currently employed public health administrators: are they prepared? *J Public Health Manage Pract.* 2001;7:30-6.

3. Gebbie KM. The public health workforce: key to public health infrastructure. *Am J Public Health.* 1999; 89:660-1.

4. Wright K, Hann N, McLeroy K, et al. Health education leadership development: a conceptual model and competency framework. *Health Promot Pract.* In press.

5. Coalition of National Health Education Organizations. *The Health Education Profession in the Twenty-First Century: Progress Report 1995–2001*. Kent, Ohio: American School Health Association; 2001.

6. Livingood WC, Auld ME. The credentialing of a population-based health profession: lessons learned from health education certification. *J Public Health Manag Pract.* 2001;7:38-45.

7. Allegrante JP, Auld ME, Butterfoss FD, Livingood C. La certification en education pour la santé: son impact sur l'évolution de la profession aux États-Unis [Credentialing and its impact on the evolution of the profession of health education in the United States]. *La Santé de l'Homme.* 2001 May-June:46–53.

MISUNDERSTANDING THE WORLD HEALTH REPORT 2000

Coyne and Hilsenrath's piece on the World Health Report 2000 is a welcome contribution to the US discourse on the nature and goals of the important document from the World Health Organization.[1] However, the depth of analysis and some of the interpretation in that article need clarification and may be contested.

Their statement that "[t]he emphasis on [disability-adjusted life expectancy] can be misleading and undermines rankings for countries with low life expectancy but otherwise good health systems" is itself misleading. Disability-adjusted life expectancy is just that—it adjusts (and does not change) the national life expectancy for the duration of time spent in states of less than perfect health. In other words, it is the duration, severity, and magnitude of disability that determine the extent of this adjustment. In addition, countries with low life expectancies have a high component of premature mortality (e.g., infant and child mortality); therefore the component of life expectancy affected by disability is quite small.[2]

Judgment of a health system has to be linked with performance and the achievement of health outcomes, as in the World Health Report 2000. A good health system is expected to increase the health status of the population, and an important measure of that status is overall life expectancy. Evaluations of health systems that consider systems "otherwise good" in the absence of measurable changes in health outcomes need to be carefully interpreted, because the core mission of the health system is not being achieved.

The authors put forth a more disturbing claim, whose basis is unclear, that "[e]quity is not universally considered desirable and is difficult to achieve in heterogeneous societies." The value placed on equity has been central to global health dialogue since the start of the primary health care movement at Alma-Ata in 1978.[3] The focus on equity has also been considered a central mission of health systems at the country level.[4] It is confusing to read a statement that reduces the operational notion of justice and fairness to a less than desirable status. Moreover, the fact that equity is difficult to achieve has no merit in terms of its value as a *vision* for health systems around the world.

In describing the case of the United States, only 1 country of 191 countries considered in the global report, the authors state that advances in health technologies have not been captured in the report. First, that was never the intent of the report and therefore was not the purpose of the methods. Second, not all of the $22 billion spent on research in the United States in 1999, or the $74 billion spent on health research in the world in 1998, is spent on the production of health technologies; rather, it funds a broad mix of basic science and operational research.[5] These investments do not all result in health technologies, nor are they predictable in terms of their output.

The authors seem to confuse the measurement of a system with the next step, which is the ability to intervene. These are distinct aspects of the analysis of a health system and should not be confused, since many problems can be measured and monitored even though interventions may not be currently available. That is where the role of research and development becomes critical.

It is common knowledge that a wide variety of health determinants, such as education, lie outside the health sector—in addition to key determinants that are within the purview of the health sector. Assessing the best actions for specific health problems or for the health system as a whole will require consideration of a multisectoral approach. However, the health system can be held responsible only for those interventions within its mandate, and this was made clear in the definition of a health system at the beginning of the World Health Report 2000.

Finally, the World Health Organization established several committees in response to valid critiques of the report, and the methodology and empirical work has progressed greatly in the past 2 years. It is critical that academia challenge such global developments in a dynamic way, and yet understanding the nature and application of methods within the global context should be part of the constructive critique. ■

Adnan A. Hyder, MD, PhD, MPH

About the Author

The author is with the Bloomberg School of Public Health, Johns Hopkins University, and the Johns Hopkins Bioethics Institute, Baltimore, Md.

Requests for reprints should be sent to Adnan A. Hyder, MD, PhD, MPH, Bloomberg School of Public Health, Department of International Health, 615 N Wolfe St, Suite E-8132, Baltimore, MD 21205 (e-mail: ahyder@jhsph.edu).

References

1. Coyne JS, Hilsenrath P. The World Health Report 2000: can health care systems be compared using a single measure of performance? *Am J Public Health.* 2002;92:30, 32–33.

2. Hyder AA, Morrow RH. Applying burden of disease methods in developing countries: a case study

from Pakistan. *Am J Public Health.* 2000;90: 1235–1240.

3. Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M, eds. *Challenging Inequities in Health: From Ethics to Action.* New York, NY: Oxford University Press; 2001.

4. *Equity: An Essential Link for Health Development.* Geneva, Switzerland: Commission on Health Research for Development; 1990.

5. *Monitoring Resource Flows for Health Research.* Geneva, Switzerland: Global Forum for Health Research; 2001.

COYNE AND HILSENRATH RESPOND

We are appreciative of Hyder's thoughtful response to our piece on the World Health Report 2000. Overall, we think the methodological concerns he raises are important, and that our differences are due not only to disciplinary approach but also to practical realities of methodologies employed by the World Health Organization (WHO).

With regard to disability-adjusted life expectancy and health system performance, the intent of adjusting life expectancy is to account for time spent in less than perfect health; however, we lack confidence in disability-adjusted life expectancy because of (1) the lack of data on disability prevalence across regions (P. Musgrove, unpublished manuscript, 2002), and (2) the fact that disability-adjusted life expectancy provides insufficient additional information compared with unadjusted life expectancy.[1]

Undeniably, health outcomes are critical, but such aggregates may conceal performance, as in South Africa, where AIDS is overwhelming a medical system with many good features. The reasons for the HIV/AIDS epidemic in South Africa are complex and cannot be laid on the doorstep of the medical, or even the health, system. These reasons have deep socioeconomic roots that transcend the measures used in the World Health Report 2000.[2]

With respect to the issues of equity and technology, the *American Heritage Dictionary of the English Language* defines equity as "the state, quality, or ideal of being just, impartial, and fair." To many, a just and fair distribution of health resources and health outcomes is not necessarily an equal one. Certainly in the United States there is no consensus that all citizens are entitled to the same health benefits, and we would submit that many finance ministries around the world, which have much to say about the allocation of health resources, do not share the egalitarian notions of equity commonly found within public health circles.

Furthermore, technology should have been a dimension of efficiency and performance captured by the WHO report. The notion of dynamic efficiency is central to economic theory, and it is technology, more than any other single factor, that has raised living standards for so much of the world's population. To ignore this element, particularly at a time when health-related technologies are a driving force in economic growth, is to disregard a very important dimension of health sector efficiency and performance—a dimension that is particularly important for future generations everywhere.

In terms of interventions, undoubtedly the WHO report could have gone further in discussing interventions for the benefit of developing health policies, as noted by the editor-in-chief of the report: "What it has yet to provide is a convincing demonstration that the information constitutes evidence for guiding or assessing health policy" (P. Musgrove, unpublished manuscript, 2002).

We support the scholars' continuing to challenge cross-country comparative studies. Although we feel that the results of our critical analysis are correct, comments such as Hyder's provide a helpful platform for such challenges. Through such exchanges, the scholarly community can help to build more complete databases and analytic tools that in turn will help in the formulation of policies for improved global health. ■

Joseph S. Coyne, DrPH
Peter Hilsenrath, PhD

About the Authors

Joseph S. Coyne is with the Department of Health Policy and Administration, Washington State University, Spokane. Peter Hilsenrath is with the Department of Health Management and Policy, School of Public Health, University of North Texas Health Science Center, Fort Worth.

Requests for reprints should be sent to Joseph S. Coyne, DrPH, Department of Health Policy and Administration, Washington State University, Spokane, 310 N Riverpoint Blvd, Box H, Spokane, WA 99202-1675 (e-mail: joecoyne@wsu.edu).

References

1. Jamison DT, Sandbu, ME. Global health: WHO ranking of health system performance. *Science.* 2001; 293:1595–1596.

2. Coyne JS, Hilsenrath P. Health system performance in South Africa vs. France: World Health Report 2000. *Safundi: The Journal of South African and American Comparative Studies.* 2001;3:1.

NAVARRO RESPONDS

I agree with Hyder's normative statement that equity should indeed be considered desirable in health policy. As Coyne and Hilsenrath indicate, however, such a goal is not universally accepted, as evidenced by the growing health inequalities between and within countries. But Coyne and Hilsenrath are just plain wrong when they justify (at least in part) the existence of those inequalities while indicating that equity is difficult to achieve in heterogeneous societies. Actually, during the 1950s, 1960s, and 1970s, China, which has more ethnic diversity and a larger population than India, had lesser regional and class health differentials and better health indicators than India, even though its GNP per capita was lower.[1]

Hyder is right when he calls for evaluating health systems' performance by looking at their impact on health outcomes, but he is wrong in believing that mortality is a good outcome measure of the effectiveness of health care. At least in the developed countries, where most morbidity is chronic, medical care does more caring than curing, and thus its impact on curing and on reducing mortality is small. The overall level of mortality in a country depends on many other interventions besides health care interventions. ■

Vicente Navarro, MD, PhD, DrPH

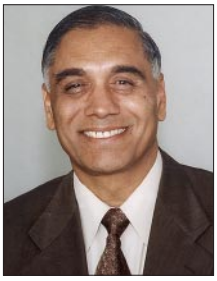
About the Author

Requests for reprints should be sent to Vicente Navarro, MD, PhD, DrPH, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, Baltimore, MD 21205 (e-mail: vnavarro@jhsph.edu).

Reference

1. Navarro V. Has socialism failed? *Int J Health Sev.* 1992;22:583–601.

EDITOR'S CHOICE



Ethics in Public Health

This month the Journal opens a forum for discussion on ethics in public health research and practice. As argued by James Thomas and his colleagues in their lead editorial, at the crux of public health ethics is "the need to exercise power to ensure the health of populations and at the same time to avoid potential abuses" (p1057). How are we to ensure that all populations benefit from recent advances in treatments, therapies, and technologies? The authors in this month's Health Policy and Ethics Forum review the heightened ethical scrutiny of public health research in the United States, pose challenging questions regarding research collaborations between developed and developing countries, and suggest mechanisms for fostering a culture of ethics in public health that will both engender public trust and reduce health inequities.

Sarah Putney and Sofia Gruskin open the forum by providing background and a thoughtful explanation for the dramatic "shutdowns" of US-based institutions by the federal Office for Human Research Protections and its predecessor, the Office of Protection Against Research Risks (page 1067). Next, Leonard Glantz challenges the public health research community to "closely examine, understand, and try to address the legitimate legal and ethical issues that so concerned [the only 2 courts that have addressed nontherapeutic research with nonconsenting subjects]." (p1073)

In a companion paper, Anna Mastroianni and Jeffrey Kahn provide a lucid overview and critical analysis of *Grimes v Kennedy Krieger Institute, Inc,* a legal case that has "sent shockwaves" through the public health and environmental justice communities (page 1073). Indeed, the Maryland Court of Appeals went so far as to compare the Kennedy Krieger Study (1993–1995), designed to measure the effectiveness of differing levels of lead abatement in housing, to the infamous Tuskegee Syphilis Study (1932–1972), wherein African American men were denied effective treatment in order to assess the natural history of their disease. While some may take issue with this analogy, it is hard to sidestep its significance. According to Mastroianni and Kahn, "Real exploitation is obviously unacceptable, and perceived exploitation serves to undermine trust in research and researchers." (p1073)

This is true not only in the United States, but throughout the world, as the last 2 forum contributions make clear. Despite the tremendous potential and promise of the so-called "genomic revolution," Tikki Pang warns, "Genomics brings with it complex new ethical, legal, social, and economic implications." (p1077) To combat global inequities in health care and benefit all concerned, he calls for the development of creative and effective research partnerships between the developed and developing world. Leslie London believes that if this goal is to be achieved, it is imperative to recognize the agency of vulnerable groups, communities, and countries in the ethical review process; otherwise, they will remain passively "in need of protection" instead of being the managers of their own health (page 1079).

In the wake of the events of September 11, controversy exists regarding the proposed recommendations in the Model State Emergency Powers Act. The prospect of unprecedented police and public health powers in the face of suspected bioterrorist activities without appropriate checks and balances has raised public fears, for example, that some populations will be quarantined while others receive medical services. In an approval process accelerated in response to these threats, APHA became the first public health organization to adopt the "Principles of Ethical Public Health Practice," published in this issue (page 1058). Although these principles alone cannot ensure the protection of populations or hold agencies accountable for abuses of power, they aim to ensure an ethical basis for public health practice. Public trust will be won only if effective protection is afforded to all. ■

Mohammad N. Akhter, MD, MPH
Executive Director, APHA
Mary E. Northridge, PhD, MPH
Editor-in-Chief

EDITORIAL

A Code of Ethics for Public Health

The mandate to ensure and protect the health of the public is an inherently moral one. It carries with it an obligation to care for the well-being of communities, and it implies the possession of an element of power to carry out that mandate. The need to exercise power to ensure the health of populations and, at the same time, to avoid abuses of such power are at the crux of public health ethics.

Until recently, the ethical nature of public health has been implicitly assumed rather than explicitly stated. Increasingly, however, society is demanding explicit attention to ethics. This demand arises from technological advances that create new possibilities and, with them, new ethical dilemmas; new challenges to health, such as the advent of HIV; and abuses of power, such as the Tuskegee study of syphilis.

Medical institutions have been more explicit about the ethical elements of their practice than have public health institutions. However, the concerns of public health are not fully consonant with those of medicine. Thus, we cannot simply translate the principles of medical ethics to public health. In contrast to medicine, public health is concerned more with populations than with individuals, and more with prevention than with cure. The need to articulate a distinct ethic for public health has been noted by a number of public health professionals and ethicists.[1–5]

A code of ethics for public health can clarify the distinctive elements of public health and the ethical principles that follow from or respond to those elements. It can make clear to populations and communities the ideals of the public health institutions that serve them, ideals for which the institutions can be held accountable.

THE PROCESS OF WRITING THE CODE

The backgrounds and perspectives of people who identify themselves as public health professionals are as diverse as the multitude of factors affecting the health of populations. Articulating a common ethic for this diverse group is a formidable challenge. In the spring of 2000, the graduating class of the Public Health Leadership Institute chose writing a code of ethics for public health as a group project. The institute provides advanced leadership training to people who are already in leadership roles in public health. Because the fellows bring a wealth of experience from a wide variety of public health institutions, they are uniquely able to represent diverse perspectives and identify ethical issues common in public health.

At the 2000 meeting of the National Association of City and County Health Officers, the group added a non-institute member (J.C. Thomas) and charted a plan for working toward a code. The plan included receiving a formal charge as the code of ethics working group at the annual meeting of the American Public Health Association (APHA); reviewing codes written by other organizations, particularly those within public health (the American College of Epidemiology and the Society of Public Health Education); and balancing open participation with efficiency in writing the code.

The latter aim was achieved by having a small number of people write an initial code, then inviting feedback on it and each successive version from progressively broader audiences. The audiences reacting to the code drafts were (1) the working group itself; (2) an additional 19 ethicists and representatives from various public health agencies gathered in a meeting at the University for Health Sciences in Kansas City to critique the code; and (3) APHA members (via the APHA Web site, where the code was posted and feedback was solicited, and the 2001 annual meeting).

THE CONTENT OF THE CODE

The consensus reached during the review process was that while people outside the public health establishment might find the code useful, it should be directed to those in traditional public health institutions, including public health departments and schools of public health. Similarly, while people working in public health throughout the world may find the code helpful, it was written with the American public health system in mind. Although touching on aspects of research, the focus of the code is principally on public health practice.

While acknowledging the value of a set of principles for individuals, and the fact that institutional policies are often carried out by individuals, the working group wrote the code for institutions. One reason was the definition of

public health first articulated in the Institute of Medicine report *The Future of Public Health* and used in the code: "What we, as a society, do collectively to assure the conditions for people to be healthy."[6] Others have also noted that one of the differences between public health and medicine is that public health is most often delivered by government institutions to a population rather than by one person to another.[3]

The writers of the code aimed for a document that could fit on a single page and be easily posted. This concise statement of 12 ethical principles (box on this page) is accompanied by a series of other documents, including a preamble that explains the purpose of the code; a list of 14 values and beliefs inherent to a public health perspective that underlie the ethical principles; and notes on the ethical principles to more fully explain their intent. (All of the components are posted on the Web, and are available at http://www.apha.org/codeofethics.)

Reviewers of the code preferred positive rather than negative wording of the ethical principles. For example, the principle addressing conflicts of interest (number 12) is worded as an affirmation of collaboration with the proviso that it be done in a way that enhances the public's trust in the institutions.

The code draws upon several ethical concepts. The more individualistic notion of human rights appears in the second principle as a necessary point of tension with the communitarian concern for the well-being of communities. Theories of distributive justice underlie the fourth principle, which speaks of the need for basic resources and conditions necessary for health among the disenfranchised. Duty as an ethical motivation is represented in several of the principles, such as the obligations to provide information in some instances and to protect it in others.

One of the beliefs inherent to a public health perspective is that each person both affects and depends upon others. This interdependence between humans underlies the most fulfilling aspects of relationships and community as well as conflicts between people. Interdependence is the complement to autonomy, a dominant principle in medical ethics. Without denying that individuals have a right to some role in decisions that affect them, a recognition of interdependence serves as a correction to an overly individualistic perspective that is inconsistent with public health's concern with whole communities and populations.

The principle of interdependence between individuals lies behind the preeminence given to the health of communities in the 2nd principle of the code. Interdependence between institutions and the need for collaboration underlies the 12th principle, and the interdependence inherent to ecological systems underlies the 9th principle, which addresses the physical and social environments.

Principles of the Ethical Practice of Public Health

1. Public health should address principally the fundamental causes of disease and requirements for health, aiming to prevent adverse health outcomes.
2. Public health should achieve community health in a way that respects the rights of individuals in the community.
3. Public health policies, programs, and priorities should be developed and evaluated through processes that ensure an opportunity for input from community members.
4. Public health should advocate for, or work for the empowerment of, disenfranchised community members, ensuring that the basic resources and conditions necessary for health are accessible to all people in the community.
5. Public health should seek the information needed to implement effective policies and programs that protect and promote health.
6. Public health institutions should provide communities with the information they have that is needed for decisions on policies or programs and should obtain the community's consent for their implementation.
7. Public health institutions should act in a timely manner on the information they have within the resources and the mandate given to them by the public.
8. Public health programs and policies should incorporate a variety of approaches that anticipate and respect diverse values, beliefs, and cultures in the community.
9. Public health programs and policies should be implemented in a manner that most enhances the physical and social environment.
10. Public health institutions should protect the confidentiality of information that can bring harm to an individual or community if made public. Exceptions must be justified on the basis of the high likelihood of significant harm to the individual or others.
11. Public health institutions should ensure the professional competence of their employees.
12. Public health institutions and their employees should engage in collaborations and affiliations in ways that build the public's trust and the institution's effectiveness.

DISSEMINATION AND ADOPTION OF THE CODE

For the code to be truly useful it must be broadly disseminated and adopted by public health institutions. Adoption by key national agencies and organizations will imbue the code with a degree of moral authority that will increase both its utility and the likelihood that it will be adopted and used by national, state, and local institutions. On February 26, 2002, the APHA Executive Board formally adopted the code, making APHA the first national organization to do so. This endorsement provides the code of ethics working group with an important tool for talking about adoption with other organizations and agencies, such as the Centers for Disease Control and Prevention, the National Association of City and County Health Officers, the Association of State and Territorial Health Officials, and the Association of Schools of Public Health. Members of these institutions contributed to the creation of the code, which should help with the next step of adoption.

Once a government agency or professional organization adopts the code, it will need to build these ethical principles into its policies and procedures, to the extent that it has not already done so, and train its employees in ways that ensure the implementation of the principles. Schools of public health should teach the code to their students. Since many public health professionals do not have a formal degree in public health, there will also be a need for continuing education or extension courses that include the code of ethics and how to use it.

For each of these tasks there will be a need for new tools. These might include materials for teaching the code, such as case studies illustrating the application of each of the 12 ethical principles; a workbook that helps

an institution consider how it might build the ethical principles into its policies and practices; and an oath to be recited by individuals as they graduate from a school of public health or as they are hired by a public health institution (the code of ethics working group is now considering writing such an oath).

FUTURE IMPROVEMENTS

The code of ethics, as it now stands, is the first explicit statement of ethical principles inherent to public health. It is a significant step forward, but it is unlikely to be the last step. Although the code was developed with broad input, we will gain new insights about its strengths and weaknesses as it is implemented. Moreover, as the world changes, public health professionals will become sensitized to new ethical issues. We anticipate, then, a time when the code will need to be updated.

To facilitate this process, the code will be posted on the Web in an interactive format that will welcome comments and will allow people to read others' comments. A standing committee of the Public Health Leadership Society will actively engage public health professionals and ethicists in the consideration of periodic updates to the code, which will incorporate lessons learned and comments received over time. In the near future, however, the code should prove to be a useful tool in clarifying the values and purposes of the public health profession and enabling it to more often achieve its high ideals. ■

James C. Thomas, MPH, PhD
Michael Sage, MPH
Jack Dillenberg, DDS, MPH
V. James Guillory, DO, MPH

About the Authors

James C. Thomas is with the Department of Epidemiology and the Program in Public Health Ethics, University of North Carolina School of Public Health, Chapel Hill. Michael Sage is with the National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Ga. Jack Dillenberg is with the School of Dentistry and Oral Health, Arizona School of Health Sciences, Phoenix. V. James Guillory is with the Department of Preventive Medicine and Division of Research, University of Health Sciences, Kansas City, Mo.

Requests for reprints should be sent to James C. Thomas, MPH, PhD, 2104-B McGavran-Greenberg Hall, CB#7435, School of Public Health, University of North Carolina, Chapel Hill, NC 27599-7035 (e-mail: jim.thomas@unc.edu).

This editorial was accepted April 10, 2002.

Acknowledgments

The writing of the code of ethics was funded in part by the Centers for Disease Control and Prevention, both directly and through funding of the Public Health Leadership Society.

The authors thank the other members of the code of ethics working group: Elizabeth Bancroft (Centers for Disease Control and Prevention, Los Angeles County), Kitty Hsu Dana (APHA), Joxel Garcia (Connecticut Department of Health), Kathleen Gensheimer (Maine Department of Health), Teresa Long (Columbus, Ohio, Department of Health), Ann Peterson (Virginia Department of Health), Liz Schwarte (Public Health Leadership Society), Kathy Vincent (Alabama Health Department), and Carol Woltring (Center for Health Leadership and Practice, Oakland, Calif).

References

1. Mann JM. Medicine and public health, ethics and human rights. *Hastings Center Rep.* 1997(May-Jun);27:6–13.

2. Beauchamp D. Community: the neglected tradition of public health. In: Beauchamp D, Steinbock B, eds. *New Ethics for the Public's Health.* New York, NY: Oxford University Press; 1999.

3. Kass NE. An ethics framework for public health. *Am J Public Health.* 2001; 91:1776–1782.

4. Callahan D, Jennings B. Ethics and public health: forging a strong relationship. *Am J Public Health.* 2002;92: 169–176.

5. Roberts MJ, Reich MR. Ethical analysis in public health. *Lancet.* 2002; 359:1055–1059.

6. Institute of Medicine. *The Future of Public Health.* Washington, DC: National Academy Press; 1988.

Critical Policy Challenges in the Third Decade of the HIV/AIDS Epidemic

Numerous policy challenges continue to face the United States in the third decade of the HIV/AIDS pandemic, in both the health and foreign policy arenas. They include longstanding questions about care, treatment, prevention, and research, as well as new ones introduced by the changing nature of the epidemic itself and the need to balance demands for limited resources.

These challenges concern the United States not only in its role as a world leader in combating a global epidemic, but in its decisions and focus at home, where the epidemic continues to take a toll. (*Am J Public Health.* 2002;92:1060–1063)

Jennifer Kates, MPA, MA, Richard Sorian, Jeffrey S. Crowley, MPH, and Todd A. Summers

THE XIV INTERNATIONAL AIDS Conference will take place in Barcelona, Spain, in July 2002. This year's conference is notable on several fronts: it follows the Durban Conference, the first international AIDS conference to be held in a developing country deeply affected by the epidemic; it comes one year after an unprecedented special session of the United Nations General Assembly on HIV/AIDS; and it is the first international conference to take place since the events of September 11, 2001.

The policy challenges facing the United States in the third decade of the pandemic are both long-standing—such as questions about care, treatment, prevention, and research—and new—including the challenges introduced by the changing nature of the epidemic itself and the need to balance demands for limited resources. The United States must meet these challenges both at home, where the epidemic continues to take a toll, and on the global front, where US leadership is needed to help combat the epidemic.

THE US EPIDEMIC

Since the beginning of the epidemic in the United States, close to 800 000 AIDS cases have been reported, and more than 450 000 people have died.[1] Nationally, 850 000 to 950 000 Americans are estimated to be living with HIV/AIDS.[2] While HIV/AIDS is a national epidemic, it has had an especially severe impact on certain groups, including gay and bisexual men, injection drug users and their sexual partners, young people, and racial and ethnic minorities. The epidemic is also increasingly affecting women and economically disadvantaged communities.[3] In addition, recent data suggest that the era of sharp declines in AIDS deaths and new AIDS diagnoses, brought on by the introduction of better therapies in the mid-1990s, may have come to an end.[4] Within this context, there are several critical challenges.

Reducing New Infections

Efforts to raise awareness about HIV/AIDS and change risky behaviors have helped to slow the number of new HIV infections in the United States from more than 150 000 per year in the mid-1980s to 40 000 today. Yet the United States has continued to experience about 40 000 new infections each year since the early 1990s.[5] The recent stabilization in the number of AIDS cases and deaths is also cause for concern.

A key aspect of HIV prevention is the frequent collision between politics and public health science. Prevention interventions have historically been mired in controversy, owing in part to the fact that HIV transmission involves sex and drugs, subjects with which many—policymakers included—are uncomfortable. This discomfort, and the absence of a national consensus, has affected the use of proven strategies for reducing the number of new infections, including targeting at-risk populations and those who are HIV-positive with tailored, culturally specific interventions[6]; reducing stigma, given that stigma may contribute to risky behavior and affect individuals' willingness to get tested or seek care[5,7–10]; integrating prevention into the clinical care setting[11]; and implementing syringe exchange as part of comprehensive prevention programs for injection drug users.[12] There is also a need for continued research to develop new behavioral and clinical prevention strategies, including topical microbicides and vaccines.

Increasing the Number of People Who Know Their Status

About 400 000 to 500 000 people with HIV/AIDS in the United States remain undiagnosed, untreated, or both, and are therefore not receiving the treatments that could forestall disease progression or the prevention supports needed to avoid passing the virus to others.[2] Many continue to face economic barriers to care, lacking insurance coverage to help them afford the high cost of HIV care, which can average as much as $20 000 a year. Increasingly, HIV affects those who are poor, are outside the workforce, and have a history of barriers to access. Even among individuals who have some resources, the high cost of HIV care can

quickly exhaust their assets and leave them impoverished.[13]

As a result, people with HIV/AIDS increasingly rely on the public sector for care, primarily Medicaid, Medicare, and the Ryan White CARE Act.[14] One major barrier to Medicaid coverage is a catch-22 in eligibility—most low-income people with HIV must wait until they become disabled by AIDS to be eligible for coverage of treatments that can prevent disability.

Strategies for addressing these issues include increasing the number of people who know their HIV status by providing more information to the public and at-risk populations about voluntary HIV counseling and testing; using new testing technologies, such as rapid testing, to better target those most at risk; furthering efforts to reduce stigma and discrimination; and increasing access to care and coverage for people with HIV/AIDS through expansions of public and private coverage. For example, Congress is considering the Early Treatment for HIV Act, which would address Medicaid's catch-22 by creating a new state option to expand Medicaid coverage to low-income people with HIV who are not yet disabled.[15,16]

Addressing the Impact of HIV in Minority Communities

HIV/AIDS disproportionately affects racial and ethnic minorities, as it has since the beginning of the epidemic. HIV is the leading cause of death for African Americans between the ages of 25 and 44 years and the third leading cause of death for Latinos in this age group.[17] People of color now represent the majority of new HIV infections (74%) and people living with AIDS (62%).[1]

The increasing concentration of the epidemic among minority Americans is due to many complex factors, including social inequalities related to income and race and stigma associated with being gay or bisexual, which exists within minority communities as well as in the larger society. These contextual forces may operate at the individual level to increase high-risk behaviors or at the societal level by compromising community infrastructure for responding to the epidemic. There is a critical need to better understand where and why these disparities occur, what factors affect receptivity to prevention messages and health care access, and whether public programs, particularly Medicaid and the Ryan White CARE Act, are adequately serving people of color. Understanding the views of minority leaders and communities toward HIV/AIDS is essential to an informed response.[18–20] The Minority HIV/AIDS Initiative, adopted by Congress in 1999 after much community pressure, has been one attempt to enhance community capacity to respond to HIV/AIDS.[21]

Addressing Rising Drug Costs

Spending on prescription drugs is one of the fastest growing components of US health care spending,[22] and spending on HIV-related therapies is no exception. Because access to medications is critical for people with HIV/AIDS and these drugs are expensive, concerns have been raised about rising expenditures and the price of prescription drugs. Policymakers are faced with several questions: Are there mechanisms for purchasing drugs at lower prices, such as purchasing in bulk or through rebate programs? Should government be involved in limiting or controlling drug prices? Should the public and private sectors' respective investments in drug research be considered in determining drug pricing? Are there ways to use existing resources more efficiently, such as purchasing or continuing private insurance coverage for people with HIV?

Stimulating Research and Development

Despite significant public investment and progress in HIV research, there is still no cure for HIV and no vaccine against the virus, and available treatments, while effective for many, do not help everyone and often have severe side effects. There is a great deal to learn about how to use existing pharmaceuticals safely and appropriately; about long-term toxicities of the multiple medications that are being prescribed for people living with HIV; and about the development of drug resistance. Priority research areas include vaccine development, prevention, microbicides, and therapeutic research.

Policymakers are faced with a complex array of decisions and choices concerning research and development: What is the role of the federal government in conducting therapeutics research vis-à-vis private pharmaceutical and biotechnology companies? What is the best way to allocate public research dollars for basic science research vs clinical research? Are public dollars—or public policies—leading to research that can answer some of the questions about long-term toxicities, resistance, and so forth? Since barriers prevent private firms from aggressively conducting vaccine research, should federal policymakers fund this research directly or provide incentives for private research (e.g., through tax credits)?

Maintaining Attention to the US Epidemic

After 2 decades of fighting the HIV/AIDS epidemic in the United States, it is not surprising that there may be some signs of "AIDS fatigue." For example, although Americans still rate AIDS as a top health concern for the nation, the proportion who see it as the number one health problem has declined over the past few years.[8] In addition, a recent report on the role of private philanthropy in responding to the epidemic found that while philanthropic support of global AIDS efforts is on the rise, support for domestic efforts has not grown.[23] Yet as the US epidemic continues to exact an increasing toll on racial and ethnic minority communities, maintaining attention to the epidemic at home remains critical, even as the global response gains attention.

THE GLOBAL EPIDEMIC

Worldwide, more than 60 million people have been infected with HIV, and 20 million have died. HIV is now the leading cause of death in Africa and the fourth leading cause of death worldwide. Most of the impact has been felt in the developing world. Young people and women are increasingly at risk.[24,25] In addition, it is estimated that more than 40 million children will have lost one or both parents to HIV/AIDS by 2010, and these children will also be at increased risk for

HIV.[26] The United States faces several challenges in addressing this global epidemic.

Identifying Appropriate Forms of US Assistance

The United States allocates funding and other resources used to address the global epidemic in several ways, including direct financial assistance to other countries, support for multilateral organizations such as the Joint United Nations Program on HIV/AIDS (UNAIDS), and broader forms of development assistance. This assistance goes toward a variety of activities, including direct prevention, care, treatment, and support services and, increasingly, impact mitigation efforts that address the larger societal consequences of the pandemic.

To date, the bulk of US foreign assistance in the fight against the global pandemic has been in the form of bilateral assistance to other nations. While the level of spending and other resources made available is clearly a fundamental component of the US response, it is also important to assess the mechanisms by which resources are allocated and their effect on recipient countries and programs. These mechanisms include US agency activity; direct assistance through government-to-government agreements and bilateral aid; contributions to multilateral programs; loans to developing countries; debt relief; and direct assistance to nongovernmental organizations. For example, since foreign debt is one of the major barriers facing developing nations' ability to respond to the epidemic, grants and debt relief may represent more viable options than loans.[27]

Shaping the Global Fund to Fight AIDS, TB, and Malaria

Total resources for addressing the HIV/AIDS epidemic in the developing world are estimated to be at least $7 to $10 billion annually.[28] In April 2001, UN Secretary-General Kofi Annan issued a call to action to create a Global Fund to fight HIV/AIDS as a mechanism for mobilizing and coordinating additional resources toward this goal. The scope of the fund was expanded to include tuberculosis and malaria and plans for the fund began later that year.[29] The first round of grants were announced in April 2002. The United States has pledged $300 million to date and has earmarked an additional $200 million for fiscal year 2003, which is awaiting congressional approval.[30]

The US government has played a critical role in shaping decisions concerning the fund, working with other governments, research and community organizations, foundations, and other private sector players. Continued leadership from the United States is needed to address ongoing challenges including mobilizing larger and sustained contributions (and articulating the appropriate role of US commitments in this regard); expediting disbursements without sacrificing oversight and accountability; establishing executive leadership and appropriate staffing; and clarifying the role of the Global Fund in the context of other global AIDS efforts (the fund is intended to represent new resource commitments, rather than funding redirected from other health and international development efforts).[29,31]

Balancing Priorities

Research, care and prevention are integral components of an effective global or national HIV/AIDS strategy, and understanding the often complex relationships between them is critical. To date, the majority of US government spending on HIV/AIDS in developing countries has been for prevention, with few resources allocated to care.[32] As the United States seeks to promote an integrated approach to the global pandemic, it will need to look at ways to foster public–private partnerships that support prevention and care, including the provision of antiretroviral therapies (to prevent mother-to-child transmission and to treat those who are living with HIV) and research in developing countries.

The issue of health care infrastructure is fundamental to these considerations. Definitions of infrastructure include such elements as the availability of health centers, facilities such as laboratories, and trained personnel; roads, equipment, supply systems, and water; security; and stability of government. There has been some reluctance on the part of the United States, other nations, and the private sector to provide increased or new support for certain interventions in developing countries because of concerns about existing infrastructure. There is a need to improve the understanding of the definition and role of infrastructure in delivery of prevention and treatment interventions in resource-poor settings and to identify ways to support infrastructure enhancements. It will be important to gain experience in implementing infrastructure development initiatives, assessing the level of infrastructure needed for different types of interventions and insuring the capacity of indigenous institutions.

Promoting Access to Treatment

The last couple of years have witnessed important progress in removing barriers to access to treatment for people living with HIV in developing countries. Nonetheless, the cost of antiretroviral and other medications far exceeds what is affordable for most individuals in these countries, raising concerns about the need to balance intellectual property rights protections with greater access to medications and propelling the discussion into the realm of US and global trade policy. Within this context, several strategies are being explored to enhance access to treatment, including the purchasing of generic drugs, bulk purchasing, parallel importation, compulsory licensure, and tiered pricing. In addition, UNAIDS and individual nations have worked with several pharmaceutical manufacturers to forge price reduction arrangements for antiretroviral and other HIV-related medications.[33]

CONCLUSION

Taken together, these challenges are formidable. Meeting them will require resources and leadership. Resources for the epidemic have always competed with other national priorities but generally have fared well on Capitol Hill. Still, total US support for international AIDS efforts represents a smaller proportion of gross national product for the United States than for many other wealthy nations.[34] In addition, the President's fiscal year 2003 budget proposes flat

funding for US prevention efforts and the Ryan White CARE Act.[30] Resources, then, remain a key, overriding challenge, underscoring the need to demonstrate that a response to HIV/AIDS is connected to numerous other areas, including national security.

There are no easy choices. Yet, as UNAIDS' Peter Piot recently noted, "the AIDS epidemic is different from any other epidemic the world has faced, and as such, requires a response from the global community that is broader and deeper than has ever before been mobilized against a disease."[35] The United States continues to be in a position to provide critical leadership to such a response. ■

About the Authors

Jennifer Kates is with the Henry J. Kaiser Family Foundation, Washington, DC. At the time of writing, Richard Sorian was with the Institute for Health Care Research and Policy, Georgetown University, Washington, DC. Jeffrey S. Crowley is with the Institute for Health Care Research and Policy, Georgetown University, Washington, DC. Todd A. Summers is with Progressive Health Partners, Washington, DC.

Requests for reprints should be sent to Jennifer Kates, MPA, MA, Kaiser Family Foundation, 1450 G St, NW, Suite 250, Washington, DC 20005 (e-mail: jkates@kff.org).

This article was accepted March 18, 2002.

Acknowledgments

We would like to acknowledge the contributions to our thinking of Drew Altman, PhD; Diane Rowland, ScD; and Marsha Lillie-Blanton, DrPh, of the Kaiser Family Foundation, Christopher Collins of the AIDS Vaccine Advocacy Coalition, amd David Winters of the Ford Foundation.

References

1. Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report*, 2001; 13 (1).

2. Fleming PL, Byers RH, Sweeney PA, Daniels D, Karon JM, Janssen RS. HIV prevalence in the United States, 2000. In: Proceedings of the 9th Conference on Retroviruses and Opportunistic Infections; February 24-28, 2002; Seattle, Wash. Abstract 11, Oral Abstract Session 5.

3. Karon JM, Fleming PL, Steketee RW, De Cock KM. HIV in the United states at the turn of the century: an epidemic in transition. *Am J Public Health.* 2001;91:1060–1068.

4. AIDS cases and deaths hold stable for second year [press release]. Atlanta, Ga: Centers for Disease Control and Prevention; August 13, 2001.

5. *HIV Prevention Strategic Plan Through 2005.* Atlanta, Ga: Centers for Disease Control and Prevention; 2001.

6. *Compendium of HIV Prevention Interventions With Evidence of Effectiveness.* Atlanta, Ga: Centers for Disease Control and Prevention; 1999.

7. Centers for Disease Control and Prevention. HIV incidence among young men who have sex with men—seven U.S. cities, 1994–2000. *MMWR Morb Mortal Wkly Rep.* 2001;50(21).

8. *The AIDS Epidemic at 20 Years: The View From America.* Menlo Park, Calif: Kaiser Family Foundation; 2001.

9. Valdiserri RO. HIV/AIDS stigma: an impediment to public health. *Am J Public Health.* 2002;92:341–342.

10. Herek GM, Capitanio JP, Widaman KF. HIV-related stigma and knowledge in the United States: prevalence and trends, 1991–1999. *Am J Public Health.* 2002;92:371–377.

11. Institute of Medicine. *No Time to Lose: Getting More From HIV Prevention.* Washington DC: National Academy Press; 2000.

12. *Evidence-Based Findings on the Efficacy of Syringe Exchange Programs: An Analysis for the Assistant Secretary for Health and Surgeon General of the Scientific Research Completed Since April 1998.* Washington, DC: US Dept of Health and Human Services; 2000.

13. *Financing HIV/AIDS Care: A Quilt With Many Holes.* Menlo Park, Calif: Kaiser Family Foundation; 2000.

14. *Federal HIV/AIDS Spending: A Budget Chartbook, FY 2000.* Menlo Park, Calif: Kaiser Family Foundation; 2000.

15. Early Treatment for HIV Act of 2001, 107th Cong, 1st sess, HR 2063 IH and S 987 IS (introduced June 5, 2001).

16. Kahn JG, Haile B, Kates J, Chang S. Health and federal budgetary effects of increasing access to antiretroviral medications for HIV by expanding Medicaid. *Am J Public Health.* 2001;91:1464–1473.

17. National Center for Health Statistics. Deaths: leading causes for 1999. *Natl Vital Stat Rep.* October 12, 2001; 49(11).

18. Friday JC, Lee DH, Lillie-Blanton M, Weinstock B, Kates J. A survey of black elected officials on HIV/AIDS in the African American community. *Minority Health Today.* April 2002(suppl).

19. *African Americans' Views of the HIV/AIDS Epidemic at 20 Years: Findings From a National Survey.* Menlo Park, Calif: Kaiser Family Foundation; 2001.

20. *Latinos' Views of the HIV/AIDS Epidemic at 20 Years: Findings From a National Survey.* Menlo Park, Calif: Kaiser Family Foundation; 2001.

21. About the initiative. Available at: http://www.omhrc.gov/omh/aids/about/abt_toc.htm. Accessed March 6, 2002.

22. *Prescription Drug Trends: A Chartbook Update.* Menlo Park, Calif: Kaiser Family Foundation; 2001.

23. Funders Concerned About AIDS. Voices from the field: remobilizing HIV/AIDS philanthropy for the 21st century. 2001. Available at: http://www.fcaaids.org/spotlight/010622.html. Accessed May 6, 2002.

24. *AIDS Epidemic Update.* Geneva, Switzerland: UNAIDS; 2000.

25. *AIDS Epidemic Update.* Geneva, Switzerland: UNAIDS; 2001.

26. *Children on the Brink.* Washington, DC: US Agency for International Development; 2000.

27. Phillips M. Treasury's O'Neill pushes plan to give grants to poor nations. *Wall Street Journal.* February 21, 2002.

28. *Fact Sheet: HIV/AIDS Financing Gap.* Geneva, Switzerland: UNAIDS; February 2002.

29. The Global Fund to Fight AIDS, Tuberculosis, and Malaria. Available at: http://www.globalfundatm.org/. Accessed March 6, 2002.

30. HHS budget for HIV/AIDS increases 8 percent [press release]. Washington, DC: US Dept of Health and Human Services; February 4, 2002.

31. *The Global Fund to Fight AIDS, TB, and Malaria: Successes and Challenges.* Washington, DC: Center for Strategic and International Studies; 2002.

32. USAID combatting HIV/AIDS: a record of accomplishment [press release]. June 22, 2001. Available at: http://www.usaid.gov/press/releases/2001/fs010420.html. Accessed March 6, 2002.

33. UN efforts broaden availability of antiretrovirals: accelerating access initiative moving forward; 72 countries worldwide express interest [press release]. Geneva, Switzerland: UNAIDS/WHO; December 11, 2001.

34. *Global Spending on HIV/AIDS: Tracking Public and Private Investments in AIDS Prevention, Care, and Research.* Menlo Park, Calif: Kaiser Family Foundation; 2001.

35. *Hearings Before the Committee on Foreign Relations of the United States Senate,* 107th Cong, 2nd sess (February 13, 2002) (testimony of Peter Piot, "Halting the Global Spread of HIV/AIDS: the Future of US Bilateral and Multilateral Responses").

Ethics in Public Health Research

Time, Place, and Consciousness: Three Dimensions of Meaning for US Institutional Review Boards

Sarah B. Putney, JD, and Sofia Gruskin, JD, MIA

In the past few years, US federal agencies governing research with human subjects and institutional review boards have taken a higher-profile path than ever before, both at home and internationally. This trend carries profound significance for US-based institutions and has implications also for the rest of the world.

What does this critical moment of heightened federal scrutiny mean for the workings of US institutional review boards? We examined board activity across 3 dimensions: time, place, and consciousness. We conclude that although institutions in all areas of biomedical and social science research are adapting their practices, the field of public health is especially well positioned to adapt to, and succeed in, new efforts to ensure protection of human research subjects. (*Am J Public Health.* 2002;92:1067–1070)

TIME

RECENTLY, DRAMATIC

shutdowns of research at US-based institutions by the Food and Drug Administration (FDA), the federal Office for Human Research Protections (OHRP), and OHRP's predecessor, the Office of Protection against Research Risks, have brought heightened attention to human subject protections. The institutional assurance mechanism is central to how a shutdown can happen.

Institutions are licensed to use federal funds to conduct research with human subjects through an assurance that is granted on the condition that the institution abide by certain terms. The funding agency can partially restrict or suspend the assurance if there is a failure in compliance on the part of the grantee. Suspension includes revoking the privilege of using federal funds to conduct research with human subjects.[1,2] Institutions are given an opportunity to remedy systemic failures that may be affecting institutional review board (IRB) operations before any suspension or shutdown occurs. The self-assessment that led to the drafting of this article was, in fact, catalyzed by an OHRP investigation into Harvard School of Public Health genetic epidemiological studies conducted in urban and rural China.

The tragedy of human deaths prompted shutdowns at the University of Pennsylvania[3] and Johns Hopkins.[4] These public punishments sent shock waves through US institutions and gave IRBs, at these institutions and elsewhere, increased internal institutional attention by providing real examples of the willingness of the government to exercise its power to secure the integrity of the system. Survivors of shutdowns are now striving to demonstrate, within their institutions and beyond, the proper ways to conduct IRB business. At IRB conferences, these individuals can be found speaking on panels with representatives of OHRP, delivering the message that the time to conduct quality improvement is *now.*

Curiously, the urgency that manifests in institutional responsiveness to the shock of shutdowns expresses the redemptive quality of IRB work. Indeed, IRBs owe their existence to efforts to reckon with the most notorious abuses of power in human studies in recent generations—the Nazi doctors, the Tuskegee Syphilis Study—as well as many lesser-known studies gone wrong.[5] The history of each of these studies is complex, raising questions about investigators' intentions, assumptions, and conflicts between conscience and rationalization, which belie the easy moral authority of hindsight.[6,7] Yet a common element among them is that researchers acted either without a legal system for enforcement of ethical behavior (as with the Nazis and the Tuskegee Study) or in spite of one.

The 1970s marked a passage into a new era of efforts to do better by human subjects, especially those recognized to be vulnerable. In 1972 the Tuskegee study was finally terminated by the US Department of Health, Education, and Welfare after the public responded to media exposure.[8]

In 1974 the federal regulations were promulgated,[9] and in 1978, the Belmont Report was issued following the work of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.[9] (The National Commission was created by the National Research Act [Pub L 93-348] in 1974. The federal regulations known collectively as the Common Rule [45 CFR §46.101 et seq] superseded the 1971 guidelines of the Department of Health, Education, and Welfare.) The federal regulations protect certain categories of "vulnerable populations" through special provisions: pregnant women, fetuses, and neonates in subpart B, prisoners in subpart C, and mi-

nors in subpart D of Title 45, CFR §46. Other groups considered vulnerable to coercion and undue influence, who do not have the benefit of a regulatory subpart, include persons with cognitive impairment, persons with low literacy skills, and the poor and politically disenfranchised (see 45 CFR §46.111(b)).

Although the public has been shaken by shutdowns and institutions complain about the regulatory burdens, the IRB system is likely to endure. It remains an appealing model because it expresses aspirational values such as individual autonomy and justice. Its operations are meant to be conducted in quasi-independence from its institutional parent and driven by conscience and rational debate. An IRB has a reasonably democratic, jury-like character. As prescribed by the Common Rule, it must consist of no fewer than 5 members, including both men and women. These members must possess both scientific and nonscientific skills and expertise and must reflect the ethnic or cultural diversity of the local community at large.[10,11] At least one member must be unaffiliated with the institution.[12]

The system requires that members, like a jury or an electorate, inform themselves, debate issues, and vote.[13] A simple majority prevails. Conflicts of interest require disclosure and recusal.[14,15] An IRB must explain its reasoning to the affected investigator if it votes to disapprove a protocol, and the investigator must be given the opportunity to respond.[16] In practice, IRBs are highly collaborative bodies; rarely is a decision made by only one person without the input of others.

"The time is now" means that now, IRBs have the responsibility to give painstaking attention to every protocol, documenting compliance rigorously. Yet urgency can meet logistical blocks—shortages of staff, space, money—that can translate into delays and crises. The costs involved in running IRBs effectively are significant, and the financial implications for institutions with limited resources have not been sufficiently considered.

Unless an industrial sponsor can be tapped (usually done through a contract budget item), institutions theoretically pay the costs of IRBs out of overhead revenue. Currently, grants do not allow for IRB costs as separate from indirect costs. The National Institutes of Health (NIH) has issued a request for applications in which it announced that $28.5 million in grant funding is available to improve institutional IRBs.[17] While this is apparently only a one-time offer, it may be of immediate use in helping to strengthen existing mechanisms and processes for eligible institutions.

This financial offer may well signal an increased effort toward collaboration and outreach by federal regulators. OHRP director Greg Koski introduced in 2001 the motto "Doing it right . . . together," which has become a hallmark of his directorship. Since then, he has rolled out a number of programs to activate the partnering role of the government, including a quality improvement unit,[18] a simpler assurance system, increased staff, and more town meetings and workshops.[19] The challenge for IRBs is, and will continue to be, "doing it right" manageably, in a way that satisfies all parties: human subjects, federal regulators, investigators, sponsors, and scientific reviewers. Doing it right means at a minimum that nothing bad happens, an outcome that is tricky to measure. The redemptive plays out as preemptive.

PLACE

Just as the time is *now*, US IRBs are increasingly finding that their place includes wherever it is that their investigators go. Everywhere US-funded investigators are engaging in human research, they and their IRBs are increasingly coordinating their approach to the ethical basics with the "local research context."[20] US-based researchers and IRBs are working on this also with their colleagues in other places, including, when necessary, supporting capacity-building efforts to help those in other countries form and operate their own ethical review boards. No matter where an ethics board or IRB is situated, it applies the same basic ethics to its work. Among IRB professionals, these basic ethical principles are sometimes called the Big Three.

The Big Three—respect for persons, beneficence, and justice—offer instant universal orientation to those concerned about the protection of human subjects within the academic research community. Uniformly in the United States and frequently abroad, we refer to the Belmont Report,[9] the primary guidance document, where one finds the elegant trio. They form the basis of the Common Rule, the US federal regulations. They inspire also the growing wad of checklists and review forms on which investigators and IRBs increasingly depend to survive audits, whether actual or anticipated. They are also being translated into the language of every location where US-based researchers are working with staff and colleagues in conducting research.

The obligations of institutions and IRBs that operate under a federal assurance include ensuring that collaborators, wherever they are in the world, are also complying with these ethical standards. Foreign institutions must also obtain a federal assurance from OHRP if they are conducting human research with US federal funds. To fulfill its obligation to monitor research under its jurisdiction anywhere it takes place,[21] every IRB is required to find a way to use existing infrastructure or develop new infrastructure for communication and to learn about the local research context, observing consent processes and assessing capacity.

Achieving compliant research operations across time zones and languages requires the establishment of solid working relationships between US IRBs and their international partners. This requires not only electronic and digital communication technologies, but also travel and other costs not previously envisioned (something wealthier institutions may be better able to manage than others). Early in 2001, the Office for International Activities was created within OHRP to support the application of the Common Rule and

the principles of respect for persons, beneficence, and justice in international research conducted by US institutions.

CONSCIOUSNESS

The time being *now*, the place *anywhere,* the mission to foster a culture of awareness, US IRBs are being greatly challenged. Not only must they review and approve protocols, monitor behavior, enforce compliance, and guide investigators, but they must do so in a way that ensures that legitimate research can get done—and without unnecessary delays. Despite their virtues, IRBs often cannot respond to submissions fast enough. Delays present obstacles to getting research started (or done at all) and can prevent probable benefits from reaching individuals or society. Investigators are therefore increasingly concerned with the role and function of IRBs and the increasing attention they require.

NIH began requiring a one-time training module for grantees in 2000, but it has since been generally agreed that a one-shot training is not enough. IRBs within the United States are grappling to offer ongoing training, using media such as videos and CD-ROMs and offering classes and workshops where attendees can ask questions, get to know their IRB administrators, and build rapport. Regular workshops led by IRB professionals can be reasonably cost-effective and are adaptable to both basic and specialized content. Employing an administrator to create and deliver a 90-minute PowerPoint presentation might cost an institution several hundred dollars the first time, but only $100 for subsequent presentations of the same slides. Investments in training can potentially save an institution millions of dollars in lost grant funding or lost contributions from donors reacting to shaken trust in the institution.

What are IRBs and federal regulators teaching to inspire investigators and to help make them more comfortable with the process? The importance of maintaining the public trust through compliance with regulations and the Big Three that inspire them: respect for persons, beneficence, and justice. It is reassuring that in the current high-pressure, grant-dependent, immensely complex and fact-driven world of scientific research on human beings, basic ethical principles continue to help make sense of the confusion. The Big Three and a small set of related regulations are being understood and applied at the individual level by researchers in designing better studies, by IRB members in doing their work, and by the public from which human subjects are recruited.

The double challenge posed by the dimensions of place and consciousness is the application of basic ethical principles not only in theory but in practice. In this respect, IRBs are increasingly trying to marshal the resources to conduct site visits to observe consent processes, research procedures, and records. Follow-up with investigators would allow IRBs to give constructive criticism, provide appropriate review, take corrective action, or all three. Such field initiatives could make the IRB real to investigators and its role better known to human subjects, while helping to keep the IRB informed about the realities of the research it reviews. Over time, IRBs may well be able to offer valuable services to investigators through this approach and thereby help to overcome the anxiety and resentment caused by some of the enforcement tasks performed by IRBs.

CONCLUSION

At this writing, there is talk on Capitol Hill about passing legislation consolidating federal oversight of human subjects protections, to strengthen the system by simplifying it.[22] (Currently, 17 federal agencies subscribe to the Common Rule, but several—most notably, the FDA—have their own rules. See the list of 17 codifications at 45 CFR §46, after the table of contents.) Comments are being solicited for modifications to FDA regulations that would require investigators to inform an IRB of previous actions on a protocol by other IRBs, in an effort to avoid "IRB shopping."[23]

Comments are also being sought for changes to the medical privacy rules to ensure a more sensible effect on research and IRBs than the current version projects.[24] There will surely be more contortions in the regulatory landscape as the movement to make compliance manageable continues. Granting agencies will need to give increased attention to the financial resources needed to pay for compliance costs; within the United States, one can ask whether NIH will allow for compliance costs in the budgets of grant applications in the near future.

The key ethical principles, the central moral force of the Belmont Report, and the IRB-based system of review and oversight do not appear to be on the table for renegotiation. On the other hand, effective management of the IRB agenda must remain at the forefront of work in this area. We are involved in developing a quality improvement plan at the Harvard School of Public Health to remedy past weaknesses and to provide a thoroughly compliant program for human research protections. Many improvements have been made and many remain to be made, but years will pass before the measure of success can fairly be taken. The rapidity with which action has been taken, however, cannot be entirely explained by the catalyzing effects of an investigation by OHRP. The public health culture is already receptive to the driving ethics of respect for persons, beneficence, and justice. Public health has, after all, always looked to maximize the social and community benefits of research and clinical interventions.

Speaking with the new regulatory voice, Koski has urged moving "beyond the culture of compliance . . . to a culture of conscience and responsibility."[25] Public health institutions are particularly well positioned in time, place, and consciousness to lead the rest of the research community in the new era of protecting human research subjects. US public health institutions and their IRBs have an opportunity to fulfill a key role in the world if we

can all manage to continue practicing respectful collaboration. ■

About the Authors

Sarah B. Putney is with the Human Subjects Committee and Sofia Gruskin is with the Program on International Health and Human Rights, François-Xavier Bagnoud Center for Health and Human Rights Harvard School of Public Health, Boston, Mass.

Requests for reprints should be sent to Sarah B. Putney, JD, Human Subjects Committee, Harvard School of Public Health, 1613 Tremont St, Boston MA 02120 (email: sputney@hsph.harvard.edu).

This article was accepted April 17, 2002.

References

1. 45 CFR §46.103.

2. Division of Assurances and Quality Improvement. Available at: http://ohrp.osophs.dhhs.gov/irbasur.htm. Accessed May 5, 2002.

3. Stolberg SJ. Gene therapy ordered halted at university. *New York Times.* January 22, 2000:A1.

4. Determination Letter from OHRP to Johns Hopkins University School of Medicine and Johns Hopkins Bayview Medical Center, July 19, 2001. Available at: http://www.ohrp.osophs.dhhs.gov/detrm_ltrs/ju01a.pdf (PDF file). Accessed May 5, 2002.

5. Levine RJ. *Ethics and Regulation of Clinical Research.* 2nd ed. New Haven, Conn: Yale University Press; 1986: 70–71, 217–219.

6. Reverby SM, ed. *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Study.* Chapel Hill: University of North Carolina Press; 2000.

7. Jones JH. *Bad Blood.* New York, NY: Simon & Schuster; 1993.

8. Brandt AM. Racism and research: the case of the Tuskegee Syphilis Study. *Hastings Center Rep.* 1978;8(6):21–29.

9. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.* Washington, DC: National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research; April 18, 1979.

10. 45 CFR §46.107(a)-(c).

11. Institutional Review Board Guidebook, chapter 1: institutional administration. Available at: http://ohrp.osophs.dhhs.gov/irb/irb_chapter1.htm. Accessed May 5, 2002.

12. 45 CFR §46.107(d).

13. 45 CFR §46.103(b)(4)(ii) and (d); §46.107; §46.111.

14. 45 CFR §46.107(e).

15. National Human Research Protections Advisory Committee. Draft interim guidance. Financial relationships in clinical research: issues for institutions, clinical investigators, and IRBs to consider when dealing with issues of financial interests and human subject protection. January 2001:3.1. Available at: http://ohrp.osophs.dhhs.gov/nhrpac/mtg12-00/finguid.htm. Accessed May 6, 2002.

16. 45 CFR 46.109(d).

17. Requests for applications, Human Subjects Research Enhancement Program. Available at: http://grants1.nih.gov/grants/guide/rfa-files/RFA-OD-02-003.html. Accessed May 5, 2002.

18. OHRP Division of Assurances and Quality Improvement. Quality Improvement Program. April 15, 2002, Available at: http://ohrp.osophs.dhhs.gov/humansubjects/qip/qipdesc.pdf (PDF file). Accessed May 6, 2002.

19. OHRP. National Human Subject Protections Education Workshop Program. Available at: http://ohrp.osophs.dhhs.gov/wrkshp.htm. Accessed May 6, 2002.

20. IRB knowledge of local research context. August 27, 1998 (updated July 21, 2000). Available at: http://ohrp.osophs.dhhs.gov/humansubjects/guidance/local.htm. Accessed May 6, 2002.

21. 45 CFR §46.109(e).

22. Kennedy human subjects research bill reportedly would fill gaps in existing security net. March 28, 2002. Available at: http://www.washingtonfax.com/ (subscription required).

23. Institutional review boards: requiring sponsors and investigators to inform IRBs of any prior IRB reviews. *Federal Register.* March 6, 2002; 67:10115–10116.

24. Proposed modifications to standards for privacy of individually identifiable health information. *Federal Register.* March 27, 2002;67:14775–14815.

25. Hitts PJ. New voluntary standards are proposed for experiments on people. *New York Times.* September 29, 2000:A14.

Nontherapeutic Research with Children: *Grimes v Kennedy Krieger Institute*

Leonard H. Glantz, JD

Research with young children raises difficult issues of law and ethics. A recent Maryland case, *Grimes v Kennedy Krieger Institute, Inc*, appears to impose restrictive rules on research with children when the subjects are put at risk but cannot derive direct benefit from their involvement in the research project. This case exemplifies the tension that exists between the goal of science to increase knowledge and the protection of the rights and welfare of nonconsenting research subjects. While some language in the opinion may be difficult to understand or apply, for the most part the case reflects the problems other courts and ethicists have had in delineating the role of children in "nontherapeutic" research. (*Am J Public Health.* 2002;92:1070–1073)

RESEARCH WITH CHILDREN starkly raises difficult issues of ethics, social policy and law.[1,2] The case of *Grimes v Kennedy Krieger Institute Inc* (KKI)[3] marked the first time a state's highest court directly addressed the issue of the authority of parents to consent to their children becoming research subjects when the research offers no prospect of direct benefit to the children. (In this article I will refer to this type of research as "nontherapeutic" research, which is also the language the court uses.) The Maryland court also addressed the legal relationships and obligations researchers and institutions that conduct nontherapeutic research have to child subjects.

The decision addresses issues raised in 2 negligence cases in which the plaintiffs were young children involved in research projects that were designed to determine the relative effectiveness of different methods of lead paint abatement. The projects were funded by the Environmental Protection Agency.

Both cases allege that the children were "poisoned" or put at the risk of being lead poisoned as a result of being research subjects. In particular, the plaintiffs alleged that KKI required that certain homes undergo only partial lead paint abatement; required landlords to rent premises to families with young children; encouraged the families of these children to remain in the study houses so that the children's blood could be tested over time; knew that one of the children had a dangerously elevated blood lead level and failed to notify the parents of this in a timely manner; and failed to provide the parents with a complete and clear explanation of the research. Only low-income children were involved as subjects, and their parents were "enticed" by food stamps, money, or other items.

The trial court granted KKI's motion for summary judgment and dismissed the case. That court based its judgment on the ground that the researchers had no legal duty to warn the plaintiffs of the presence of lead dust. The specific issue the plaintiffs appealed is the following:

> Was the trial court incorrect . . . that as a matter of law a research entity conducting an ongoing nontherapeutic study does not have a duty to warn a minor volunteer participant and/or his legal guardian regarding the dangers present when the researcher has knowledge of the potential for harm to the subject and the subject is unaware of the danger?[3]

On its face, this is a relatively narrow question of law, simply involving a duty to warn. However, in answering it the court of appeals covered a great deal of legal territory.

THE COURT OF APPEALS RULING

The court of appeals, reversing the trial court's ruling, held that "the very nature of nontherapeutic scientific research on human subjects can, and normally will, create special relationships out of which duties arise."[3] It then went on to describe the various legal theories that could create and delineate the obligations of researchers.

1. *Contractual obligations.* The court found that "[r]esearcher/subject consent can, and in this case did, create a contract." Whether researchers and subjects enter into a contract as a result of the consent process has been a matter of some question. Here the court held that in the nontherapeutic research circumstance in which subjects are compensated, no matter how small the compensation, a contact is created as a result of the consent process. This means that the consent form can create legally binding obligations on researchers. The court specifically reserved judgment about whether such contractual obligations are created in therapeutic research.

2. *Regulatory obligations.* The court held that the federal regulations that govern federally funded research might create affirmative duties for researchers that are enforceable in state negligence lawsuits. This is an important ruling because the federal regulations are merely conditions on the receipt of federal funding, with the remedy for failure to meet these conditions being suspension of federal funding as a result of governmental action. Enabling individuals to essentially enforce the federal regulations provides a new means for ensuring subjects' rights.

3. *The Nuremberg Code.* The Nuremberg Code is a set of principles that came out of the trials of Nazi doctors after World War II. There has been significant controversy as to whether courts would perceive this code as applicable to US researchers.[4] The Maryland court stated that the Nuremberg Code "might well support actions sounding in negligence in cases such as those at issue here."[3] This is the first time any US court has so explicitly adopted the Nuremberg Code as a source of legally enforceable ethical standards.

These 3 rulings standing alone would make this a notable case. But the part of the case that has engendered the most concern is the section dealing with the legality of enrolling young children in research projects that present a risk of harm but no possibility of direct benefit to the individual research subjects. The court held:

> In our view, otherwise healthy children should not be the subjects of nontherapeutic research that has the potential to be harmful to the child. It is, first and foremost, the responsibility of the researcher and the research entity to see to the harmlessness of such nontherapeutic research. Consent of parents can never relieve the researcher of this duty.[3]

This language would appear to impose a requirement of "harmlessness" on researchers and institutions that wish to perform nontherapeutic research with subjects who are incapable of consenting.

Similarly, the court restricted parental authority to consent to their children's participation in such research. After noting that children are not "the equivalent of rats, hamsters, monkeys and the like," that parents are supposed to act in their children's best interests, and that it is not in the best interests of a healthy child to be "intentionally put in a nontherapeutic situation where his or her health may be impaired, in order to test methods that may ultimately benefit all children," the court stated,

> We hold that in Maryland a parent, appropriate relative, or other appropriate surrogate, cannot consent to the participation of a child or other person under legal disability in nontherapeutic research or studies in which there is *any risk* of injury or damage to the health of the subject [emphasis added].[3]

It is the restrictiveness of this standard that has caused a great deal of concern at KKI and elsewhere, because research that compares the effectiveness of preventive interventions with children is widely conducted by public health investigators. This *no*-risk standard is considerably stricter than current federal regulations, which permit research that presents *no greater than minimal* risk.[5] Even research that presents greater than minimal risk with no prospect of benefit to the

individual child subject is permissible under the federal regulations in certain circumstances.[6]

The no-risk standard adopted by the court led KKI, supported by several academic organizations, to ask the court to reconsider its ruling, essentially arguing that the court's strict standard would prohibit important research. The court denied the motion for reconsideration but attempted to clarify its position by noting,

> [B]y "any risk," we meant any articulable risk beyond the minimal kind of risk that is inherent in any endeavor. The context of the statement was a nontherapeutic study that promises no medical benefit to the child whatsoever, so that any balance between risk and benefit is necessarily negative.[3]

It is difficult to reconcile the first sentence of this statement, which seems to permit some minimal risk, with the second sentence, which seems to mean that *any* risk is too great in the context of nontherapeutic research.

ARTICULATING ACCEPTABLE RISK

This court is not alone in being bedeviled by the problem of being unable to clearly delineate the type of risk to which a nonconsenting subject of nontherapeutic research may ethically or legally be exposed. The problem can be summarized by asking if there is an acceptable risk:no-benefit ratio that justifies the enrollment of nonconsenting subjects in nontherapeutic research.

In 1996 an intermediate appellate court in New York confronted a similar issue regarding nontherapeutic research involving incompetent subjects. In that case the plaintiffs brought an action to challenge the regulations of the New York Department of Mental Health that sanctioned the performance of greater-than-minimal-risk nontherapeutic experiments on incapable adults and children. The court found,

> [A] parent or guardian . . . may not consent to have a child submit to painful and/or potentially life-threatening research procedures that hold no prospect of benefit for the child and that may have the same result as a denial of necessary medical treatment.[7]

It is easy to accept the notion that parents may not submit their children to life-threatening research procedures that offer no prospect of direct benefit, but the prohibition on "painful" procedures would seem to prohibit even the drawing of blood.

The federal research regulations define minimal risk as a situation in which "the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests,"[8] but the reference to daily life is nebulous at best. The question is, "whose 'daily life' is the reference point?"[9] Since the daily lives of inner-city adolescents are likely to be riskier than those of middle-aged suburban adults, a literal application of this standard would appear to permit riskier research with the adolescents than with the adults.

A report by the National Bioethics Advisory Commission recommended, "Minimal risk should be defined as the probability and magnitude of harms that are normally encountered in the daily lives of the general population."[10] But trying to apply this to the KKI case makes it apparent that this definition does not resolve the issue. What is the risk of lead paint exposure in the "daily lives of the general population"? While inner-city children in Baltimore have a significant risk of such exposure, this is certainly not typical of the general population in the United States.

Another report by the commission suggests a procedural resolution of the problem.[11] This solution would permit nontherapeutic research with incompetent subjects that presents greater than minimal risk if the research were approved by a special standing panel that would be appointed by the Secretary of Health and Human Services—approval by a local institutional review board would not be sufficient to authorize such research.

In one of the great intellectual debates concerning the ethics of involving children in nontherapeutic research, Paul Ramsey argued that the use of young children as research subjects when they could derive no benefit from participating is never justified,[12] while Richard McCormick argued that research on children is ethically permissible so long as it involves "no discernible risks, no notable pain, [and] no notable inconvenience."[13] But McCormick's formulation is not particularly helpful. Is there a difference between "no risk" and "no *discernible* risk?" How much pain constitutes "notable" pain? Anyone who has experienced the reaction of an infant to an injection would have to conclude that the pain was notable to the child, although perhaps not notable to a typical adult.

The KKI court of appeals is in good company both in its attempt to protect child subjects and in its failure to fashion a satisfactorily clear standard.

CONCLUSION

This case will be criticized because it will be perceived as prohibiting some research. But this is true of all research regulation. Such a consequentialist criticism lacks ethical weight. If the court went too far in its protection of children, one needs to be able to say how far is correct.

The court of appeals was clearly correct in reversing the trial court's general ruling that institutions and researchers do not have a duty to protect subjects. It was also correct in arguing for a very cautious approach toward nontherapeutic research with children. Indeed, it is in agreement with the earlier New York court in adopting this protective approach. What we do not know, however, is how the court would apply its approach to a particular factual circumstance—and no facts have yet been proven in the KKI case.

It is important to remember that the plaintiffs' allegations

have not been proven, because the trial court dismissed the case before the trial was held. When a defendant asks for summary judgment, the judge is supposed to view the plaintiff's evidence "in a light most favorable to them." The appeals court applied this rule in rendering its decision, which gives the impression that the appeals court believed the plaintiffs' allegations. The appeals court sent the case back to the trial court. The outcome of the trial is far from certain. As the appeals court noted in the last line of its order rejecting the defendants' motion for reconsideration, "the determination of whether the study in question offered some benefit, and therefore could be regarded as therapeutic in nature, or involved more than that minimal risk is open for further factual development on remand."[3]

While some may characterize this case as an example of an overly zealous court's intruding into the prerogatives of the research community, such a characterization would be unfortunate and shortsighted. There is no doubt that parts of the opinion are vague and potentially difficult to apply. It is not uncommon for courts to encounter these types of problems when they enter a new area of jurisprudence with little precedent upon which they can rely. But it is important to keep in mind that the only 2 courts that have addressed the issue of nontherapeutic research with nonconsenting subjects have voiced strikingly similar concerns. The research community would do well to closely examine, understand, and try to address the legitimate legal and ethical issues that so concerned these courts. ■

About the Author

Leonard H. Glantz is Professor of Health Law at the Boston University School of Public Health. Requests for reprints should be sent to Leonard H. Glantz, JD, Boston University School of Public Health, 715 Albany St T-3, Boston, MA 02215 (e-mail: lglantz@bu.edu).

This article was accepted March 19, 2002.

References

1. Glantz LH. Research with children. *Am J Law Med.* 1998;24:213–244.

2. Grodin MA, Glantz LH, eds. *Children as Research Subjects: Science, Ethics and Law.* New York, NY: Oxford University Press; 1994.

3. 366 Md 29; 782 A2d 807; 2001 Md LEXIS 496 (2001). Also available at: http://www.courts.state.md.us/opinions/coa/2001/128a00.pdf (PDF file). Accessed May 2, 2002.

4. Annas GJ. Mengele's birthmark: The Nuremberg Code in United States courts. *J Contemp Health Law Policy.* 1991;7:17–45.

5. 45 CFR §46.404.

6. 45 CFR §46.406.

7. TD v NY State Office of Mental Health, 228 AD 2d 95, 650 NYS 2d 173 (1996) rev on other grounds, 690 NE2d 1259 (NY 1997).

8. 45 CFR §46.102(i).

9. Freedman B, Fuks A, Weijer C. In loco parentis: minimal risks as an ethical threshold for research upon children. *Hastings Center Rep.* 1993;23:13–19.

10. National Bioethics Advisory Commission. *Ethical and Policy Issues in Research Involving Human Participants–Volume 1.* August 2001. Available at: http://www.ntis.gov. Accessed April 30, 2002.

11. National Bioethics Advisory Commission. *Research Involving Persons with Mental Disorders That May Affect Decisionmaking Capacity.* December 1998. Available at: http://www.ntis.gov. Accessed April 30, 2002.

12. Ramsey P. *The Patient as Person.* New Haven, Conn: Yale University Press; 1970:11–12.

13. McCormick RA. Proxy consent in the experimentation situation. *Perspect Biol Med.* 1974;18:2–20. Cited by: Mnookin RH, Weissberg DK. *Child, Family and State.* 3rd ed. New York, NY: Aspen Publishers Inc; 1995:667.

Risk and Responsibility: Ethics, *Grimes v Kennedy Krieger,* and Public Health Research Involving Children

Anna C. Mastroianni, JD, MPH, and Jeffrey P. Kahn, PhD, MPH

The legal case of Grimes v. Kennedy Krieger Institute, Inc, has raised concerns in the public health research community regarding the acceptable level of risk in research involving children, parental authority for informed consent, and exploitation of research subjects for the benefit of public health. We provide an overview of the case and discuss the impact of the court's decision and its possible effect on future research protection policies and practices. (*Am J Public Health.* 2002;92:1073–1076)

THE AUGUST 2001 DECISION in the case of *Grimes v Kennedy Krieger Institute, Inc*,[1] sent shockwaves through the public health research community. The court had challenged the acceptable level of risk in pediatric research studies, concluding that parents in the state of Maryland could not consent to their minor children's participation in research that posed even a minimal risk of harm if it offered no prospect of direct medical benefit to the subjects. Researchers feared that valuable public health research that complied with long-standing federal standards for research on children would be halted altogether or subjected to judicial oversight and intervention.

Two months after the initial decision was handed down, the court indicated that it had not intended to apply a zero-risk standard to "nontherapeutic" pediatric research studies. The court articulated acceptable risk in this

context as including the "minimal kind of risk that is inherent in any endeavor," which appears to be consistent with the federal regulatory standard.[2]

Despite the welcome clarification, *Grimes* deserves analysis of the ethical issues it raises and its relevance to future research. Here we provide a brief overview of the case, discuss some of the ethical issues raised by the court (risk–benefit assessment, informed consent, and exploitation), and offer opinions on the impact of the court's decision on research involving children and its possible effect on future research protection policies and practices.

OVERVIEW OF THE CASE

In *Grimes,* the Maryland Court of Appeals permitted the parents of minor children to bring a negligence action against the Kennedy Krieger Institute (KKI) for lead-related health injuries allegedly contracted by their children during a KKI research study.[1] Researchers at KKI, a Baltimore-based children's health facility and research institute affiliated with Johns Hopkins University, conducted a 2-year study (1993–1995) to measure the effectiveness of differing levels of lead abatement procedures in housing. Identifying the minimal effective level of lead abatement was considered important because of the need to preserve availability of low-rent urban housing that might be abandoned by landlords if they had to pay for the expensive repairs needed to eliminate lead using standard abatement methods. The study was funded by the US Environmental Protection Agency and local Maryland organizations.

In the project, 108 Baltimore rental properties were classified into 5 groups. Three groups of housing received less than full lead abatement (a different level for each group); housing in the 2 control groups either had previously undergone full lead abatement or had been constructed without lead paint. Over a 2-year period, the researchers were to measure and compare lead dust levels collected in the housing with lead levels in blood samples drawn from children living in those homes. Informed consent was to be obtained, and parents were to be notified of their children's blood levels and the results of lead dust collection in their homes. Blood levels would provide evidence regarding the effectiveness of a particular level of abatement. The study was approved by a Johns Hopkins University institutional review board (IRB).

Two families in housing that received less than full lead abatement brought the cases against KKI. One of the families was described as being on public assistance and from a minority group. The parents alleged that KKI designed a study that placed children at unnecessary risk; that KKI had discovered hazardous conditions in their homes and delayed reporting of test results that would have allowed them to prevent their children from being exposed to high levels of lead; and that KKI failed to completely and accurately inform them of all the hazards and risks of the study.

As will be discussed below, in the course of its lengthy opinion, the court willingly stepped into the shoes of investigators, IRBs, and regulators to scrutinize assessments of risk and benefit, oversee informed consent, and raised issues of exploitation. The judges also challenged a parent's authority to act in the child's best interests in consenting to research participation. One of the primary conclusions of the appeals court was that, in Maryland, a parent or other legal surrogate could not consent to a child's participation in "nontherapeutic research or studies in which there is *any* risk of injury or damage to the health of the subject."[3]

This decision would have had the effect of proscribing research on children in Maryland even if the research was in compliance with federal regulations on research using children. However, as noted above, the court's later clarification appears to comport with the federal regulations. It stated that its conclusion referred only to risk in a "nontherapeutic study that promises no medical benefit to the child whatever, so that any balance between risk and benefit is necessarily negative."[2]

It is now up to the trial court to apply the conclusions of the appeals court in light of a full presentation of the facts at trial, and further, to determine how the benefit and risk of the study should be characterized. In the interim, 3 additional families seeking $22 million dollars in damages have brought cases against KKI, raising claims similar to those currently under consideration.[4]

RISKS AND BENEFITS

Research on human subjects must be understood as an inherently risky endeavor, no matter how thoughtfully it is designed and how carefully it is carried out. This understanding is reflected in the long history of federal policies created to protect research subjects from risk, with additional protections for population subgroups deemed particularly vulnerable, including children.[5] These policies are anchored in the ethical principles of respect for persons, beneficence, and justice—articulated in the Belmont Report[6]—which translate to the very issues that the court focuses on in its ruling: informed consent, appropriate balancing of risks and benefits, and avoiding the unjust use of research subjects.

These same ethical principles animate current federal requirements for prospective review by local IRBs to ensure that risks and benefits are appropriately balanced and fairly distributed and for provision of informed consent.[7] Additional protections for minor children include parental (or guardian) consent and limits on risk levels. In particular, when research does not offer the prospect of direct benefit to the subjects themselves (a category of research that the Maryland court referred to as "nontherapeutic"), regulations limit participation by children to research that carries "not greater than minimal risk." In practice, this risk criterion is interpreted as no more risk than the subjects would encounter in their daily

lives, as ultimately judged by IRBs. When the proposed research does offer the prospect of direct benefit to the subjects, this minimal risk criterion does not apply, and IRBs are charged with ensuring that risk and benefit are appropriately balanced.

The trial court's characterization of the risks and benefits to the children who were subjects in KKI's research will be key to the final judgment about both the legal and ethical acceptability of the research. In *Grimes,* the court based conclusions on an assumption that the research had offered the subjects no benefits but a significant risk of lead poisoning, which would have made the risk–benefit balance unacceptable even under federal regulations. Assessment of the risk–benefit ratio on the basis of the facts that will be presented by the parties to the legal action will rest on whether the children would have experienced higher or lower levels of lead in their homes had they *not* participated in the research—a question that may be impossible to answer. Media accounts suggest that children involved in the research ended up living in homes with lower lead exposure than homes of families not enrolled in the study, as the improvements were made in neighborhoods where about 95% of the homes contain lead hazards.[8,9]

INFORMED CONSENT

Even if the children were to benefit, on balance, from participation in the research, their recruitment and inclusion would not be ethically acceptable if their parents did not adequately understand the risks and potential benefits of the research when they consented to their children's participation. The Maryland judges expressed their concern that the study's consent forms did not clearly indicate the risk that children would likely ingest hazardous lead dust particles and that the lead levels in their blood, whether increasing or decreasing, would be used as a way of assessing the effectiveness of the lead abatement measures.

The court was also concerned about the timeliness of disclosure of testing results. Although the parents were notified of their children's blood lead levels very soon after collection, they claimed that reporting of the results of dust collection in their homes was significantly delayed. In at least one of the cases, these results were reported 9 months after collection, and only after the child's blood revealed elevated lead levels.

If the researchers did withhold information from parents that could have helped them reduce the risks to their children, thus allowing identified risks to remain unaddressed, this raises significant ethical concerns. The situation would be especially problematic if the researchers knew that the research posed substantial health risks to the children in the study.

EXPLOITATION

Even if the children realized some benefit from their research participation, and especially if they did not, the court was concerned that such studies have the possibility to exploit economically disadvantaged populations, given their limited access to lead-free housing. Comparing the KKI study to the infamous so-called Tuskegee Syphilis Study, the court said:

> Otherwise healthy children . . . should not be enticed into living in, or remaining in, potentially lead-tainted housing and intentionally subjected to a research program, which contemplates the probability, or even the possibility, of lead poisoning or even the accumulation of lower levels of lead in blood, in order for the extent of contamination of the children's blood to be used by scientific researchers to assess the success of lead paint or lead abatement measures. . . . [P]arents, whether improperly enticed by trinkets, food stamps, money or other items, have no more right to intentionally and unnecessarily place children in potentially hazardous nontherapeutic research surroundings than do researchers. In such cases, parental consent, no matter how informed, is insufficient.[10]

The court's comparison of the KKI research to the Tuskegee study may not be entirely apt, but it is important for what it says about how research and researchers are viewed, particularly within communities that may have been exploited in the past and that feel vulnerable to possible exploitation in the future. Real exploitation is obviously unacceptable, and perceived exploitation works to undermine trust in research and researchers. While it is difficult to regulate or adjudicate away mistrust, creating clear policies that ensure respect for subjects fosters trust in researchers, the research oversight process, and the research enterprise generally.

DISCUSSION

It is important to recognize that the case has not run its course through the judicial system. However, even at this stage of judicial inquiry, the case clearly highlights the need for true partnership in the research enterprise, particularly when proposed research involves vulnerable communities. Collaboration and open communication should begin with the development and design of research and continue throughout the processes of subject recruitment, data collection, and eventual disclosure of research results.

Further, community partnerships and collaborations provide an important test of the ethical propriety of the research. If researchers cannot convince the community in which they propose to conduct the research that it is acceptable, then they must reevaluate and revise their proposal until the community's concerns are adequately addressed.

Public health researchers already recognize the value of community partnerships in preventing misunderstandings. The extent of community participation in this particular research project is not indicated in the court's opinion, but given the court's views, a participatory approach can be seen to have a secondary benefit: as a tool to prevent judicial second-guessing of researchers' motives. In the absence of community participation, researchers can be per-

ceived, as they were in *Grimes,* as placing the investigators' and institution's interests above the interests of research subjects and the communities to which they belong.

This is not to suggest that community partnerships are a substitute for improved oversight of research—both are important priorities. Certainly no child or parent should be taken advantage of in the context of research, and significant policy protections have been designed to prevent exploitation. However, no matter how well written, policies cannot prevent poor implementation in practice. We can always improve research practices, and we need only look to historical deficiencies to understand the need for doing so.

In the research at issue in *Grimes,* we do not know yet whether shortcomings were inherent in the research protocol (which ought to have been caught by peer review or the IRB) or occurred in the implementation of the research (meaning the investigators deviated from the protocol or there was a failure to adequately communicate with the research subjects). Whatever the eventual findings, protection of subjects would benefit from improved education of researchers, ongoing oversight of research in addition to prospective review, and mechanisms for assuring that research protocols are faithfully followed, all of which could be part of the IRB's responsibility.

Throughout its opinion, however, the court expressed its concern that IRBs are in positions of conflict of interest because they are committees of the research institution, with most of their members drawn from the institution they are charged to oversee, and therefore that the peer review system cannot adequately police itself. Our view is that the problems with oversight are less related to conflicts of interest than to the fact that IRBs have not received adequate priority and attention, financially and otherwise; therefore they have increasingly unmanageable workloads and inadequate resources to carry out their efforts. Clearly, IRBs can do a better job—and the court's strong words should provide an impetus for IRBs to redouble their efforts to carry out their task in ways that are above reproach—but they are only part of the overall research subject protection process.

With pointed references to many of the worst research scandals in history, such as the Tuskegee Syphilis Study, the ruling in *Grimes* seems to conclude that public trust in research is not warranted. While researchers and institutions can and must do a better job of respecting research protection policies, instances of failure do not indicate that investigators or the system of oversight are corrupt, or that intervention by the courts is the most effective remedy. Such interventions will lead only to greater obstacles to carrying out research, not to greater protections for subjects.

The court's decision in *Grimes* is not so much a call to rethink the ethics of research involving children as it is a warning to make sure that existing protections are adequately applied and overseen.

Wherever the source of research protections and oversight, creating and fostering a culture of ethical research among researchers and research institutions will lessen the need for emphasizing compliance and will go a long way toward warranting the trust of the public. Public trust may be the key, since without it there can be no research—no subjects for research projects, and no funding to support them.

However the case is ultimately decided, the court's involvement in *Grimes* should serve as a wake-up call for all those involved in research and its oversight—investigators, IRBs, peer review panels, and research institutions themselves. The age of judicial oversight is upon us, and responsible research protocols and scrupulous oversight are the way not only to conduct research in the most ethical manner, but to avoid judicial intervention as well. ■

About the Authors

Anna C. Mastroianni is with the University of Washington School of Law and the Institute for Public Health Genetics, Seattle. Jeffrey P. Kahn is with the University of Minnesota Center for Bioethics, Minneapolis.

Requests for reprints should be sent to Jeffrey P. Kahn, PhD, MPH, University of Minnesota Center for Bioethics, N504 Boynton, 410 Church St SE, Minneapolis, MN 55455 (e-mail: kahnx009@tc.umn.edu).

This article was accepted March 19, 2002.

References

1. *Grimes v Kennedy Krieger Institute, Inc.* 366 Md 29, 782 A2d 807 (2001), reconsideration denied (October 11, 2001). Also available at: http://www.courts.state.md.us/opinions/coa/2001/128a00.pdf (PDF file). Accessed May 2, 2002.

2. 782 A2d at 862.

3. 782 A.2d at 858.

4. 3 new complaints seeking total of $22 million filed against KKI over study. *Mealey's Litigation Reports: Lead.* September 19, 2001. Available at: http://www.mealeys.com/ (fee for access). Accessed May 7, 2002.

5. 45 CFR §46, subparts A–D.

6. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.* Washington, DC: National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research; April 18, 1979.

7. 45 CFR §46, subpart A (the Common Rule).

8. Pelton T. Two universities, Krieger Institute seek reconsideration of lead paint opinion. *Baltimore Sun.* September 18, 2001:3B (2001 WL 6169982).

9. Pelton T. Groups target study limits. *Baltimore Sun.* September 30, 2001:B1 (2001 WL 6171119).

10. 782 A2d at 814.

The Impact of Genomics on Global Health

Tikki Pang, PhD, FRCPath

Ensuring that advances in genomics are applied to the health improvement of people living in developing countries is an important contemporary challenge. In the near term, such advances are likely to alleviate infectious diseases, with longer-term benefits envisaged for chronic disorders.

To ensure that benefits are shared by developing countries, attention must be paid to complex ethical, legal, social, and economic issues, as well as to public education and engagement. Creative and equitable international mechanisms and goodwill are needed to turn high hopes into reality and allow the use of genomics to reduce health inequities between rich and poor nations. (*Am J Public Health.* 2002;92:1077–1079)

THE ONGOING GENOMICS revolution, highlighted by the sequencing of the human genome, promises to change how diseases are diagnosed, prevented, and treated. It has tremendous potential to improve health globally. Despite the flush of excitement about its potential, drugs and interventions derived from genomics are likely to be expensive, and of particular interest is how these advances will affect the health of people living in the developing countries. The reality is that many of the advances in genomics were made, and in part are owned, by the developed world, and this has given rise to the concern that a "genomics divide"[1] will be created that will further widen the equity gap in health between rich and poor nations.

Instead, genomics and related technologies should be used to narrow the existing unethical inequities in global health. A report recently released by the World Health Organization focuses on this inequity. It points out, for example, that approximately 80% of investments in genomics in 2000 were made in the United States, and 80% of the DNA patents in genomics in the period 1980 through 1993 were held by US companies.[2] Of the 1233 new drugs marketed between 1975 and 1999, only 13 were approved specifically for tropical diseases.[2]

POTENTIAL FOR HEALTH IMPROVEMENT

In recent years the genomes of nearly 50 microbial pathogens have been sequenced, and ongoing efforts to sequence the genomes of mosquito vectors (e.g., *Anopheles gambiae,* the malaria vector, and *Aedes aegypti*, the main vector for dengue fever) promise benefits in the shorter term for the control of communicable diseases.[3] Fosmidomycin, originally developed for treatment of recurrent urinary infections, showed effective anti-malarial activity when genome sequence information from *Plasmodium falciparum* revealed a common biochemical target, present in the parasite and not in the human host[4]; the drug has gone into clinical trials in less than 2 years. Clinical trials have also begun in Africa of a pre-erythrocytic DNA-based vaccine that gave significant protection against natural *P falciparum* infection.[5]

Although the benefits of alleviation of infectious diseases are obvious, it is now believed that the information generated by genomics will, in the long term, also play a major role in the prevention, diagnosis, and management of many diseases which hitherto have been difficult or impossible to control, including cardiovascular disease, cancer, diabetes, the major psychoses, dementia, rheumatic disease, and asthma.[6] From a public health perspective, the genomics revolution may present new opportunities for the prevention of these diseases, but before these opportunities can be realized we will need to know more about what combination of genetic and environmental factors predispose people to such diseases.[7]

New approaches to population-based epidemiological studies, such as "genomic epidemiology" to chart the molecular, metabolic, and disease profiles of thousands of subjects, may be the path to the future. A new consortium has been formed to pursue this approach, which aims to scale the relevant technologies to sample sizes appropriate for epidemiological studies.[8] The initial focus will be on diabetes and cardiovascular disease, but the goal is to develop generic tools and protocols that can ultimately be applied to other diseases.

The recent announcement of the genome sequences of 2 varieties of rice, *indica* and *japonica*,[9,10] marks another milestone in the genomics revolution with tremendous potential implications for health. Three billion people, mainly in the developing world, depend on rice as their staple diet. The sequencing of the rice genome may pave the way for better strains of rice with enhanced yields, nutritional value, and disease resistance.

IMPLICATIONS AND CONCERNS

Aside from the complex scientific and technical problems of bringing genomics to the clinic, ensuring that its benefits will be reaped by developing countries will require attention to many equally challenging issues. Genomics brings with it complex new ethical, legal, social, and economic implications, as well as concerns about risks and hazards.[11] Issues of confidentiality, stigmatization, and misuse of genetic information are high on the list of concerns, particularly the potential for creating a genetic underclass that may be denied

medical insurance as a result of genetic testing and screening. Genomics has also been associated with the prospect of "designer babies," and there is a concomitant concern about creating a genetically engineered overclass and a disease-prone underclass; the higher likelihood of the former being associated with richer people in the developed world is obvious.

Issues of intellectual property rights associated with DNA sequences[12] and the potential exploitation of developing-country populations by creating genetic databases, often at the behest of companies based in the developed world,[13] are other areas of concern. While industry believes that without strong and effective global intellectual property rules, the gap between developed and developing countries will only grow in the future, there are plenty of concerns about the patentability of DNA sequences and the applications derived from them, and what implications this will have for the developing countries.

Most important, the relatively rich product pipeline of genomics-based drugs will mean a tremendous increase in the demand for clinical trial sites, many of which will be in the developing countries; this area represents an ethical minefield relating to issues such as informed consent, standard of care, and continuing availability of the drug being tested, the price of which is often beyond the reach of poor people.[14] Finally, in the aftermath of the tragedy that took place in the United States on September 11, 2001, the utilization of advances in genomics for acts of bioterrorism and biological warfare similarly occupy the minds of many.

INTEGRATION WITH RESEARCH AND PRACTICE

Despite the tremendous potential and promise of genomics, it is very difficult to predict when its benefits for health will be realized; there are so many critical things we do not yet know about how gene products interact. Many people were surprised to learn that we have only twice as many genes as a fly or a worm.[15] Hence it is vitally important for the developing countries to maintain focus on the basics of what can be done now, particularly in the fields of public health and the development of more functional health care systems.

The main message of the World Health Organization report[2] is that medical practice will not change overnight as a result of new technologies spawned by genomics, but the long-term possibilities are such that both developing and developed countries must prepare themselves for this new technology and carefully explore its possibilities, always looking at its cost-effectiveness compared with more standard approaches to medical care. It is also vital that genomics research not be pursued to the detriment of the well-established methods of clinical practice and clinical and epidemiological research. Indeed, for its full exploitation it will need to be integrated into clinical research involving patients and into epidemiological studies in the community. It is crucially important to maintain a balance in medical practice and research between genomics and these more conventional and well-tried approaches.

In addition, it is crucial to increase the quality of education in genetics and genomics at all levels of society. If this is not achieved it will be impossible to develop an informed debate about the various issues involved, and there is a danger that those who administer health services will be unable to distinguish between hyperbole and reality in a new, uncertain, and rapidly expanding research field.

STRATEGIES FOR EQUITABLE SHARING

What strategies and actions are needed in the future to ensure that the benefits of the genomics revolution are shared by the developing countries? Strong international leadership by the scientific community, international organizations, governments, and industry is required through promotion of innovative partnerships and cooperation strategies. A key issue in the postgenomics era will be who will pay to test, develop, and deliver important vaccines, drugs, and diagnostic procedures for diseases of the developing world, and who will ensure equitable access to those who need it most.

The "Millennium Challenge Account" to improve health in the developing world, discussed at the recent Monterrey summit on financing for development,[16] could be partly used for this purpose. Given the ethical concerns associated with many of the key issues and the significant commercial interest, a proposal has also been made for a Commission on Global Genomics Governance to make recommendations for genome-related issues and activities.[17] At a higher political level, the potential of genomics to generate economic and health benefits for developing countries should be highlighted to the world's leaders. Attention to these problems at the June 2002 meeting of the G8 (the world's wealthiest nations—Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States—and Russia), which is to focus on Africa, would be a visionary move on the part of these countries.

Such a call for action acknowledges that while most of the incentives to develop new drugs and vaccines are primarily of interest to markets in the industrialized world, there are enormous opportunities to apply knowledge of the genome to diseases of the poorest people as well, and that we all have a responsibility to help make these opportunities into realities.

In particular, the medical profession in the developed countries has a vital role to play. Many of the important infectious killers are being encountered with increasing frequency in richer countries and, as the provision of basic health care improves, many poorer countries are making the epidemiological transition toward a pattern of disease similar to those of the devel-

oped countries. Globally, heart disease is now the most common cause of death. The globalization of disease is a message that must be clearly understood by medical schools, research funding bodies, industry, and governments of rich countries.

The development of effective and equitable research partnerships between developed and developing countries will not only help to combat the global inequity of health care, but will also be of enormous mutual benefit to both parties. As Donald Kennedy, editor-in-chief of *Science*, aptly and succinctly put it, "What can First World science do, not for the West, but for the Rest."[18] ■

About the Author

Requests for reprints should be sent to Tikki Pang, Research Policy and Cooperation (RPC/EIP), World Health Organization, Ave Appia, CH-1211 Geneva 27, Switzerland (e-mail: pangt@who.int).

This article was accepted April 11, 2002.

References

1. Singer P, Daar A. Harnessing genomics and biotechnology to improve global health equity. *Science.* 2001;294: 87-89.

2. *Genomics and World Health.* Report of the Advisory Committee on Health Research. Geneva, Switzerland: World Health Organization; 2002.

3. Hoffman SL, Subramanian GM, Collins FH, Venter JC. *Plasmodium,* human and *Anopheles* genomics and malaria. *Nature.* 2002;415:702–709.

4. Jomaa H, Wiesner J, Sanderbrand S, et al. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science.* 1999;285: 1573–1576.

5. Bojang KA, Milligan PJ, Pinder M, et al. Efficacy of RTS,S/ASO2 malaria vaccine against *Plasmodium falciparum* infection in semi-immune adult men in The Gambia: a randomised trial. *Lancet.* 2001;358:1927–1934.

6. Wright AF, van Heyningen V. Shortcut to disease genes. *Nature.* 2001; 414:705–706.

7. Steinberg K, Gwinn M, Khoury MJ. The role of genomics in public health and disease prevention. *JAMA.* 2001; 286:1635.

8. Butler D. Epidemiology set to get fast-track treatment. *Nature.* 2001; 414:139.

9. Yu J, Hu S, Wang J, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science.* 2002;296: 79–92 .

10. Goff SA, Ricke D, Lan TH, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science.* 2002;296:92–100.

11. Fukuyama F. *Our Posthuman Future: Consequences of the Biotechnology Revolution.* New York, NY: Farrar Strauss & Giroux; 2002.

12. Cook-Deegan RM, McCormack SJ. Intellectual property. Patents, secrecy and DNA. *Science.* 2001;293:217.

13. Burton B. Proposed genetic database on Tongans opposed. *BMJ.* 2002; 324:443.

14. Killen J, Grady C, Folkers GK, Fauci AS. Ethics of clinical research in the developing world. *Nature Rev Immunol.* 2002;2:210–215.

15. Lewis F. Science's latest discovery: how much we don't know. *International Herald Tribune.* April 3, 2002:6.

16. A feast of giving. *Economist.* March 23, 2002:73.

17. Loff B. Africans discuss ethics of biomedical research. *Lancet.* 2002; 359:956.

18. Kennedy D. Science and development. *Science.* 2001;294:2053.

Ethical Oversight of Public Health Research: Can Rules and IRBs Make a Difference in Developing Countries?

Leslie London, MD, MB ChB, MMed (Comm Health)

Controversies in the conduct of international research continue to pose challenges for the system of ethical review, particularly for developing countries. Although the concept of vulnerability is key to addressing these challenges, ethical review has typically ignored the agency of vulnerable participants and groups in determining what kind of review process is needed. Concurrent with developments shaping the new public health that seek to operationalize empowerment of communities by placing them as initiators and organizers of their own health, ethical review of public health research must find ways to recognize the agency of vulnerable individuals, groups, and communities in the review process if it is to address effectively the ethical dilemmas currently evident in collaborative international research. (*Am J Public Health.* 2002;92:1079–1084)

ALTHOUGH THE CONTROVERSIES over the ethical conduct of international research have in recent years generated increasingly copious[1–9] and at times acrimonious[10–12] debate, there is little doubt that many key issues remain unresolved,[13] particularly for developing countries.[14] Even in developed countries, the track record of ethical review remains open to criticism, particularly around the disputed definition of what constitutes "minimal risk."[15]

The current institutional framework for collaborative international health research stems largely from US government–driven directives to establish ethical oversight over research involving human subjects, most particularly in the form of federal regulations, and systematized in the form of institutional review boards (IRBs).[15–17] While the role of an IRB is viewed as having broad responsibilities for considering all ethical dimensions of a research protocol, it is recognized that IRBs in practice tend to focus on informed-consent documents.[14,16,18] Indeed, rules-driven ethical oversight[19] has created a disjuncture between ethical codes, which are essentially standards for professional behavior with limited ac-

countability, and legal standards, which have the potential for high levels of sanction and enforceability.[13] As a result, the system has driven a concern about procedural correctness[19] rather than a substantive approach to ethical reasoning.[13]

It is particularly in the setting of international collaborative research, involving researchers from developed and developing countries, that the fault lines have been most exposed.[1,4,13,18,20–23] For example, cross-country research protocols (i.e., research conducted in developing countries by developed country researchers) have set regulation-driven US standards against local interpretations of ethical codes, with the former routinely superceding the latter because of regulatory requirements that substituted provisions must be "at least equivalent" to those in US policies.[16] The responsibility of the IRB—or of its equivalent in a developing country—to span the tension between "legalist" and "ethical" reasoning approaches to safeguarding the rights and welfare of participants is complicated where there are profound cross-cultural differences, inadequate health care provision, high levels of social inequality, and social systems that fail to respect human rights.[16]

It is perhaps not surprising that, in considering ethical oversight, informed consent is most typically identified as the area on which "local" expertise is best placed to comment,[16] rather than on issues of the ethics of study design or justice. For example, locally specific cultural values are said to be more likely to influence preferences in how a study should be conducted (recruitment, disclosure, consent) than to influence the "overall acceptability" of a study.[14] Centers for Disease Control and Prevention (CDC) IRB interviewees expressed the view that national ethics committees were better placed "to address political and cultural acceptability" while CDC IRBs were better placed to address research design and biological risk.[24]

Accordingly, informed consent has become a particularly important focus of attention in ethical considerations in international research. One the one hand, this is welcome. Careful and culturally sensitive attention to matters of consent and confidentiality is essential for protecting the rights and welfare of study subjects[16,18,23,25,26] and finding fair resolution to the problems associated with interpreting what constitutes meaningful informed consent across different cultures.[14,16] However, the requirements of federal regulations and donor agency "rules"[27] have also resulted in the application of extremely detailed and elaborate procedures for consent that, from the point of view of researchers and subjects in developing countries, may simply be inappropriate in their assumptions and expectations of what research participants would see as important protections of their rights. It is not surprising that these stringent rules of consent may be seen as primarily protection for the researcher or institution rather than the study subject.[24,25,27] Even where consent is carefully addressed, it is usually in the design and planning, with little or no attention to monitoring actual implementation or how effective the consent is.[18,26]

Furthermore, as many commentators have pointed out,[13,15,18,26,28] reliance on procedural aspects of consent is wholly insufficient to guarantee ethical standards in research, and such narrow interpretations of ethical obligations have often had disastrous consequences for subjects in developing countries. Concentration on procedural aspects of consent to the exclusion of issues of distributive justice[13,18,21] makes it easier to justify or overlook exploitative research.

VULNERABILITY AND AGENCY

The concept of vulnerability of study populations, and the corollary obligations of the researchers, is central to the well-grounded fear of the exploitation of study subjects—particularly in, but not limited to,[26] developing countries. Traditionally, ethical codes have viewed the inability or incapacity to make independent decisions as the cornerstone of vulnerability[29,30] and characterized vulnerable populations by their need for special protection.[29,31]

Such formulations, while intended to redress imbalances in power underlying conditions of vulnerability, do not necessarily recognize the mutuality of the researcher–subject relationship,[26] inherent in newer conceptions of public participation[32,33] and community agency in determining need[34] in public health. The process of "speaking for others," so central to IRB processes in relation to vulnerable groups,[17] is coming under increasing criticism in development-related analyses of public heath where the "need for people to negotiate their own inclusion"[33] is critical to the success of the new public health.

For example, Eckenwiler[17] defines vulnerability in terms of threats to self-development, self-determination, and equality that exist independent of research and suggests that variables other than those "traditionally" cited in research codes (gender, disability, children, poverty, etc.) should be considered within the ambit of vulnerability, based on an analysis of the participant's own experiences and their particularity. Similarly, Zion et al.[26] define vulnerability in terms of the lack of basic rights and freedoms required for participants' free choices. Brody[25] has argued that failure to make informed consent meaningful has occurred because of a preoccupation with autonomy as a right rather than as a value, resulting in an ethical principle that generates no obligation on the part of the researcher to help the participant make autonomous decisions.

Institutional review processes, therefore, need to develop new ways to recognize and strengthen the agency of individuals, groups, and communities, whom institutional review has thus far only viewed as candidates for protection.[35] As in broader public

health policy, rather than "doing things *for* or *to* the poor, [ethical review should] start strengthening the capacities of the poor *to do* things themselves."[36] By asking difficult but appropriate questions of researchers relating to the role their research will play in facilitating empowerment of vulnerable groups, ethical review should aim to strengthen the connection between autonomy and freedom.[26]

For example, reproductive health studies that have the potential risk of gender violence could, if appropriately planned and managed, afford significant opportunities for intervening to reduce vulnerability and empower marginalized women, both within and outside of the study. Researchers should recognize a positive obligation to actualize participants' autonomy[25] and therefore improve the situation of vulnerable populations. Underresearching of the problems of vulnerable communities is itself an ethical issue that should not be aggravated by the very process of ethical review. The recognition of, and opportunity to support, participants' agency potentially turns institutional review into a mechanism for redress of inequality, rather than framing ethical discourse in relation to vulnerable populations as a negative conditionality (i.e., in terms of conditions under which research might be "ethically acceptable"[26]). Clearly, designing and implementing such studies demand additional resources and represent the added costs of community participation and empowerment. However, in evaluating the ethics of a study, public health research should not accept resource constraints as a given, and certainly should not allow such acceptance to set the terms in which ethical questions in public health research are resolved.

Moreover, one of the implications of the biomedical community's embrace of the randomized controlled trial as the core criterion for any public health evidence base is that other kinds of research, better suited to empowering vulnerable communities and groups, may be systematically neglected. For example, participatory action research[37,38] and social epidemiology[39] are research methods more attuned to assessing community harms and benefits and empowering study subjects. In examining challenges facing public health in developing countries, the strongest evidence for people-driven development in public health is often to be found in case studies,[33] replete with internally subjective assessments of success. Yet such evidence rarely meets traditional standards for scientific objectivity. If ethical review processes are to take seriously the participation of those who are labeled vulnerable, it has to address the passivity imposed on subjects by the objectification of participants inherent in many research designs.

PARTIALITY AND POWER

Linked to the positivist framework of most biomedical research is the view that ethical review requires careful assessment by an impartial external agency. However, commentators have observed that, in reality, the very process of ethical review is often replete with partiality[17] and—in Africa, for example—the colonial legacy leaves significant potential for personal bias.[23] Moreover, decisions of IRBs take place in the context of significant differences in interests among stakeholders, which include governments, funders, academic institutions, researchers, and communities, both within countries and between countries.[26] These interests, depending on the particular context, are expressed through uneven power relations between stakeholders,[40] between stakeholders and IRBs,[8,20,23,40,41] between IRBs and researchers in developed and developing countries,[16,27,42] and even within IRBs.[14,16]

Money, power (political or economic), prestige, custom, indifference, or even simple lack of awareness may all influence, sometimes decisively, the decisionmaking process of ethical review, frequently complicated by conflicts of interest (situations of dual loyalty) in which researchers[17,18,40,41] or IRB members[17,22,23,40,43] might find themselves. It is not uncommon for developing-country IRBs to be pressured by funders or researchers to approve the local arm of a multicenter study that has already met with approval in the funder's home country. Financial incentives to researchers and whole institutions may be sufficiently powerful to shape entire research agendas in the developing world[18,20–22,42] and the developed world.[18,40,41] Thus, while an idealized view of IRBs holds that their decisionmaking is always independent of influence, the reality may often be very different.

Some critics have argued that impartiality in ethical review is simply not possible and that claims for impartiality are in fact part of the problem.[17] What enables differences in power to persist, and to insert themselves into institutional review, is largely the lack of explicit recognition of these diverse and often powerful sectoral interests by the review process. Eckenwiler[17] has referred to the "flattening" process that occurs within IRBs that results in the hiding of inherent subjectivities, prejudice, and misunderstandings in the review process. Therefore, different frameworks for adjudicating conflicting interests, which are perhaps more effective at making the implicit explicit, may be useful if integrated in ethical review. Specifically, human rights approaches, by setting clear terms of reference for adjudicating conflicting rights in relation to widely accepted human rights standards, rely on exactly that process of making diverse and potentially divergent interests explicit, so that value judgments can be applied within a framework that has credibility with all stakeholders.[44,45] So, for example, rights analyses may help to clarify diverging interests, allowing greater weighting to be placed on the rights of participants from vulnerable groups, and thereby making the work of IRBs easier.

Such approaches may also be useful in helping to answer an-

other difficult controversy—to whom are IRBs accountable when funders, governments, participants, and researchers all look to IRBs to ensure that in the process of ethical review, their particular interests are met? IRBs can only balance these diverse pressures by making such pressures explicit, and negotiating a priori how consideration should be given to different kinds of stakeholders in the context of prioritizing the interests of study participants, particularly those from vulnerable groups. However, in the absence of any organized, credible, representative structure, often the interests of the participants are represented in the abstract, by consultants, or by token participation from individuals said to be representing the patients/participants.[8,16,35,43] Here again, the need to recognize the agency of vulnerable groups in representing themselves is paramount.

The problem of "moral relativism"[14,16,18,46] remains a serious obstacle in resolving Western vs non-Western differences in interpreting ethical standards. For example, how should consent be operationalized in societies that do not share Western constructions of autonomy?[14,16,18] How do different values influence the judgments made regarding the justice of the distribution of burdens and benefits in a particular study?[13,18,21,46] For research located in a developing country, US IRBs may be at a substantive disadvantage, having little familiarity with some of the key contextual issues in a developing country needed for assessing the ethical standing of a study.[13,16] Indeed, IRB members may have little if any knowledge about people who enroll in studies, whether in developing or developed countries.[17]

Arguments that it is "ethical imperialism" for outsiders to dictate ethical standards to African researchers and IRBs[12,13,14,47] are correct in drawing attention to uneven power relations between Northern and Southern institutions and researchers. However, such arguments ignore the heterogeneity that exists within countries, within committees, between researchers and communities, and among researchers and among communities. Moreover, the use of these arguements as a sole criterion for ethical justification of studies represents a false "procedural defense"[13] that opens the door further to the internalization of power differentials within the work of IRBs, rendering them more vulnerable to exploitative practices. Adherence to study procedures is a necessary but insufficient condition for the ethical conduct of a study,[13] as much in developed as in developing countries.

Finding an institutional system for ethical oversight that is able to account for all morally relevant factors particular to a local situation[46] without imposing "Western" rules, and in a way that affirms the agency of vulnerable participants, is hugely challenging. One suggestion to address cross-cultural differences in interpreting ethical principles has taken the form of a negotiated ethical standard that recognizes the "important truths" reflected in both approaches based on moral relativism and those based on moral fundamentalism.[16] However, even in its careful algorithm seeking to adapt "Western standards" to best protect the interests of local study populations, the core requirement is still to "satisfy Western standards of respect for persons, beneficence, and justice," which implies that "the sponsoring IRB should have the last word on whether the protocol is approved."[16]

MAKING THE DIFFERENCE

What, then, can be done? The work of IRBs in both developed and developing countries could be greatly facilitated by the establishment of effective and timely independent mechanisms for monitoring the conduct of research[8,18,28,48] and the availability of whistle-blower mechanisms to protect individuals who identify system failures.[18,48] Increasing the transparency of IRB decisions by opening IRB meetings to members of the public[8,14,28] and allowing public access to records of IRB decisions could be effected without compromising proprietary information or professional integrity. Most importantly, expanding IRB membership to include participants specifically drawn from vulnerable communities[14,28] would contribute significantly to critical and challenging perspectives being heard, which could also ensure the incorporation of "particularity into the review process."[17]

Of course, community participation is fraught with complexities; it may end up that only the most powerful—or minority—voices in heterogeneous communities are heard,[16,37] particularly where communities lack the infrastructure or organization to manage outsiders' requests for participation.[16] Recruitment of "lay" participants tends to privilege participants of higher educational and social standing,[16,17] and there is evidence that nonscientist members tend to be less active and to lack influence in ethical review processes.[14,17] Therefore, expanding IRB membership to include significant participation from groups that include vulnerable populations will be effective only if complemented by active steps to empower nonmedical participants in the review process.[35] Complementary to general calls for ethical training of developing country research ethics committees[6,18,43,46] should be measures that enhance nonmedical participants' understanding of how to organize themselves to offer ethical input that "empower[s] citizens to understand and utilize their potential for participating."[17] Moreover, such training must be situated within a context of policy and legislative changes to strengthen the role of civil society in ethical oversight.

The concept of vulnerability has been extended to apply to governments in developing countries, in recognition of the growing inequities generated by the process of globalization and the implications for health-related research.[49] Is it ethically tolerable that only 10% of global

spending on health research is directed toward diseases that contribute 90% of the global burden of disease?[50] And is it the task of the IRB system to address what Benatar[18] has called "macro issues"—the best interests of whole populations and science that concerns itself with the amelioration of "the miserable conditions in which the majority of the world's population live"? A narrow view might regard the IRBs' responsibilities as solely to protect human subjects from the risks posed by their participation in research.[14] However, such an exclusive focus is to lose sight of the broader challenges facing public health research. Vulnerable groups, vulnerable communities, and vulnerable countries will remain passively "in need of protection" until they gain the type of agency "that locates organized and active communities at the center as initiators and managers of their own health."[33] The process of ethical review in public health research must facilitate such agency if it is to address effectively the ethical dilemmas currently emerging in collaborative international research. ■

About the Author

Leslie London is with the Department of Public Health and Primary Health Care, University of Cape Town, Cape Town, South Africa.

Requests for reprints should be sent to Leslie London, MD, MB, ChB MMed (Comm Health), Department of Public Health and Primary Health Care, University of Cape Town, Private Bag Rondebosch, 7700 South Africa (e-mail: ll@cormack.uct.ac.za).

This article was accepted April 18, 2002.

Acknowledgments

The Fulbright Foundation and the Takemi Program at the Harvard School of Public Health provided support for this work.

The comments of Louis Reynolds, Marietjie de Villiers, and Peter Lurie are gratefully acknowledged, as is the encouragement and support of Sofia Gruskin in the completion of this work.

References

1. Lurie P, Wolfe SM. Unethical trials of interventions to reduce perinatal transmission of human immunodeficiency virus in developing countries. *N Engl J Med.* 1994;331:1173–1180.

2. Levine RJ. The "best proven therapeutic method" standard in clinical trials in technologically developing countries. *IRB.* 1998:20:5–9.

3. Varmus H, Satcher D. Ethical complexities of conducting research in developing countries. *N Engl J Med.* 1997; 337:1003–1005.

4. Angell M. The ethics of clinical research in the Third World. *N Engl J Med.* 1997;337:847–849.

5. The ethics industry [editorial]. *Lancet.* 1991;350:897.

6. *The Ethics of Clinical Research in Developing Countries.* London, England: Nuffield Council on Bioethics; 1999. Available at: http://www.nuffieldbioethics.org/publications/pp_0000000013.asp. Accessed April 26, 2002.

7. Annas GJ, Grodin M. Human rights and the maternal–fetal HIV transmission prevention trials in Africa. In: Mann JM, Gruskin S, Grodin MA, Annas GJ, eds. *Health and Human Rights: A Reader.* 1st ed. New York, NY: Routledge; 1999:373–379.

8. Schuklenk U. Protecting the vulnerable: testing times for clinical research ethics. *Soc Sci Med.* 2000;51: 969–977.

9. London AJ. The ambiguity and the exigency: clarifying "standard of care" arguments in international research. *J Med Philos.* 2000;25:379–397.

10. Weijer C, Anderson JA. The ethics wars: disputes over international research. *Hastings Cent Rep.* 2001;31: 18–20.

11. Lansang MA, Crawley F. The ethics of international biomedical research. *BMJ.* 2000;321:777–778.

12. Brunet-Jailley J. The ethics of clinical research in developing countries. *IRB.* 1999;21:8–11.

13. Macklin R. Unresolved issues in international research. *Kennedy Inst Ethics J.* 2001;11:17–36.

14. Macpherson CC. Research ethics committees: a regional approach. *Theor Med Bioeth.* 1999;20:161–179.

15. Woodward B. Challenges to human subject protections in US medical research. *JAMA.* 1999;282: 1947–1952.

16. White MT. Guidelines for IRB review of international collaborative medical research: a proposal. *J Law Med Ethics.* 1999;27:87–94.

17. Eckenwiler L. Moral reasoning and the review of research involving human subjects. *Kennedy Inst Ethics J.* 2001;11: 37–69.

18. Benatar SR. Reflections and recommendations on research ethics in developing countries. *Soc Sci Med.* 2002; 54:1131–1141.

19. Ellis GB. Keeping research subjects out of harm's way. *JAMA.* 1999;282: 1963–1965.

20. Edejer TT. North-South research partnerships: the ethics of carrying our research in developing countries. *BMJ.* 1999:319:438–441.

21. Benatar SR. Distributive justice and clinical trials in the Third World. *Theor Med Bioeth.* 2001;22:169–176.

22. Wolffers I, Adjei S, van der Drift R. Health research in the tropics. *Lancet.* 1998;351:1652–1654.

23. Loue S, Okello D, Kawuma M. Research bioethics in the Ugandan context: a program summary. *J Law Med Ethics.* 1996;24:47–53.

24. Santelli J, Ginn E, Speers MA. An evaluation of human subject protection at CDC/ATSDR. *IRB.* 2001;22:1–6.

25. Brody BA. Making informed consent meaningful. *IRB.* 2001;23:1–5.

26. Zion D, Gillam L, Loff B. The Declaration of Helsinki, CIOMS and the ethics of research on vulnerable populations. *Nat Med.* 2000;6:615–617.

27. Hardy E. Factors often not considered before a multicenter trial is started. *Clin Pharmacol Ther.* 1996;60: 121–123.

28. Annas GJ. Questing for grails: duplicity, betrayal, and self-deception in postmodern medical research. In: Mann JM, Gruskin S, Grodin MA, Annas GJ, eds. *Health and Human Rights: A Reader.* 1st ed. New York, NY: Routledge; 1999:312–335.

29. *World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects.* Geneva, Switzerland: World Medical Association (WMA); 2000.

30. *CIOMS International Ethical Guidelines for Biomedical Research Involving Human Subjects—Revised Draft, January 2002.* Geneva, Switzerland: Council of International Organisations of Medical Sciences (CIOMS); 2002.

31. *Ethical and Policy Issues in Research Involving Human Participants. Recommendations of the National Bioethics Advisory Commission.* Washington, DC: National Bioethics Advisory Commission; 2000. Available at: http://www.bioethics.gov. Accessed April 23, 2002.

32. Loewenson R. Commentary: people centred science and globalisation: putting the public back into public health policy. *Int J Occup Environ Health.* 1999;5:65–71.

33. Macfarlane S, Racelis M, Muli-Musiime F. Public health in developing countries. *Lancet.* 2000;356:841–846.

34. Robertson A. Critical reflections on the politics of need: implications for public health. *Soc Sci Med.* 1998;47: 1419–130.

35. Hurst MH. The value of difference. *IRB.* 2001;23:6–8.

36. Equinet Steering Committee. Equity in health in Southern Africa: turning values into practice. Equinet Policy Series No. 7. September 2000. Equinet, Harare. Available at: http://www.equinetafrica.org. Accessed April 25, 2002.

37. Macaulay AC, Commanda LE, Freeman WL, et al. Participatory research maximises community and lay involvement. *BMJ.* 1999;319:774–778.

38. Mergler D. Worker participation in occupational health research: theory and practice. *Int J Health Serv.* 1987;17: 151–167.

39. Kriger N. A glossary for social epidemiology. *J Epidemiol Community Health.* 2001;55:693–700.

40. Emanuel EJ, Steiner D. Institutional conflict of interest. *N Engl J Med.* 1995;332:262–268.

41. Kelch RP. Maintaining the public trust in clinical research. *N Engl J Med.* 2002;346:285–287.

42. Butler D. Call for North/South code of research ethics. *Nature.* 2000; 406:337.

43. Rivera R, Ezcurra E. Composition and operation of selected research ethics review committees in Latin America. *IRB.* 2001;23:9–12.

44. Ignatieff M. *The Rights Revolution.* 1st ed. Toronto, Ontario: Anansi Press; 2000.

45. Mann JM. Human rights and AIDS: the future of the pandemic. In: Mann JM, Gruskin S, Grodin MA, Annas GJ, eds. *Health and Human Rights: A Reader.* 1st ed. New York, NY: Routledge; 1999:216–226.

46. Benatar SR, Singer PA. A new look at international research ethics. *BMJ.* 2000;321:824–826.

47. Pragmatism in codes of research ethics [editorial]. *Lancet.* 1998;351:225.

48. British Medical Association. *The Medical Profession and Human Rights. Handbook for a Changing Agenda.* 1st ed. London, England: Zed Books; 2001.

49. Greco Dirceu B. Revising the Declaration of Helsinki: ethics vs economics or the fallacy of urgency. *Can HIV/AIDS Policy Law Rev.* 2000;5:94–97.

50. Commission on Health Research for Development. *Health Research: Essential Link to Equity in Development.* 1st ed. Oxford, England: Oxford University Press; 1990.

The Effects of the Corset

Elizabeth Fee, Theodore M. Brown, Jan Lazarus, and Paul Theerman

THIS ILLUSTRATION APPEARED as a foldout in a little book, *Über die Wirkungen der Schnürbrüste* (On the Effects of the Corset), written by Samuel Thomas von Sömmerring.[1] The essay was published in 1793 and republished in an expanded edition in 1803. Translated into several languages, it became a best-seller.[2]

Von Sömmerring, a physician and well-known anatomist, argued that the back-laced corset, as worn by fashionable ladies of the time, constituted a heath hazard by compressing the ribs and other internal organs and leading—he claimed—to tuberculosis, cancer, and scoliosis, or curvature of the spine. His illustration contrasts the natural shape of the female body with the artificial hourglass shape produced by a tightly laced corset. In Germany, England, and the United States, dress reformers advocated looser lacing, pantaloons or "bloomers," and clothes that allowed more natural movement. Throughout the late 19th century, however, these reformers belonged to the radical fringe of the feminist movement, and their arguments led to much merriment in the popular press. Most middle- and upper-class women continued to compete for the tiniest waists, regardless of their impact on health. In more recent years, hiatus hernias caused by overly tight girdles or corsets have been termed "Sömmerring's syndrome" in tribute to the first physician to warn of the dangers of tight lacing.[3] ■

About the Authors

Elizabeth Fee, Jan Lazarus, and Paul Theerman are with the History of Medicine Division, National Library of Medicine, National Institutes of Health, Bethesda, Md. Theodore M. Brown is with the Department of History and the Department of Community and Preventive Medicine, University of Rochester, Rochester, NY.

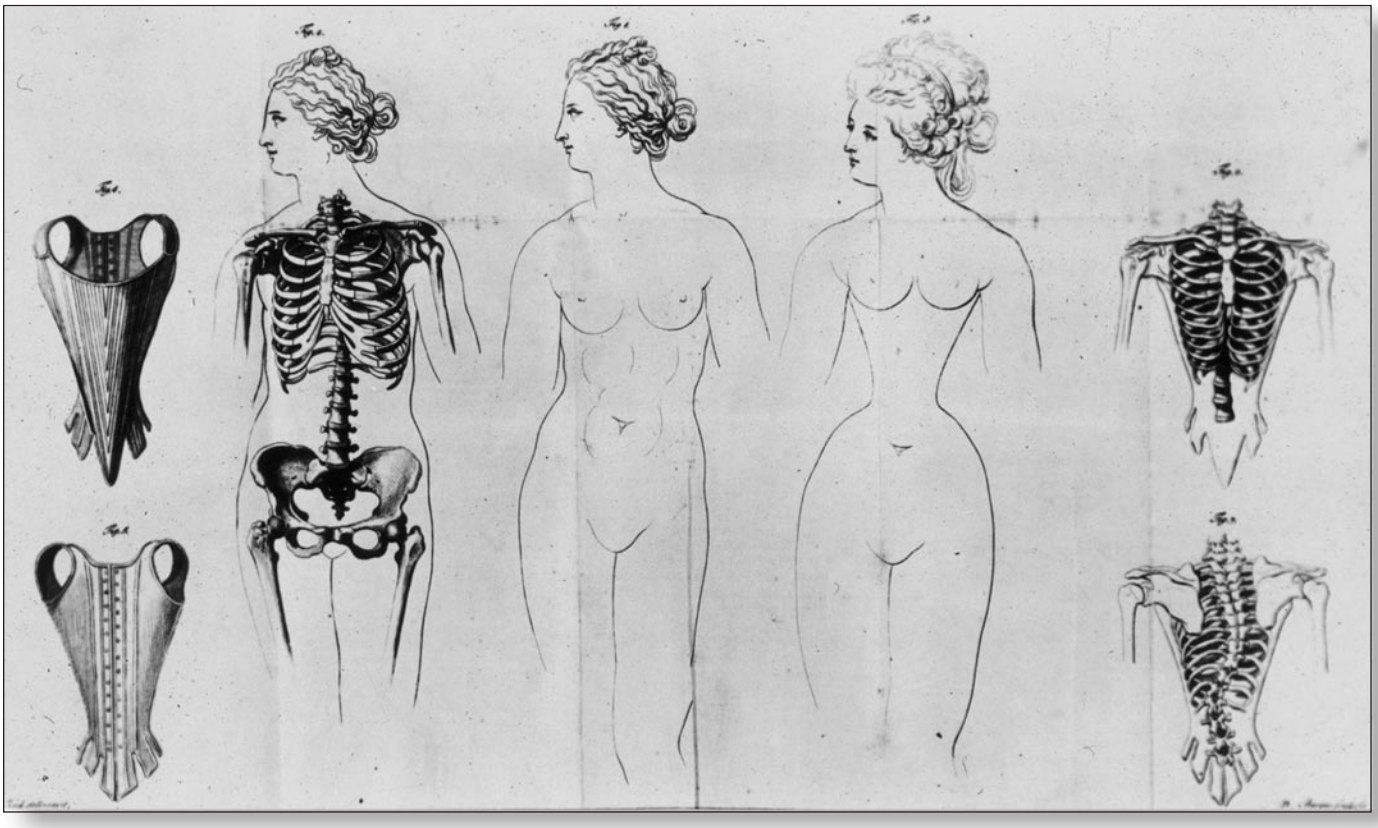
Requests for reprints should be sent to Elizabeth Fee, PhD, Bldg 38, Room 1E21, 8600 Rockville Pike, Bethesda, MD 20894 (e-mail: elizabeth_fee@nlm.nih.gov).

Note. *Most of the Prints and Photographs Collection of the History of Medicine Division of the National Library of Medicine may be viewed through the online database "Images From the History of Medicine" at http://wwwihm.nlm.nih.gov/. The Web site also provides information on ordering reproductions of images. If you have a print, photograph, or other visual item that might be appropriate for this collection, please contact the History of Medicine Division.*

References

1. von Sömmerring ST. *Über die Wirkungen der Schnürbrüste.* Berlin: Neue Auflage; 1793.
2. Schwarz GS. Society, physicians, and the corset. *Bull N Y Acad Med.* 1979;55:551–590.
3. Bell JA, Bell GC. "Tight girdle" or Sömmerring's syndrome. *N Engl J Med.* 1973;289:698.

Source. **Prints and Photographs Collection, History of Medicine Division, National Library of Medicine.**



Burning Love: Big Tobacco Takes Aim at LGBT Youths

Secret tobacco industry documents lay bare the industry's targeting, seduction, and recruitment of minority groups and children. They also unmask Big Tobacco's disdain for its targets.

| Harriet A. Washington

A DECADE AGO, FORMER Winston honcho David Goerlitz sneered that the R.J. Reynolds Tobacco Company had built its fortune by marketing to "the young, poor, black, and stupid."[1] A tobacco executive who risked such candor today might add, "the lesbian, gay, bisexual, transgender, and Hispanic." But such loose talk is very unlikely today because a series of legal reversals and losses in the court of public opinion have created an acutely circumspect tobacco consortium.

Decades ago, flagrant, disrespectful stereotypes marked the industry's initial courting of African Americans. Sports sponsorships, cartoon characters, and trinkets clearly labeled yesterday's marketing efforts to children and youths. But by the late 1980s, tobacco firms could read the writing on the billboard. Public health advocates and African American activists joined to protest such egregious forms of targeted marketing as the saturation of urban communities with billboards. Even more vociferous protests castigated the design and marketing of cigarettes and tobacco blends targeted exclusively at African Americans.

By the mid-1990s, Minnesota Attorney General Hubert Humphrey III wrested a legal settlement from the nation's major tobacco companies into which he incorporated a brilliant public relations stealth bomb: He forced the release and publication of Big Tobacco's secret internal marketing and research documents on the Internet for all to read. These documents laid bare, in the industry's own damning words, the oft-denied targeting, seduction, and recruitment of minority groups and children. They also unmasked Big Tobacco's disdain for its targets.

The ensuing spate of state and federal legal victories over Big Tobacco has, among other things, specifically banned traditional means of marketing to young people, such as cartoons, billboards, and advertisements in periodicals with significant youth readership.

These developments, while public health successes, have also served to drive tobacco's youth recruitment efforts underground, where they continue, shrouded in coded language and the all-too-familiar denials. The tobacco industry still boasts a marketing budget of $8.4 billion per year for the United States alone,[2] and the hidden truth about today's targeted marketing of vulnerable groups such as lesbian, gay, bisexual, and transgender (LGBT) youths will be even harder to excavate than was yesterday's.

Big Tobacco's past approaches toward targeting minorities, especially African Americans, are illuminating and may prove instructive for detecting its current ploys toward forbidden markets such as youths, especially LGBT youths. For example, Big Tobacco made loyal customers (and defenders) of many African Americans by lavishing positive attention on them when the rest of corporate America still considered Blacks to be marketing pariahs. Tobacco firms hired African Americans before other elements of corporate America welcomed or even accepted them, and tobacco firms infused languishing African American media, cultural, and advocacy groups with desperately needed financial support (box on page 1091).

But as the targeted-marketing backlash in the African American community has limited and sometimes stymied tobacco firms' influence, these companies have sought out lucrative new markets. Internal marketing memos show that the tobacco industry has scoured the globe for new target communities as

starved for corporate attention and acceptance today as African Americans were yesterday. These new targets include Hispanics, the fastest growing element of the population, and sexual minorities.

Today, tobacco firms are emerging from their corporate closets to openly engage in every type of marketing targeted at gay adults. Most alarmingly, the targeted marketing focuses on LGBT youths, but the cynical marketing snares for the young are carefully hidden and slyly labeled. Today, tobacco's corporate language is sanitized in a Newspeak of acronyms and is bowdlerized to delete any overt reference to youth marketing. The industry whose internal memos once blithely spoke of recruiting Black 14-year-olds is now careful to refer in print and in public only to "young smokers 18 and older."

SLAUGHTER OF THE INNOCENTS

The tobacco industry has never admitted targeting LGBT youths or even targeting youths at all, despite Joe Camel, cartoons, logo-rich children's trinkets, and sponsorship of youth-oriented sporting and music events. In the face of denials and the recent absence of loose-lipped memos, how can one know that Big Tobacco markets to LGBT youths?

The first clue is a look at the fruits of that targeting, because the overwhelming majority of adult smokers were once underage smokers, and almost no smokers, gay or straight, take up the habit after age 20. Tobacco companies know that they must hook a smoker as a child or not at all, and the US tobacco industry invests $23 million *every day* to ensure that they do.[6–8] Twenty-eight percent of high school students smoke, as opposed to 23% of adults. Every day, 5000 children take their first puff; 2000 are unable to stop and thus swell the ranks of the nation's smokers. One third of the addicted will die from their smoking habits—and this doesn't count the 14% of boys who become addicted to the smokeless tobacco popularized by generations of sports heroes who chew, dip, and spit very publicly.[2]

> *"Tobacco companies have always been aware of sexual minorities as customers who smoked at extremely high rates, but only in the last decade have they embraced marketing targeted at gays in earnest and in large numbers."*

A number of studies have determined that children are 3 times as susceptible to tobacco advertising than adults and that such advertising is a more powerful inducement than is peer pressure.[9,10]

Ugly as this picture is, the prospects of avoiding tobacco addiction are much bleaker for LGBT youths. The prevalence of smoking is around 46% for gay men and 48% for adult lesbians,[11,12] twice as high as for their peers. Data on bisexual and transsexual smoking behavior are sparse, as are data on the smoking behavior of sexual minorities who are also members of high-risk racial and ethnic minority groups. Still, smoking prevalence is likely to be disastrous at the intersection of such high-risk groups.

The smoking rates of LGBT youths are just as high as those of adults, which is hardly surprising. And not only do twice as many LGBTs as other Americans take up smoking; they find it harder to quit, although most want to. Eighty percent of the 1011 adult respondents in a 2001 American Medical Association (AMA)–Robert Wood Johnson (RWJ) Foundation poll said they had tried to stop smoking but could not.[13] Like African Americans, members of sexual minority groups pay a much

Gay and lesbian adults smoke twice as much as their peers.

"The tobacco industry has scoured the globe for new target communities as starved for corporate attention and acceptance today as African Americans were yesterday. These new targets include Hispanics, the fastest growing element of the population, and sexual minorities."

Linking cigarettes, alcohol, and the bar scene has been a staple of tobacco marketing strategies.

higher medical price for their tobacco addiction. The direct health effects of smoking for gays and lesbians are legion, although the exact figures are still a matter of debate.[11] Lesbians who use tobacco face risks of breast cancer, colorectal cancer, and other cancers 5 times higher than those of other women.[11] The Centers for Disease Control and Prevention (CDC) calculates that the life expectancy of a homosexual man is 8 to 20 years less than that of other men[14] and that high smoking rates contribute directly and indirectly to their early deaths. Some researchers fear that cigarette smoking may increase the risk of HIV infection and accelerate the progression to AIDS.[15] HIV-positive men have the highest smoking rates of all,[15] and heavy smoking triggers immune function changes that also worsen the prognosis for other infectious diseases such as sexually transmitted diseases and hepatitis C.[15,16]

Despite all these studies, surveys reveal that gay smokers do not believe that smoking a pack a day constitutes a health risk.[17] This is a dangerous attitude, because the health effects of smoking are not always related to the dose.

The scourge of tobacco addiction doesn't wait for adulthood to erode the mental and physical health of LGBT youths. These youths already face much higher vulnerabilities to violence, suicide, and risk-taking behavior (including risky sexual behaviors) than their peers. Young gay smokers are the greatest risk takers; their higher rates of alcohol and drug use[15] lead many experts to characterize tobacco as a "gateway drug" for gay youths.

LGBT youths also pay a high price in direct health effects of their smoking addiction, such as 30% to 87% higher rates of cancer. The synergistic effect of tobacco and alcohol encourages a constellation of other respiratory diseases and of ear, nose, and throat diseases; early smoking also inflates the lifetime risks not only of premature death but also of impairments such as blindness and infertility.

Ninety percent of smokers start in their teens,[17] and LGBT smokers start even younger; in one survey, 13 years was the median age for girls.[18] Tobacco firms therefore know that their efforts to target gays and lesbians will work only if they target gay youths younger than 18, the legal smoking age in most states.

The tobacco industry's targeting of the LGBT communities is a matter of record. Alcohol companies, many owned by tobacco firms, have targeted gays[12] as far back as the 1950s, when Joseph Cotton, with a hand resting on his double, slyly touted Smirnoff vodka, "mixed or straight."[19] Tobacco companies have always been aware of sexual minorities as customers who smoked at extremely high rates, but only in the last decade have they embraced marketing targeted at gays in earnest and in large numbers.

The secret tobacco documents placed on various Web sites afford revealing insights into the industry's changing perception of the LGBT communities. Internal memos reveal that tobacco companies sought gay voters' support as early as 1983,[2] when they wished to repeal workplace smoking bans in San Francisco.[12] An internal Philip Morris memo from 1985 reveals grudging admiration at how views of gays and lesbians as customers were changing: "It seems to me that homosexuals have made enormous progress in changing their image in this country. . . . A few years back they were considered damaging, bad and immoral, but today they have become acceptable members of society. . . .We should research this material and perhaps learn from it."[20]

A few years later, when African Americans were successfully protesting targeted community saturation by tobacco firms, the tobacco industry began openly contemplating sexual minorities as a less troublesome market. In the 1990s, just after the protests that aborted the marketing of Uptown to urban African Americans, a "Top Secret Operation Rainmaker" memo listed gays as a marketing "issue" to be discussed.[21] Paradoxically, this marketing attention was catalyzed by

a concerted political attack mounted on the tobacco consortium by gays. When the AIDS Coalition to Unleash Power (ACT-UP) organized a 1990 boycott of Philip Morris over its support of Jesse Helms, tobacco companies responded by donating large funds to AIDS organizations in appeasement efforts,[12] just as they have showered politically pivotal African American organizations with money. Tobacco firms swiftly followed these overtures to gays and lesbians with national advertising campaigns. In 1991, a *Wall Street Journal* headline trumpeted, "Overcoming a deep-rooted reluctance, more firms advertise to [the] gay community." The story called gays and lesbians "a dream market" and focused on the tobacco industry's courtship of LGBT media giants such as *Genre*.[22]

Between 1990 and 1992, a series of ads for American Brands' Montclair featured an aging, nattily dressed man sporting an ascot, a pinky ring in lieu of a wedding ring, a captain's hat, and an orgiastic expression. This was perceived as a gay or effeminate persona by many readers and media analysts, a reading that American Brands denies. Benson & Hedges, for its part, touted a series of gay-themed ads for its Kings brand in *Genre*, a gay fashion and lifestyle magazine, as well as in *Esquire* and *GQ*, which have significant gay male readerships.

Recent marketers have not overlooked lesbian and bisexual women.[12] Philip Morris's Virginia Slims ads send messages of independence, camaraderie, and iconoclasm that appeal to feminists as well as lesbians. But the ads, appearing in such magazines as *Essence* and *Ms.*, have departed from women in lipstick and heels to feature more androgynous and sexually ambiguous portrayals of women. Women couples are shown fishing in plaid shirts, mesh vests, and hip boots; women duos in tailored clothes are captured in a tête-à-tête over coffee; women even throw appreciative glances at each other on the street over text that exhorts them not to "follow the straight and narrow."

Marlboro, the most popular cigarette brand among gay men, has flaunted the brand's rugged hypermasculine image in venues calculated to appeal to gays. For example, one large billboard features a close-up of a substantial male crotch clad in weathered jeans with a carton of Marlboros slung in front of it. The image hangs between 2 gay bars in San Francisco's Mission district.[12]

The health advocacy backlash to Big Tobacco's first flirtation with gay media was swift and sure. The Coalition of Lavender Americans on Smoking and Health (CLASH) issued a press release that read in part, "This is a community already ravaged by addiction: we don't need the Marlboro man to help pull the trigger."[12]

BEYOND BILLBOARDS

If the experiences of racial minorities—notably African Americans—serve as a guide, targeted advertising will be just the tip of the tobacco iceberg. The underwriting of key cultural institutions is another insidious route to the control of minority lungs. So is control over the specialized news media that minorities trust. Until 5 or 6 years ago, nearly all African American publishers retained, and sometimes defended, alcohol and tobacco advertisers. For example, in 1998, Dorothy Leavell, who was then president of the National Newspaper Publishers Association, said, "Adult African Americans are mature enough to make [their] own decisions unless government makes tobacco illegal." Leavell acknowledged that tobacco money provided key support for the 121 publications her group represented: "The tobacco-settlement negotiations have hurt our publications dollarwise." She lamented the fact that tobacco advertising revenues fell "from a peak of 12–20 million a year [in 1995] to less than 6 million [in 1998]" (D. Leavell, oral communication, 1998).

Many African American publications still embrace and defend tobacco advertising, not from choice, but out of desperation. Alcohol and tobacco corporations have long showered African American publications with advertising and philanthropic revenue while other national corporate advertisers shunned their pages, taking African American consumers' money but refusing to advertise in such important publications as *Ebony*, *Essence*, or *Jet*. Smaller magazines and newspapers found corporate support even more elusive and were often completely dependent upon the tobacco industry (which owns many alcohol companies). Curious news policies ensued; medical writers, for example, were ordered to avoid the topic of smoking, and articles about cardiovascular disease were published that did not mention tobacco use as a risk factor.

Key African American advocacy organizations such as the Urban League and the National Association for the Advancement of Colored People (NAACP) were gratified not only by large be-

Blacks and other racial/ethnic groups have been targeted by the tobacco industry through ads such as this Virginia Slims ad, with its suggestive undertones.

"Alcohol and tobacco corporations have long showered African American publications with advertising and philanthropic revenue while other national corporate advertisers shunned their pages."

quests but the presence of highly placed, highly visible African American tobacco company staff. For African Americans, tobacco has been a good corporate friend in a hostile corporate universe. In 1992, "R.J. Reynolds also decided "to improve [the employment] recruitment process by the development of an attractive corporate image to be systematically utilized in recruitment ads. . . . [F]eel-good 'social-responsibility' campaigns by tobacco companies help the industry not only to sway political and public opinion but to continue to recruit effective salespeople and boost employee morale." This thrust includes hiring more sexual minorities, supporting their organizations, and addressing issues of importance to LGBT communities.[23] In this context, consider the insights of Janelle Lavelle, whose gay advocacy group opposed Senator Jesse Helms's 1990 reelection. At that time, Lavelle told *The Advocate*, "The protests make working with other minority groups . . . harder because Philip Morris is one of the most labor-positive and minority-positive corporations in North Carolina." She added, "Philip Morris has openly gay people working at several places. As North Carolina companies go, Philip Morris is a jewel."[24]

Today, the targeting of gays and lesbians is escalating because gays constitute a very attractive consumer market. They boast a high disposable income and are "attention-starved and very loyal," according to Jeff Vitale, president of Overlooked Opinions, an LGBT marketing firm.[12] For example, 94% of gay readers in one survey said that they would support advertisers in gay magazines and contributors to gay organizations.[12]

And, of course, LGBTs have high smoking rates.

LGBT YOUTHS IN THE CROSSHAIRS

This supremely attractive LGBT market has an important feature in common with both heterosexuals and with the racial minorities on which Big Tobacco cut its targeted-marketing teeth: It can be captured and retained only by attracting potential smokers while they are very young.

A 1981 secret Lorillard memo asks rhetorically, "Where should our marketing thrust be?" and replies, "keep riding with Newport" because it is "heavily supported by Blacks and under-18 smokers. We are on somewhat thin ice should either of these two groups decide to shift their smoking habits."[15]

Unfortunately, federal and state lawsuits have not ended the seduction of children by Big Tobacco. By 1998, a flurry of successful state and federal lawsuits had snatched from the industry's bag of marketing tricks such options as advertising at sporting events and on billboards and in children's publications. However, Big Tobacco quickly replaced such traditional sponsorships with increased product placement in films, logo-rich announcements, sponsorship of alternative music clubs, and moving billboards on taxis. Such venues are more likely to reach and to appeal to some LGBT youths than traditional sporting events and ads in heterosexual women's and men's magazines. A 2002 University of Chicago report documents that despite the explicit 1998 prohibitions, the 3 largest US tobacco companies have selectively increased youth targeting.[26] "Cigarette companies had to become slightly more subtle about it, but they continue to aim their advertising at people under 18," avers Paul Chung, MD.[26]

Another reason for the tobacco industry's subtlety in targeting LGBT youths may be that it found young teens responded negatively to overtly sexual messages, according to secret docu-

Sold Down Tobacco Road

WHAT WERE AFRICAN AMERICANS' favorite radio stations in 1967—alphabetized, by city? How many owned automobiles, and where were these car owners most likely to live? What was the African American median income in 1971? How many hours of television did the average Black man watch that year? Did he prefer flavored, filtered, or menthol cigarettes? And exactly how cool a menthol cigarette did he like to smoke–to *the degree*?

Ask a tobacco company.

The marketing and biochemical research documents that were released as part of Minnesota's 1996 settlement with Big Tobacco reveal a staggeringly exhaustive research dossier on African American culture, habits, physiology, and biochemistry. This intimate portrait of Black America enabled the design of special tobacco products developed with African Americans in mind. The special blends were then painstakingly test marketed down to the smallest cynical details; these included packaging with an Afrocentric red, black, and green color scheme, an "X" logo that evokes political hero Malcolm X, and packs that opened from the bottom because demographic surveys revealed that this was a favored practice of Black men.

The evolution of marketing targeted at African Americans is too complex to describe in great detail here, but its history reveals some important parallels to the way in which Big Tobacco is making overtures to the LGBT communities.

Secret tobacco company documents available on the Internet reveal how the industry meticulously researched African American habits and manipulated corporate and media leadership. The early efforts of the 1950s, 1960s, and 1970s were crudely stereotypical. Newspapers were to be eschewed in favor of musical advertisements, because "The beat, the tempo, and the 'feeling' of the 'Soul' music is almost instinctively identifiable to the Negro ear, which is accustomed to this sound." [3]

"... 'Outdoors' (hunting, skiing, sailing) is not felt to be suitable, as these are considered unfamiliar to the Negro..."[3] The Montclair ads of the 1990s and the Project Scum campaign (see text) mirror such offensive stereotypes.

However, later documents revealed an appalling sophistication both in content and in quality as the targeted marketing became subtler. With the aid of hundreds of surveys, focus groups, and social and even biochemical studies of African Americans, tobacco companies researched and planned very effective marketing campaigns. They even launched special brands, mostly menthol, specifically targeted to African Americans, just as Marlboros, Virginia Slims, and Camels are marketed specifically to LGBT youths today, through marketing campaigns such as Project Scum.

African American communities have been saturated with billboards and placards featuring ethnically diverse smokers, frequent cigarette giveaways, and nefarious tobacco products such as "blunt wraps"—rolls of tobacco used to hold marijuana. For a quarter, children and the poor can buy "loosies," or illegal single cigarettes, much more frequently in communities of color than in other areas.

But Big Tobacco's siege of the African American community didn't stop there. It proceeded to kill with corporate kindness. In the 1950s, boardrooms and marketing plans tended toward the male and monochromatic. But the top 5 American tobacco companies had discovered that there was green in Black communities, and the tobacco industry offered African Americans influential work and well-paid careers when other companies barred Blacks from their boardrooms.

Just as tobacco companies today hire openly gay staff and pour funds into LGBT organizations, both local and national, they filled the cash-starved coffers of nearly every influential African American organization from the National Urban League to local churches. Tobacco has been the best corporate friend that Black America ever had, pouring so much money into political, social, artistic, and religious organizations that, like hooked smokers, these pivotal African American organizations cannot function without tobacco. Tobacco companies did not forget to invest in important politicians and to shower advertising and other financial support on influential African American news and entertainment media, which have become dependent on this habitual largess.

Today, the 28% smoking rate among African Americans is higher than the national norm. After dropping briefly during the antitargeting backlash of the 1990s,[4] smoking rates among African American youths are climbing again. What's worse, 75% of male African American smokers use menthol cigarettes. This is no coincidence, suggests Philadelphia lawyer William Adams, who has brought targeted marketing cases against Big Tobacco.[5] He recalls that "In 1957, only 5% of African Americans who smoked consumed menthol products. This represents fantastic growth. The taste for menthol was carefully cultivated by the tobacco industry itself" (W. Adams; oral communications; February 20, 1998, and May 14, 1999).

The 1990s saw a burgeoning African American opposition to targeted marketing, especially to cynically specialized cigarette brands such as Uptown, Camel Menthol, and X, and to the billboards and ads that target minority children. Such opposition has made African Americans increasingly difficult and expensive consumers. The first Surgeon General's Report to focus on the smoking habits of racial and ethnic groups further damaged the relationship between Big Tobacco and African Americans by documenting the persistent targeting of young African Americans.

But the African American community remains vitally important to the cigarette industry. African Americans represent nearly 14% of the population and 38% of its smokers (vs 31% for the total population). African Americans still constitute a large total market share of 8 top brands, especially Kool (20.5%).[4]

ments generated as early as 1978. A trial advertising campaign for Old Gold filters found that sexually referential advertising "produces no increase in brand switching or awareness and that it does not contribute to the success of Old Gold lights." The report cited children's negative reactions to "the whole sexual erection thing . . . 'get it on' . . . ugh . . . I wouldn't go for it. . . ."[27]

The news media made much of the high recognition of Joe Camel among young children and the plans to use sweet flavorings in cigarettes. More apropos to LGBT youths, however, are marketing strategies directed at youths who are beginning to recognize their sexual identities. For example, in one internal memo, tobacco companies muse how helpful it would be if they could discover that tobacco helps a common health concern such as acne.[28]

Access to cigarettes also remains easy for children. Although cheap "loosies" are more available in minority neighborhoods to African American children, children of all ethnicities can easily buy tobacco products from vending machines and the Internet. A 1983 study showed that 25% of children younger than 13 purchased cigarettes from vending machines.[29]

Big Tobacco's canon of Newspeak includes an important veiled reference to young smokers: FUBYAS, the acronym for First Usual Brand Younger Adult Smokers. The first brand smoked by a youth is tremendously important to marketers. Smokers are notoriously brand loyal, making it expensive and difficult to induce them to change brands. Therefore, tobacco companies seek to induce seduction by and loyalty to a brand, not to the smoking habit. The tobacco documents reveal a great deal of finely detailed research to tailor tobacco brands to very specific populations and their subgroups. "FUBYAS" are youths because youths constitute the vast majority of "first" smokers. Thus, targeting FUBYAS is, by definition, marketing to youths. For gay markets, this is marketing targeted at gay youths. We know this because tobacco companies do not target children in broad strokes but rather narrowly, by race, by income, and by personality, as illustrated by tobacco documents full of stereotyped descriptions. "The elements that FUBYAS know are their social groups. These are large, loosely knit but highly labeled sub-societies from which FUBYAS draw their identity. . . . The FUBYAS readily classifies others into the groups, and knows what his 'membership' is."[30] "This is good news, because therein lies differentiation and opportunity," concludes a tobacco company document cited in the *Washington Post*.[30]

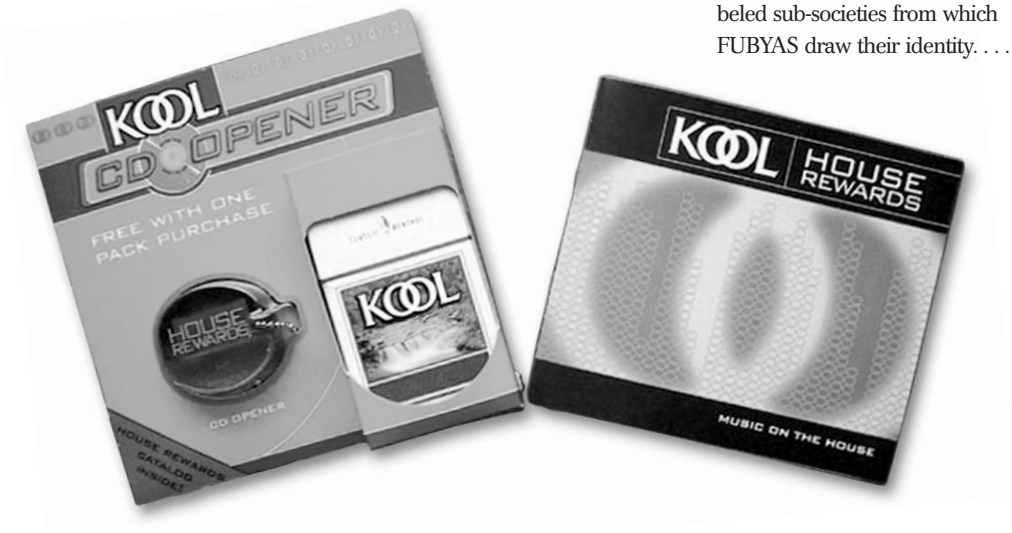
For example, a June 1994 tobacco document observes, "research indicates that adult smokers are of different sexes, races and sexual orientations."[31] The descriptions of these targets are rarely flattering. For example, More brand cigarettes are targeted at women—but not just any women, according to internal marketing memos: "Woman—liberated, but not ball-busting." Users of generic cigarettes are described as "not mentally stable . . . imageless, hobos, tramps, rag pickers."[32,33]

Companies also target children by sexual orientation. As with lesbian Virginia Slims and gay male Montclair campaigns on the early 1990s, the targeting is covert, but it exists. Strategies by which tobacco companies have increased the targeting of LGBT youths include the following:

• *Exploiting the bar culture.* In one survey, 32% of lesbians and gays cited the bar culture as a factor in their nicotine addiction.[12] Drinking in bars fosters the smoking habit by lowering inhibitions and fuels LGBT alcoholism rates, which are 3 times the national average.[12] The synergy of alcohol and smoking is a special hazard for gays, as it multiplies the risks of cancer and other diseases.

The importance of the bar culture is not lost on the industry, whose documents show a heavy marketing investment in bar crowds, including "African American" and "alternative" bars

Cigarettes are sometimes packaged with free gifts that appeal to both young adults and youths.



where tobacco company representatives distribute free cigarette samples, hold contests with tobacco premiums as prizes, and sometimes take over the bar to buy free drinks.

• *Exploiting drug use rates,* which are higher among LGBT youths, a fact that Big Tobacco has not hesitated to use to its advantage.

• *Targeting geographical areas* where many young LGBTs congregate socially, such as the Castro and Tenderloin districts of San Francisco.

PROJECT SCUM

The industry that has always denied that it targets youths, despite the detailed statements, reports, and youth marketing plans in its secret internal memos, today refuses to admit its targeting of LGBT youths. Unfortunately, in the absence of a confession, hard evidence is hard to come by. But hard evidence of plans to target LGBT youths can sometimes be excavated from internal tobacco marketing documents that were never meant for consumers' eyes.

Between 1995 and 1997, R.J. Reynolds internal documents recorded corporate-wide overtures to the young LGBT community in what must be one of the least flattering targeted marketing plans in history. In "Project Scum," R.J. Reynolds tried to market Camel and Red Kamel cigarettes to San Francisco area "consumer subcultures" of "alternative life style." R.J. Reynolds's special targets were gay people in the Castro district, where, as the company noted, "The opportunity exists for a cigarette manufacturer to dominate."[34] The gay Castro denizens were described as "rebellious, Generation X-ers," and "street people." Both the coded labeling of targets as "Generation X-ers" in the mid-1990s and as "rebellious" indicates their youth. Project Scum also planned to exploit the high rates of drug use in the "subculture" target group by saturating " 'head shops' and other nontraditional retail outlets with Camel brand." More telling than the guardedly worded report itself are the scrawled marginalia such as "Gay/Castro" and "Tenderloin." The observation "higher incidents [sic] of smoking in subcultures" has the phrase "and drugs" written in.[34–36]

In one copy, the word "Scum" is crossed out and the word "Sourdough" substituted by a belatedly cautious executive. After such careful sanitizing, the final document could have emerged as Project Sourdough with no clear written evidence that young LGBTs had ever been targeted.

The readiness such documents evinced to exploit the higher drug use rates among LGBT youths sheds marketing light on some dubious promotional devices. For example, Philip Morris has given away key rings with its Alpine Extra Light cigarettes that conceal a screw-top glass vial. Surveys by the smoking cessation group Quit, from Melbourne, Australia, suggest that most youths think these premiums are for holding drugs. Unprompted, 10 of the 13 groups surveyed suggested that the vials were meant to carry "drugs," "coke," "stash," and "speed." One youth summed it up: "It's a key ring, and it's what people typically use to carry drugs."[37]

STRANGE BEDFELLOWS

In the early 1990s, tobacco companies began forging legislative and political ties with the LGBT community, just as they had with many African American politicians, the African American news media, and with groups such as the National Urban League. The tobacco industry began introducing antidiscrimination bills that purported to offer protection against sexual-orientation bias but whose addenda and riders protected the rights of smokers—and of the companies that feed smokers' habits. According to the Tobacco Institute, by September 23, 1992, the industry had enlisted the help of local gay and lesbian alliances to pass smoker protection/"antidiscrimination" laws in 20 states.[38]

> *"Philip Morris sought to expand its pursuit of the US Hispanic community by sponsoring soccer events, Latin music events, and Hispanic festivals and by giving away cigarettes."*

Sometimes, however, tobacco companies knew that anti-tobacco LGBT groups troubled by high rates of LGBT addiction would not support them, so they formulated plans to cut these leadership groups out of the picture and appeal directly to gay voters.

In 1998, California voters were offered Proposition 10, a measure to substantially hike cigarette taxes. A concerned Tobacco Institute (an industry consortium) sought professional advice from marketers, researchers, and consultants on how best to garner the gay vote. Consultant David Mixner advised the Tobacco Institute "to bypass" gay organizations: "Since it is apparent that we are not going to have the endorsement of most Gay and Lesbian leadership, it is important to use these campaign tools to bypass

that and go directly to the Gay and Lesbian voter with a message that will resonate."[39]

Just as struggling African American media tended to defend (or to conveniently ignore) targeted marketing by Big Tobacco, some LGBT media view tobacco's growing attention to gay and lesbian customers as a boon.

In response to a August 17, 1992, *Wall Street Journal* story by Joanne Lipman, Don Tuthill, the publisher of *Genre*, wrote, "The Philip Morris/Genre story is not a story of

> 'the tobacco industry . . . turning its marketing muscle on another minority'; it is a story of inclusion . . . a story about a major marketing company recognizing, including and supporting a frequently disenfranchised and overlooked segment of the U.S. population. . . . this is exactly the kind of support that gay media has sought for years."[40]

Joe Landry, publisher of *Out* and *The Advocate*, crowed that Big Tobacco's advertising "shows we're making progress. . . . Lots of companies are adding diversity marketing to their budgets, which used to mean money for advertising mainly to blacks and Hispanics, but now it's meant largely, and sometimes mainly, for gay and lesbian customers."[41]

However, the targeting of racial and ethnic groups is hardly a thing of the past; indeed, it is escalating and gaining precision. For example, a perusal of the online tobacco documents shows that tobacco companies have added surgically precise research on Hispanic markets to their ethnic marketing mix.

As early as 1988, Philip Morris sought to expand its pursuit of the US Hispanic community by sponsoring soccer events, Latin music events, and Hispanic festivals and by giving away cigarettes. The proposed budget for the Hispanic Marlboro campaign alone was $3.5 million.[42–44]

Recently, however, such targeting has achieved sophistication; it is no longer in broad strokes but rather surgical strikes. For example, although Lorillard market research details that African American men are the largest market for menthol cigarettes, in some Los Angeles and New York Hispanic communities, menthol is perceived as the choice of gay men. But in others menthol is seen as macho, and in still others as an acceptable choice for anyone, man or woman.[42]

In sum, one can see dramatic parallels between tobacco's courtship of the LGBT community today and its targeted marketing of African Americans decades ago. The industry gave African Americans the corporate attention and acceptance they craved, advertising in their media and neighborhoods. It became a generous and attentive corporate friend, investing heavily in African American organizations of every stripe. It hired many African Americans and non-White women in positions of responsibility in an era when boardrooms were male and monochromatic.

Behind the scenes, however, Big Tobacco's secret documents reveal that it slyly targeted children and knowingly sabotaged the health of a community already ravaged by high disease rates.

Today, tobacco companies view both racial and sexual minorities as key markets. Targeted marketing of tobacco to sexual minorities seems likely to mirror the process that worked so powerfully to addict and sicken African Americans.

Marketing targeted at LGBTs will probably accelerate, for several reasons:

- The LGBT community is an extremely attractive new consumer base, with high disposable income, a very brand-loyal culture, a tradition of tobacco use, and a hunger for corporate attention and acceptance.
- The Internet will make it irresistibly quick and easy to precisely target and survey specific demographic groups within the LGBT community for specific tobacco products, much easier than detailing racial minorities has been during the past 40 years.
- The Internet also offers a way to market and sell cigarettes to underage youths, including LGBT youths.

All this means that we may see marketing aimed at LGBTs—and, by implication if not definition, at *young* LGBTs—escalate in scope and directness. The wave of the future may be revealed by Germany's Reemstma tobacco company, whose advertisements have thrown yesterday's coded ambiguity to the winds. Ads for its New West brand have featured a gay male marriage and copy that exhorted, "Men! . . . taste how strong this one is."[12] Reemstma's Web sites include one labeled "Queer," which offers a dazzling cornucopia of advertisements, news, fashion, music, games, chat rooms, explicit gay pornography, and, of course, tobacco products, all tailored to the tastes and interests of LGBT consumers.[45]

What better illustration of how Big Tobacco increasingly views the LGBT community with lust rather than apprehension? As a growing number of companies court the gay market, analysts predict that this desirability will achieve widespread corporate acceptability. Then tobacco advertising may become omnipresent

within the LGBT communities, just as tobacco billboards, ads, promotions, sponsorship, employees, and products did in African American communities by the 1970s, seducing new generations of smokers. ■

About the Author

Harriet A. Washington is with the Harvard Medical School, Boston, Mass.

Requests for reprints should be sent to Harriet A. Washington, 4049 Broadway, #123, New York, NY 10032 (e-mail: haw95@aol.com).

This article was accepted April 16, 2002.

References

1. Tobacco Institute memo from Samuel Chilcote, Jr, to Rep. Thomas Luken's Subcommittee on Transportation and Hazardous Materials [re: Protect Our Children From Cigarettes Act of 1989]. July 26, 1989. Bates No. 515687450. Available at: http://tobaccodocuments.org/rjr/515687449-7757.html. Accessed May 15, 2002.

2. The campaign for tobacco-free kids. Available at: http://tobaccofreekids.org/adgallery. Accessed April 21, 2002.

3. Holland G. A study of ethnic markets [by R. J. Reynolds marketing officials]. September 1969. Bates No. 501989230. Available at: http://www.tobaccodocuments.org. Accessed June 17, 1998.

4. Stolberg SG. Rise in smoking by young blacks erodes a success story. *New York Times.* April 4, 1998: A24.

5. Slobodzian J. Civil rights suit names tobacco. *Natl Law J.* November 2, 1998:A10.

6. DiFranza JR, Librett JJ. State and federal revenues from tobacco consumed by minors. *Am J Public Health.* 1999;89:1106–1108.

7. The toll of tobacco in the United States of America. Available at: http://tobaccofreekids.org/research/factsheets/pdf/0072.pdf. Accessed May 23, 2002.

8. Cummings KM, Pechacek T, Shopland D. The illegal sale of cigarettes to us minors: estimates by state. *Am J Public Health.* 1994;84:300–302.

9. Evans N, Farkas A, Gilpin E, Berry C, Pierce JP. Influence of tobacco marketing and exposure to smokers on adolescent susceptibility to smoking. *J Natl Cancer Inst.* 1995;87:1538-1545.

10. Pierce JP, Choi WS, Gilpin EA, Farkas AJ, Berry CC. Tobacco industry promotion of cigarettes and adolescent smoking [with erratum in *JAMA.* 1998; 280:422]. *JAMA.* 1998;279:511-505.

11. *Lesbians and Smoking Fact Sheet.* Washington, DC: NOW Foundation; October 16, 2001. Available at: http://www.nowfoundation.org/health/lybdkit/factsheet.html. Accessed March 4, 2002.

12. Goebel K. Lesbians and gays face tobacco targeting. *Tob Control.* 1994;3: 65-67.

13. Author unknown. Most smokers try to quit, but can't [poll]. September 11, 2001. Available at: http://www.umich.edu/~umtrn/websites.html. Accessed September 11, 2001.

14. Bosworth, Brandon. A last drag? *Am Enterprise Online.* November 15, 2001. Available at: http://www.theamericanenterprise.org/hotflash011115.htm. Accessed May 23, 2002.

15. Ardey DR, Edlin BR, Giovino GA, Nelson DE. Smoking, HIV infection and gay men in the United States. *Tob Control.* 1993;2:156–158.

16. Newell GR, Mansell PW, Wilson MB, Lynch HK, Spitz MR, Hersh EM. Risk factor analysis among men referred for possible acquired immune deficiency syndrome. *Prev Med.* 1985;14:81-91.

17. Paralusz KM. Ashes to ashes: why FDA regulation of tobacco advertising may mark the end of the road for the Marlboro Man. *Am J Law Med.* Spring 1998;24:89–122.

18. Rosario M, Hunter J, Gwadz M. Exploration of substance abuse among lesbian gay and bisexual youth: prevalence and correlates. *J Adolesc Res.* October 1997;12:454–476.

19. Steele BC. From pickups to SUVS: winking ads that appealed to gay men as early as the 1950's lead the way for today's bold appeals to lesbian and gay consumers. *The Advocate.* August 15, 2000:14.

20. Philip Morris marketing memo, March 29, 1985. Bates No. 2023268366/8374. Available at: http://www.pmdocs.com. Accessed March 1, 2002.

21. Top Secret Operation Rainmaker, Tuesday, 20 March 1990. Bates No. 2048302227–2048302230. Available at: http://tobaccodocuments.org/usc_tim/1251.html. Accessed May 20, 2002.

22. Rigdon JE. Overcoming a deep-rooted reluctance, more firms advertise to gay community. *Wall Street Journal.* July 16, 1991: B1.

23. Institute for Consumer Responsibility. Georgians Against Smoking Pollution newsletter. 1984 (est.). Bates No. 2045748385-8398. Available at: http://tobaccodocuments.org/usc_tim/2045748385-8398.html. Accessed March 6, 2002.

24. Institute for Consumer Responsibility. [No title]. 1984 (est.). Bates No. 2045748385-2045748398. Available at: http://tobaccodocuments.org/pm/2045748385-8398.html. Accessed March 28, 2002.

25. Replies to 5-year plan questionnaire. 1981. Bates No. 03490806-03490844. Available at: http://tobaccodocuments.org/youth/AmYoLOR00000000.Rm.html. Accessed May 20, 2002.

26. Cigarette ads target youth, violating $250 billion 1998 settlement [press release]. Chicago, Ill: University of Chicago Hospitals & Health System, Public Relations Office; March 12, 2002.

27. OGF Vitality Test WAVE II. Bates No. 91145517-91145643. Available at: http://tobaccodocuments.org/lor/91145517-5643.html. Accessed May 23, 2002.

28. Memo suggesting consideration of formation of a scientifically oriented group to review work made at tobacco industry expense. Tobacco Institute. June 19, 1974. Bates No. TIMN0068332. Available at: http://www.smokescreen.org. Accessed March 7, 2002.

29. Kessler DA, Witt AM, Barnett PS, et al. The Food and Drug Administration's regulation of tobacco products. *N Eng J Med.* 1996;335:988–994.

30. [Author unknown.] Cigarette marketers' decades-old stereotypes may threaten today's deal. *Washington Post.* January 19, 1998: A23.

31. Landman A. Confidential—recap of Hispanic meeting. Bates No. 2060286766. Available at: http://tobaccodocuments.org/landman/2060286766.html. Accessed March 8, 2002.

32. Bender G. Image attributes of cigarette brands. For Brown and Williamson Tobacco Corp. January 1982. Bates No. 670575136-670575158. Available at: http://tobaccodocuments.org/usc_tim/247684.html. Accessed April 1, 2002.

33. Mindset segments. January 3, 1991. Bates No. 510320848– 510320876. Available at: http://tobaccodocuments.org/landman/510320848-0876.html. Accessed April 1, 2002.

34. Simkins BJ. Project Scum. R. J. Reynolds Tobacco Co. December 12, 1995. Bates No. 518021121–518021129. Available at: http://tobaccodocuments.org/landman/518021121.html. Accessed April 4, 2002.

35. Project Scum. R. J. Reynolds Tobacco Co. April 2, 1996. Bates No. 519849940–519849948. Available at: http://tobaccodocuments.org/rjr/519849940-9948.html. Accessed April 4, 2002.

36. O'Conner LM. Project Scum. R. J. Reynolds Tobacco Co. 1997. Bates No. 517060265–517060272. Available at: http://tobaccodocuments.org/rjr/517060265-517060272.html. Accessed January 28, 2001.

37. Chapman S. Campaigners accuse tobacco firm of dubious ploy. *BMJ.* 2000;320:1427.

38. Anti-discrimination measures. The Tobacco Institute. September 23, 1992. Bates No. TIMN0030964-TIMN0030965. Available at: http://tobaccodocuments.org/ti/TIMN0030964-0965.html. Accessed May 20, 2002.

39. Mixner D. Memo from DBM Associates [consultant to the Tobacco Institute] to Grant Gillham, Joe Shumate & Associates, Inc. Bates No. TCAL0404793. Available at: http://tobaccodocuments.org/landman/TCAL0404793.html. Accessed April 14, 2002

40. Letter to Joanne from Don Tuthill. August 17, 1992. Bates No. 2023439114. Available at: http://tobaccodocuments.org/pm/2023439114-9115.html. Accessed March 1 2002.

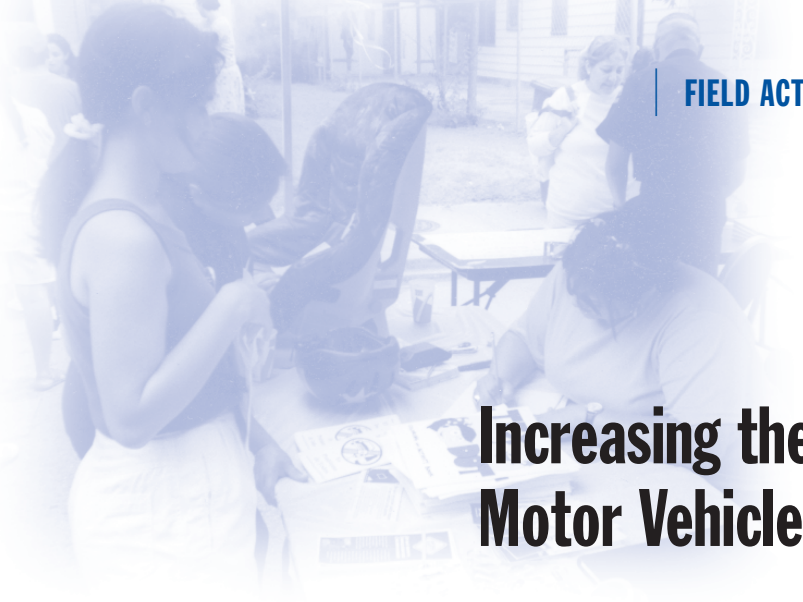
41. Pertman A. Gay market, ads target big dollars, not big change. *Boston Globe.* February 4, 2001:E1.

42. Marlboro Hispanic promotions 880000 marketing plan. 1988 (est). Bates No. 2048679289/294. Available at: http://www.tobaccodocuments.org/pm/2048679289-9294.html#images. Accessed April 21, 2002.

43. Hispanic market opportunity assessment. Progress update presentation. January 1985. Bates No. 504617452–504617507. Available at: http://tobaccodocuments.org/rjr/504617452-7507.html. Accessed April 21, 2002.

44. The Hispanic market. February 10, 1983. Bates No. 2042463401–2042463437. Available at: http://tobaccodocuments.org/pm/2042463401-3437.html. Accessed March 10, 2002.

45. Reemtsma West brand cigarettes "Queer" Web page. Available at: http://www.west.de/queer. Accessed May 15, 2002.



Increasing the Use of Child Restraints in Motor Vehicles in a Hispanic Neighborhood

Gregory R. Istre, MD, Mary A. McCoy, Katie N. Womack, MS, Linda Fanning, MS, Laurette Dekat, MD, MPH, and Martha Stowe, MSW

The use of restraints in motor vehicles is less common in minority and low-income populations than in the general population. A preliminary survey of Hispanic preschool-aged children in west Dallas, Tex, conducted in 1997 showed much lower child restraint use (19% of those surveyed) than among preschool children of all races in the rest of the city (62%).

Because there are few reports of successful programs to increase child restraint use among Hispanics, we undertook to implement and evaluate such a program. The program was conducted by bilingual staff and was tailored for this community. It was successful in increasing both child restraint use and driver seat belt use.

MINORITY AND LOW-INCOME populations are less likely than the general population to use restraints in motor vehicles.[1–4] A preliminary survey of Hispanic preschool-aged children in west Dallas, Tex, conducted in 1997 showed that the percentage using child restraints was much lower in this population (19%) than among preschool children of all races in the rest of the city (62%).[5]

Information obtained from focus groups of Hispanic parents in the community led to development of a program tailored to that population. Activities were carried out at neighborhood parties and in a local community health center, local day care centers, churches, community centers, and *botanicas* (stores that sell traditional Hispanic remedies and often employ a folk healer). Child safety seats were distributed through parent education classes in a variety of locations.

Child motor vehicle restraint use was evaluated through structured observational surveys, which showed a significant increase in child restraint use in the community. By 2000, restraint use among Hispanic preschool-aged children attending the clinic (72%) had surpassed use in a comparison population of preschool-aged children in the rest of Dallas (69%).

THE PROGRAM

Three adjacent zip codes (75208, 75211, and 75212) in the west sector of Dallas were chosen for the survey because of their predominantly Hispanic population (population 110 000, 60% Hispanic). Preliminary surveys had shown that child restraint use among Hispanic preschool-aged children was lower than 20% in several settings in these zip codes.

We used the Safe Communities model to develop community interventions.[6] Components of the program are listed in Table 1. The interventions were developed from standardized educational programs, with modifications based on information obtained from 6 focus groups in the Hispanic community. These included establishing a child safety seat loaner program, educating parents in small classes, identifying mothers as authority figures to help communicate the message, addressing the issue of fatalism or destiny, and using videos that graphically showed what happens to a child held on an adult's lap in a car crash.

Since child safety seat use was not part of the tradition or culture of the Hispanic community in the target area,[7,8] the interventions were incorporated into various aspects of the culture. For example, local priests were asked to bless the child safety seats in a ceremony before they were distributed; pamphlets about the program were distributed through local *botanicas*, churches, and community centers; and educational materials about child safety seats were presented on local Spanish-language radio and television shows.

Child safety seat classes were conducted in Spanish and English in the target area throughout

the project, beginning in May 1997. The classes were taught by certified child passenger safety technicians and were held biweekly at the only county-sponsored community primary care health center in the target area and at other locations in the community on request. Parents were required to attend an hour-long training class on the proper use and installation of child safety seats before they received a seat. They were asked to pay a $10 deposit for the seat, but they were not denied a seat if they were unable to pay. More than 3000 child safety seats were distributed to Hispanic families in the target area during the survey period.

Interventions also were implemented in several day care centers and neighborhoods in the target area (Table 1). A key component was the traffic safety workshops, which included information about vehicle safety, driver's licenses, immigration, and social security laws, as well as demonstrations of proper installation of child safety seats. At local schools, churches, and neighborhood events, a Hispanic policewoman known as *La Protectora* ("The Protector")[9] held classes in Spanish and English for parents and children to explain child safety laws and procedures. Information was also provided through activities at churches, community centers, and *botanicas.* Trained bilingual staff, most of whom were also residents of the target area, conducted all activities.

EVALUATION

Observers were trained in the use of a standardized observation survey form that had been used by the Texas Transportation Institute (TTI) for the past 13 years for longitudinal studies of restraint use throughout Texas.[10] Beginning in February 1997, surveys were conducted as vehicles entered parking lots at 3 types of locations in the target area: (1) the community health center where the intervention was done; (2) day care centers that were sites of interventions, and (3) the parking lots of 8 grocery stores,

KEY FINDINGS

- A program to increase use of child safety restraints in motor vehicles in a Hispanic neighborhood was successful because it incorporated religion, cultural beliefs, and community into the interventions, and because it was ongoing and multifaceted.
- The program was most successful among persons who attended the community health center and in the youngest age group (children younger than 2 years).
- Use of child safety seats and restraints was closely linked to drivers' use of seat belts.

TABLE 1—Components of Program to Increase Child Restraint Use in Hispanic Community, Dallas, Tex

Program development and cultural issues

- The program used the Safe Communities model. It was developed from standardized educational programs, adapted to the Hispanic community, and was further modified on the basis of results from 6 community focus groups.
- Activities were conducted in Spanish and in English.
- Classes were taught by certified child passenger safety technicians who were bilingual.
- Activities addressed issues of importance to the community other than child safety restraints.

Health center

- Classes were held biweekly.
- Parents were required to attend an hour-long training class on proper use and installation before receiving a child safety seat.
- Parents were not denied a seat if they were unable to pay the $10 deposit.
- Pediatricians at the health center promoted child safety seat use by distributing "prescriptions" for proper child safety seat use to patients.
- Class instructors also participated in health fairs and special events sponsored by the health center, distributed pamphlets about child safety seat and seat belt use, and conducted child safety seat inspections and demonstrations.

Day care centers

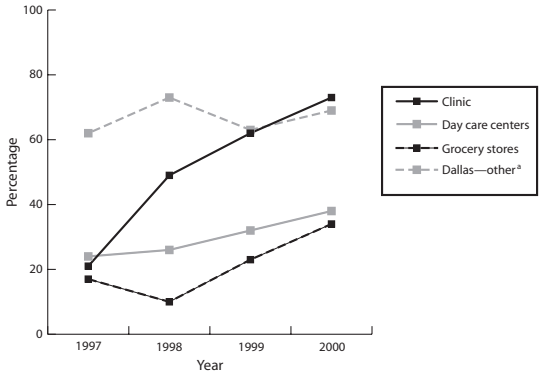
- A week-long intervention for children, parents, and day care staff included a presentation by a Hispanic policewoman, seat belt demonstrations, coloring contests, child safety seat training for day care center employees, and traffic safety workshops for parents. It emphasized the importance of seat belt use by parents as well as use of child safety seats.
- A second intervention involved hiring 3 local mothers as liaisons to promote child safety seat use over a 9-week period in 5 day care centers. Liaisons were responsible for developing an ongoing relationship with the day care centers and implementing interventions tailored to the centers' needs and interests. Strategies developed by the liaisons included information booths, raffles, and games designed to promote child safety seat use.

Neighborhoods

- Interventions implemented in several neighborhoods in the target area included neighborhood block parties, health fairs, child safety seat inspections, traffic safety workshops, and delivery of educational messages at local festivals.
- The traffic safety workshops included information about vehicle safety, driver's license and traffic laws, immigration and social security laws, and proper installation and use of child safety seats.
- At local schools, churches, and other neighborhood events, a Hispanic policewoman held classes in Spanish and English for parents and children to explain child safety laws and procedures.
- Educational pamphlets were distributed at churches, community centers, and local *botanicas*.

Evaluation

- Child restraint surveys were performed by trained observers, using survey forms developed by the Texas Transportation Institute.
- The target-area sites were a community health center, day care centers, and grocery store parking lots. The comparison-area sites were day care centers and shopping centers in other parts of the city.



Note. Data from the health center, day care centers, and grocery stores in the target area are for Hispanic children only.
[a]Dallas data were obtained from Texas Transportation Institute surveys for Dallas, Tex, excluding the target area, and include all races/ethnicities.

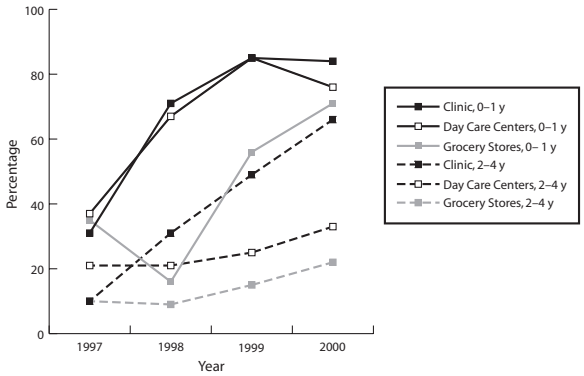
FIGURE 1—Child restraint use for children from birth to age 4 years in various settings in the target area, compared with Dallas data, Dallas, Tex, 1997–2000.

which were patronized predominantly by Hispanics. Observations at the grocery stores were considered to be most representative of the community as a whole. Children who were restrained in accordance with current Texas state law were considered properly restrained.[11] The safety seats were not examined in detail to determine whether they were appropriately tightened and tethered.

A total of 7413 observations among preschool-aged Hispanic children (<5 years) were conducted from 1997 through 2000: 2246 (30%) at the health center, 2735 (37%) at day care centers, and 2432 (33%) at grocery store parking lots. Additionally, 4137 comparison observations were done by TTI on preschool-aged children of all races in other parts of Dallas.

Child restraint use among preschool-aged Hispanic children increased significantly in all 3 settings between 1997 and 2000 ($P<.0001$ by χ^2 for trend; Figure 1). By 2000, use of restraints among Hispanic children attending the health center was higher than use in the rest of the city as measured in the TTI survey (72% vs 68%). There was substantially higher use among younger children (from birth to 1 year) than among children aged 2 to 4 years (Figure 2). Nevertheless, the trend of increasing use was significant in all 3 settings and for both age groups ($P<.001$ by χ^2 for trend).

Observed driver seat belt use also increased significantly in each of the 3 settings ($P<.001$ by χ^2 for trend), whereas the TTI survey showed little change in driver seat belt use for other parts of Dallas (not significant). There was a strong association between child restraint use and driver seat belt use at all the observation sites. The association remained strong after results were stratified by year, setting, age of the child, and type of vehicle (summary risk ratio=5.7, 95% confidence interval=5.0, 6.4; $P<.0001$).



Note. Data from the health center, day care centers, and grocery stores in the target area are for Hispanic children only.

FIGURE 2—Child restraint use for Hispanic children from birth to age 1 year and aged 2 to 4 years in various settings in the target area, Dallas, Tex, 1997–2000.

DISCUSSION

We believe that the success of this program was the result of its ongoing nature,[12] its integration of cultural and religious factors,[13,14] the use of Hispanic teachers in child safety seat classes, the efforts of the community health center staff to integrate safety messages into the clinical routine,[15] and the feedback that program staff received from surveillance data about progress in the program.[16] In addition, it appears clear that a successful program to increase child restraint use must target driver seat belt use.[17,18] We saw little increase in child restraint use in vehicles in which drivers did not wear a seat belt. Driver seat belt use may be a necessary factor in child restraint use, although it is not the sole determinant.

Several caveats are in order. First, although we found a significant increase in child restraint use in the community (as measured by the grocery store and day care center surveys), use in the community remained significantly lower than use at the health center, indicating that there is much work still to be done. Some of the increase in restraint use at the health center may have been due to a "social desirability" effect—parents who knew they were going to the health center, the primary site of the intervention activities, may have been more likely to practice car seat safety.

Second, although the overall trend was relatively flat, there were some fluctuations in restraint use in the rest of Dallas from year to year, which may have reflected other community factors at work that influenced child restraint use. Third, the predominant impact of the program in the overall community (as measured by the surveys in grocery store parking lots) was seen among children younger than 2 years. Finally, the program may not be generalizable to other populations and ethnic groups.

Despite these possible shortcomings, the program appears to have been successful. With a multifaceted program, child restraint use and driver seat belt use in the Hispanic population may reach levels that equal or exceed those of the general population. ■

About the Authors

Gregory R. Istre, Mary A. McCoy, Linda Fanning, and Martha Stowe are with Injury Prevention of Greater Dallas, Dallas,

Tex. Gregory R. Istre is also with PID Associates, Dallas, Tex. Katie N. Womack is with the Texas Transportation Institute, College Station. Laurette Dekat is with the Department of Pediatrics and the Department of Community and Family Medicine, University of Texas Southwestern Medical Center, Dallas.

Requests for reprints should be sent to Gregory R. Istre, MD, Injury Prevention of Greater Dallas, PO Box 36067, Dallas TX 75235.

This report was accepted February 18, 2002.

Contributors

G. Istre participated in conception and design of the intervention and evaluation, analyzed and interpreted the data, and wrote the manuscript. M. McCoy participated in conception and design of the intervention and evaluation, analyzed data, and assisted in interpreting the analysis. K. Womack participated in conception and design of the intervention and evaluation, helped with data analysis, and assisted in interpreting the analysis. L. Fanning participated in design of the intervention and evaluation. L. Dekat participated in the design and implementation of the interventions. M. Stowe participated in the conception and design of the intervention and evaluation, and assisted in interpreting the analysis. All authors helped revise the manuscript.

Acknowledgments

This work was supported in part by cooperative agreement DTNH22-96-H-15166 from the National Highway Traffic Safety Administration and the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

We are indebted to Salvador Manzo and Isabel Colunga, who performed observational surveys; Dawn Cano and Martha Maldonado for their work at the health center and in the community; Allen Bolton for his work in guiding the early stages of the project; the staff of the Texas Transportation Institute for data from the annual child restraint surveys for Texas; physicians, nurses, and staff of the deHaro Saldivar Health Center for their dedication to these ongoing safety interventions (in particular, Esperanza Martinez and Antonia Garza, who maintained the sign-up lists for classes); the directors and staff of the day care centers and the managers of the grocery stores in the study; Mike West and the Texas Dept of Transportation and the members of the Safe Riders program of the Texas Department of Health for their roles in acquiring child safety seats for the program; Paul Boumbulian and Ron Anderson for their support of the many prevention efforts of the IPC; the Dallas Police Department and Debra Rafferty for implementing the La Protectora program in the Hispanic community; the staff and members of Avance Dallas and Wesley-Rankin Community Center for their partnership in reaching the Hispanic community; and Kristen Howell for reviewing the manuscript.

References

1. Louis B, Lewis M. Increasing car seat use for toddlers from inner-city families. *Am J Public Health.* 1997;87:1044–1045.

2. Hanfling MJ, Mangus LG, Gill AC, Bailey R. A multifaceted approach to improving motor vehicle restraint compliance. *Injury Prev.* 2000;6:125–129.

3. Harper JS, Marine WM, Garrett CJ, Lezotte D, Lowenstein SR. Motor vehicle crash fatalities: a comparison of Hispanic and non-Hispanic motorists in Colorado. *Ann Emerg Med.* 2000;36:589–596.

4. Lee P, Orsay E, Lumpkin J, Ramakrishman V, Callahan E. Analysis of Hispanic motor vehicle trauma victims in Illinois, 1991–1992. *Acad Emerg Med.* 1996;3:221–227.

5. Womack KN. *2000 Survey of Front Seat Occupant Restraint Use in Eighteen Texas Cities.* College Station: Texas Transportation Institute; January 2001.

6. Svanstrom L, Ekman R, Schelp L, Lindstrom A. The Lidkoping Accident Prevention Programme—a community approach to preventing childhood injuries in Sweden. *Injury Prev.* 1995; 1: 169–172.

7. Stiles MC, Grieshop JI. Impacts of culture on driver knowledge and safety device usage among Hispanic farm workers. *Accid Anal Prev.* 1999; 31: 235–241.

8. Shin D, Hong L, Waldron I. Possible causes of socioeconomic and ethnic differences in seat belt use among high school students. *Accid Anal Prev.* 1999;31:485–496.

9. *How to Implement an El Protector Traffic Safety Program.* Washington, DC: National Highway Traffic Safety Administration; 1994. Publication DOT-HS-808-110.

10. Kiegle CL, Waller AE, Pendleton OJ, Bremer R. *A Statewide Sampling Survey of Child Restraint Usage in Texas.* College Station: Texas Transportation Institute; April 1985.

11. *Traffic Safety Facts 1998.* Washington, DC: National Highway Traffic Safety Administration, National Center for Statistics and Analysis; 1999. Publication DOT-HS-808-983.

12. Grossman DC, Garcia CC. Effectiveness of health promotion programs to increase motor vehicle occupant restraint use among young children. *Am J Prev Med.* 1999;16(suppl 1):12–22.

13. Agran PF, Winn DG, Anderson CL, Del Valle C. Family, social, and cultural factors in pedestrian injuries among Hispanic children. *Injury Prev.* 1998;4:188–193.

14. Klassen TP, MacKay JM, Moher D, Walker A, Jones AL. Community-based injury prevention interventions. *Future Child.* 2000;10:83–110.

15. DiGuiseppi C, Roberts IG. Individual-level injury prevention strategies in the clinical setting. *Future Child.* 2000;10:53–82.

16. Dowswell T, Towner EM, Simpson G, Jarvis SN. Preventing childhood unintentional injuries—what works? A literature review. *Injury Prev.* 1996; 2: 140–149.

17. Agran PF, Anderson CL, Winn DG. Factors associated with restraint use of children in fatal crashes. *Pediatrics.* 1998;102:e39.

18. Miller TR, Spicer RS, Lestina DC. Who is driving when unrestrained children and teenagers are hurt? *Accid Anal Prev.* 1998;30:839–849.

A priest reads a prayer before the blessing ceremony. Parents are standing next to their vehicle, and the child safety seats are in place.

Zip Code Caveat: Bias Due to Spatiotemporal Mismatches Between Zip Codes and US Census–Defined Geographic Areas—The Public Health Disparities Geocoding Project

Nancy Krieger, PhD, Pamela Waterman, MPH, Jarvis T. Chen, ScD, Mah-Jabeen Soobader, PhD, S. V. Subramanian, PhD, and Rosa Carson, BA

Use of zip codes in US public health research is on the rise. As of February 2002, 230 articles were indexed by zip code in PubMed,[1] all published since 1989. Fifty-two of these articles (23%) involved the use of census-derived zip code socioeconomic data (e.g., median household income) to investigate the effects of socioeconomic position on specified health outcomes (article citations are available on request from the authors).

To date, discussions regarding the use of zip code socioeconomic data for US public health research have focused chiefly on whether zip codes' larger population size (average: 30 000) and potentially greater socioeconomic heterogeneity would attenuate estimates of socioeconomic gradients in health detected using zip codes in comparison with estimates obtained via census tract (average population: 4000) or block group (average population: 1000) socioeconomic data.[2–7] Unacknowledged in the public health literature, however, is the fact that zip codes differ from census tracts and block groups in other important ways, including spatiotemporal definition and stability.

Unlike census tracts, defined by the US Bureau of the Census as "small, relatively permanent statistical subdivision[s] of a county . . . designed to be relatively homogeneous with respect to population characteristics, economic status, and living conditions,"[8(ppG-10–G-11)] zip codes are "administrative units established by the United States Postal Service . . . for the most efficient delivery of mail, and therefore generally do not respect political or census statistical area boundaries."[9(pA-13)] Spanning in size from a single building or company with a high volume of mail to large areas that cut across states, "carrier routes for one zip code may intertwine with those of one or more zip codes" such that "this area is more conceptual than geographic."[10(p22)]

To "overcome the difficulties in precisely defining the land area covered by each zip code,"[11] the US Census Bureau created a new statistical entity built from census blocks: the 5-digit zip code tabulation area (ZCTA), first used in the 2000 census.[12] Of note, ZCTAs and zip codes sharing the same 5-digit code may not necessarily cover the same area (Table 1),[13] so that zip codes obtained via self-report or from addresses in medical records cannot be assumed to correspond to census-defined ZCTAs.

Even before introduction of the ZCTAs, there were 2 types of spatiotemporal discontinuity that could conceivably affect health studies linking zip codes to census-derived data: (1) changes in zip code delivery routes—and hence in population covered by the affected zip code—and (2) discontinuation and addition of zip codes in nondecennial years.[14–16] Between 1997 and 2001 alone, the US Post Office added approximately 390 new zip codes nationwide and discontinued 120 (oral communication, Meg Ausman, US Post Office Data Center, February 5, 2002). One implication of these changes is that persons could be correctly geocoded to a zip code that did not exist in the preceding decennial census.

Findings from the Public Health Disparities Geocoding Project[17] illustrate the potential problems for health research of spatiotemporal zip code–census mismatches, even those dating from before the creation of ZCTAs. This project was designed to assess which area-based socioeconomic measures at which levels of geography (census tract, block group, and zip code) are most appropriate for monitoring socioeconomic inequalities in health. Health data from 2 states (Massachusetts and Rhode Island) and the 1990 census were used. Records were geocoded in 1999 by a firm whose accuracy we ascertained to be high (96%),[18] and the firm, following standard practice, returned the most recent geocodes available.

TABLE 1—Technical Definitions of and Distinctions Between Zip Codes and Zip Code Tabulation Areas (ZCTAs)

Definition of ZCTAs[11]

"ZIP Code Tabulation Areas (ZCTAs™) are a new statistical entity developed by the US Census Bureau for tabulating summary statistics from Census 2000. This new entity was developed to overcome the difficulties in precisely defining the land area covered by each ZIP Code®. Defining the extent of an area is necessary to accurately tabulate census data for that area. ZCTAs are generalized area representations of US Postal Service (USPS) ZIP Code service areas. Simply put, each one is built by aggregating the Census 2000 blocks, whose addresses use a given ZIP Code, into a ZCTA which gets that ZIP Code assigned as its ZCTA code. They represent the majority USPS five-digit ZIP Code found in a given area. For those areas where it is difficult to determine the prevailing five-digit ZIP Code, the higher-level three-digit ZIP Code is used for the ZCTA code. Since we take the ZIP Code used by the majority of addresses in an area for the ZCTA code, some addresses will end up with a ZCTA code different from their ZIP Code. Also, some ZIP Codes represent very few addresses (sometimes only one) and therefore will not appear in the ZCTA universe."

Distinction between ZCTAs and Zip Codes[12]

"Even though the codes may appear the same, the addresses and areas covered by these areas may not be the same. We strongly advise data users who wish to compare 1990 and 2000 data to determine and evaluate any coverage differences that exist before making any comparisons. There are several reasons for this caution: The USPS has extensively modified ZIP Codes over the last ten years. Even though a 1990 ZIP Code matches a Census 2000 ZCTA code, there is no guarantee that these cover the same geographic area. Also, some ZIP Codes in the 1990 data products were discontinued by the USPS, and new ZIP Codes were created; ZCTAs and the 1990 data products were discontinued by the USPS, and new ZIP Codes were created; ZCTAs and the 1990 census ZIP Code areas were delineated using different methodologies and therefore may not have comparable coverage area or size; and the Census 2000 ZCTAs will include some dedicated PO box ZIP Codes. All dedicated PO box ZIP Codes were excluded as ZIP Code areas in 1990. The resulting 1990 areas include data for both PO box ZIP Codes and the ZIP Codes that provide street or rural route delivery to the surrounding area."

TABLE 2—Incident Colon Cancer Counts by Geographic Level: Massachusetts, 1987–1993

	Geocoded to			
			Zip Code	
No. of Cases of Colon Cancer	Block Group No. (%)	Census Tract, No. (%)	Total, No. (%)	Zip Code Changed or Established After the 1990 Census, No. (%)
17 266	15 792 (91.5)	17 265 (100.0)	17 266 (100.0)	1784 (10.3)

Cancer incidence rates were one of the health outcomes addressed. We found that in Massachusetts (474 zip codes listed in the 1990 census), 17 376 (10.4%) of the 166 730 cancer cases occurring during 1987 to 1993 were geocoded to 193 zip codes not included in the 1990 census; 15 774 (90.8%) of these 17 376 cases were in one of 30 zip codes changed or established after the 1990 census.[19–21] By contrast, in Rhode Island (70 zip codes listed in the 1990 census), only 0.7% (148) of the 19 766 geocoded cancer incidence records were matched to zip codes not included in the 1990 census.

In the case of colon cancer incidence in Massachusetts, moreover, the impact of excluding persons linked to zip codes not included in the 1990 census was substantial. Zip code–level analyses yielded socioeconomic gradients contrary to those observed via data at the tract and block group levels and contrary to those reported in the literature (Tables 2–4).[22]

Given the growing interest in linking geographic and health data,[23,24] we urge researchers, when using geocoded records, to pay careful attention to the potential for spatiotemporal mismatches between census-derived and zip code data as well as to changes in zip code boundaries and years in which boundaries were established. Public health projects and programs that use zip code data should likewise be alert to potential new issues stemming from the replacement of zip codes with ZCTAs in the 2000 census. ■

About the Authors

The authors are with the Department of Health and Social Behavior, Harvard School of Public Health, Boston, Mass.

Requests for reprints should be sent to Nancy Krieger, PhD, Department of Health and Social Behavior, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115 (e-mail: nkrieger@hsph.harvard.edu).

This brief was accepted March 15, 2002.

Contributors

N. Krieger conceived and designed the study, directed data analysis, interpreted the data, and wrote the article. P. Waterman, J.T. Chen, M. Soobader, and S.V. Subramanian contributed to the conception and design of the study and analyzed and assisted with interpretation of the data. R. Carson assisted with presentation of the study results.

Acknowledgments

This work was funded by the National Institute of Child Health and Human Development, National Institutes of Health (grant 1 R01 HD36865-01).

We thank Dr Daniel Friedman (assistant commissioner, Bureau of Health Statistics, Research and Evaluation, Massachusetts Department of Health) and Dr Jay Buechner (chief, Office of Health Statistics, Rhode Island Department of Health) for facilitating the Public Health Disparities Geocoding Project by providing data from their respective health departments and for their helpful comments on the article. We likewise thank the following individuals for helping us to access the mortality and cancer registry data: Alice Mroszczyk, Dr Susan Gershman, Mary Mroszczyk, and Ann R. MacMillan of the Massachusetts Department of Public Health and Dr John Fulton of the Rhode Island Department of Health.

References

1. National Library of Medicine. PubMed. Available at: http://www.ncbi.nlm.nih.gov/PubMed/. Accessed February 5, 2002.

2. Krieger N, Williams D, Moss N. Measuring social class in US public health research: concepts, methodologies and guidelines. *Annu Rev Public Health.* 1997;18:341–378.

3. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health.* 1992;82:703–710.

4. Geronimus AT, Bound J. Use of census-based aggregate variables to proxy for socioeconomic group: evidence from national samples. *Am J Epidemiol.* 1998;48:475–486.

5. Diez Roux AV, Kiefe CI, Jacobs DR Jr, et al. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Ann Epidemiol.* 2001;11:395–405.

6. Fiscella K, Franks P. Impact of patient socioeconomic status on physician profiles: a comparison of census-derived and individual measures. *Med Care.* 2001;39:8–14.

TABLE 3—Colon Cancer Incidence Rates, Stratified by Area-Based Socioeconomic Measures, Among Persons in Areas With the Least and Most Resources, Along With Age-Adjusted Comparisons (Incidence Rate Ratio and Relative Index of Inequality): Massachusetts, 1987–1993

Selected Area-Based Socioeconomic Measure[a]	Rate: Least Resources[b]			Rate: Most Resources[b]			Incidence Rate Ratio (95% Confidence Interval): Least/Most			Relative Index of Inequality (95% Confidence Interval)		
	BG	CT	ZC	BG	CT	ZC	BG	CT	ZC	BG	CT	ZC
Working class (categorical)	41.3	42.5	41.1	45.8	48.3	27.9	0.90 (0.76, 1.06)	0.88 (0.73, 1.06)	1.47 (1.14, 1.90)	0.89 (0.84, 0.95)	0.85 (0.80, 0.90)	1.28 (1.20, 1.36)
Median household income (quintile)	41.0	42.5	42.3	46.3	48.9	37.2	0.89 (0.75, 1.04)	0.87 (0.74, 1.03)	1.14 (0.97, 1.34)	0.87 (0.82, 0.93)	0.88 (0.83, 0.93)	1.19 (1.12, 1.27)
Poverty (categorical)	41.7	45.6	44.8	43.9	47.4	41.6	0.95 (0.80, 1.13)	0.96 (0.81, 1.15)	1.08 (0.88, 1.32)	0.94 (0.88, 1.00)	0.95 (0.89, 1.01)	1.06 (0.99, 1.13)
Low education (categorical)	39.5	40.8	43.8	45.2	48.0	39.3	0.87 (0.73, 1.05)	0.85 (0.70, 1.03)	1.11 (0.90, 1.38)	0.84 (0.79, 0.90)	0.90 (0.85, 0.96)	1.15 (1.08, 1.22)
Index of local economic resources (quintile)	40.3	42.6	43.1	45.4	48.7	33.6	0.89 (0.76, 1.04)	0.87 (0.74, 1.03)	1.28 (1.09, 1.50)	0.86 (0.81, 0.91)	0.88 (0.83, 0.94)	1.27 (1.19, 1.35)

Note. The relative index of inequality is a measure of effect that takes into account both the population distribution of the exposure and the magnitude of the rate ratio detected in each socioeconomic stratum, thereby permitting meaningful comparison of gradients across different socioeconomic measures.[25–27] BG = block group; CT = census tract; ZC = zip code.

[a]The area-based socioeconomic measures and their cutpoints for these analyses are defined in Table 4.[17]

[b]Average annual rate (per 100 000) age standardized to the year 2000 standard million.[28]

TABLE 4—Area-Based Socioeconomic Measures and Cutpoints Used in Data Analysis

Selected Area-Based Socioeconomic Measure	Operational Definition and Cut Points Used
Working class[2] (categorical)	Percentage of persons employed in predominantly working class occupations (i.e., as nonsupervisory employees), operationalized as percentage of persons employed in the following 8 of 13 census-based occupational groups: administrative support; sales; private household service; other service (except protective); precision production, craft, repair; machine operators, assemblers, inspectors; transportation and material moving; handlers, equipment cleaners, laborers; cutpoints: C1 = 0%–49.9%, C2 = 50%–69.9%, C3 = 66%–74.9%, C4 = 75%–100%
Median household income (quintile)	Median household income in year before the decennial census (US in 1989: $30 056); cutpoints:
	Massachusetts BG: Q1 = $4999–$26 110, Q2 = $26 111–$33 749, Q3 = $33 750–$40 798, Q4 = $40 799–$49 903, Q5 = $49 904–$150 001
	Massachusetts CT: Q1 = $4999–$26 471, Q2 = $26 472–$33 162, Q3 = $33 163–$39 286, Q4 = $39 287–$47 124, Q5 = $47 125–$102 797
	Massachusetts ZC: Q1 = $9726–$30 624, Q2 = $30 625–$36 246, Q3 = $36 247–$41 396, Q4 = $41 397–$48 841, Q5 = $48 842–$94 898
Poverty (categorical)	Percentage of persons below federally defined poverty line, a threshold that varies by size and age composition of the household and, on average, equaled $12 647 for a family of 4 in 1989[9]; cutpoints: C1 = 0%–4.9%, C2 = 5.0%–9.9%, C3 = 10.0%–19.9%, C4 = 20%–100%; areas with a poverty rate of ≥ 20% are federally defined poverty areas[2]
Low education (categorical)	Percentage of persons 25 years and older with less than a 12th grade education; cutpoints: C1 = 0%–14.9%, C2 = 15.0%–24.9%, C3 = 25.0%–39.9%, C4 = 40%–100%
Index of local economic resources[29] (quintile)	A "summary index" used by the Centers for Disease Control and Prevention and based on "white collar employment, unemployment, and family income"; cutpoints:
	Massachusetts BG: Q1 = 0–6, Q2 = 7–11, Q3 = 12–15, Q4 = 16–20, Q5 = 21–27
	Massachusetts CT: Q1 = 0–5, Q2 = 6–10, Q3 = 11–15, Q4 = 16–19, Q5 = 20–26
	Massachusetts ZC: Q1 = 0–8, Q2 = 9–12, Q3 = 13–15, Q4 = 16–19, Q5 = 20–26

Note. C = category; BG = block group; Q = quintile; CT = census tract; ZC = zip code.

7. Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health.* 2001;91:1783–1789.

8. US Bureau of the Census. Geographical areas reference manual. Available at: http://www.census.gov/geo/www/garm.html. Accessed February 12, 2002.

9. *Census of Population and Housing, 1990: Summary Tape File 3 Technical Documentation.* Washington, DC: US Bureau of the Census; 1991.

10. US Bureau of the Census. Geographics changes for census 2000 + glossary. Available at: http://www.census.gov/geo/www/tiger/glossary.html. Accessed July 3, 2001.

11. US Bureau of the Census. Census 2000 zip code tabulation areas (ZCTAs). Available at: http://www.census.gov/geo/ZCTA/zcta.html. Accessed February 12, 2002.

12. US Bureau of the Census. Zip code tabulation area (ZCTA) frequently asked questions. Available at: http://www.census.gov/geo/ZCTA/zctafaq.html. Accessed February 12, 2002.

13. US Bureau of the Census. Census 2000 zip code tabulation areas technical documentation. Available at: http://www.census.gov/geo/ZCTA/zcta_tech_doc.pdf. Accessed February 12, 2002.

14. *Address Information Products Technical Guide.* Memphis, Tenn: US Postal Service National Customer Support Center; 2001.

15. Office of Social and Economic Data Analysis. The zip code resource page: tools and resources related to US zip codes. Available at: http://www.oseda.missouri.edu/uic/zip.resources.html. Accessed: February 24, 2002.

16. Range L, Clay T, Oliva G. *California Child and Youth Injury Hot Spot Project 1995–1997, Volume 3: Technical Guide.* Sacramento, Calif: California Dept of Health Services, Maternal and Child Health Branch; 2000:59–67.

17. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R. Geocoding and monitoring US socioeconomic inequalities in mortality and cancer incidence: does choice of area-based measure and geographic level matter?—the Public Health Disparities Geocoding Project. *Am J Epidemiol.* In press.

18. Krieger N, Waterman P, Lemieux K, Zierler S, Hogan JW. On the wrong side of the tracts? Evaluating accuracy of geocoding for public health research. *Am J Public Health.* 2001;91:1114–1116.

19. *National Five-Digit Zip Code and Post Office Directory.* Memphis, Tenn: US Postal Service; 1993:8-3. Publication 65.

20. *National Five-Digit Zip Code and Post Office Directory.* Memphis, Tenn: US Postal Service; 2000:3-1328–3-1371. Publication 65.

21. *Zip Code Changes for the Boston Area Effective July 1, 1998.* Dorchester, Mass: US Post Office; 1998.

22. Kogevinas M, Pearce N, Susser M, Bofetta P, eds. *Social Inequalities in Cancer.* Lyon, France: International Agency for Research on Cancer; 1997. IARC scientific publication 136.

23. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. Geographic information systems and public health: mapping the future. *Public Health Rep.* 1999; 114:359–373.

24. Moore DA, Carpenter TE. Spatial analytic methods and geographic information systems: use in health research and epidemiology. *Epidemiol Rev.* 1999;21: 143–161.

25. Pamuk ER. Social class inequality in mortality from 1921 to 1972 in England and Wales. *Popul Stud.* 1985;39:17–31.

26. Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. *Soc Sci Med.* 1991; 33:545–557.

27. Davey Smith G, Hart C, Hole D, et al. Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Community Health.* 1998;52:153–160.

28. Anderson RN, Rosenberg HM. Age standardization of death rates: implementation of the year 2000 standard. *Natl Vital Stat Rep.* October 7, 1998;47(3).

29. Casper ML, Barnett E, Halverson JA, et al. *Women and Heart Disease: An Atlas of Racial and Ethnic Disparities in Mortality.* Morgantown, WVa: Office for Social Environment and Health Research, West Virginia University; 1999.

Rural–Urban Differences in the Public Health Workforce: Local Health Departments in 3 Rural Western States

Roger A. Rosenblatt, MD, MPH, Susan Casey, PhD, and Mary Richardson, PhD, MHA

Most local health departments or districts are small and rural; two thirds of the nation's 2832 local health departments serve populations smaller than 50 000 people.[1] Rural local health departments have small staffs and slender budgets, yet they are expected to provide a wide array of services[2] during a period when

the health care system of which they are a part is undergoing change.[3]

This study provided quantitative, population-based data on the supply and composition of the rural public health workforce in 3 extremely rural states: Alaska, Montana, and Wyoming. The study focused on the relative supply of personnel in the principal public health occupational categories, differences across states in staffing levels, and difficulties experienced in recruiting and retaining personnel.

METHODS

We identified all local health departments in the 3 states with assistance from the state health departments—52 in Montana and 23 in Wyoming. In Alaska, each of the 22 local offices of the state public health department was treated as a separate local health department.

The survey instrument was based on work performed by the American Public Health Association,[4,5] as modified by the Center for Health Policy Study of the University of Texas.[6] The survey was mailed directly to the administrator of every local health department in 1999 and 2000. We used follow-up contacts until every local health department had responded, for a 100% response.

We defined a local health department as rural if it was within a county with fewer than 50 000 people. In Alaska, which does not have county governments, we designated the Anchorage and Fairbanks local health departments as urban.

In every state, some local services are also provided by state or regional public health personnel. We specifically excluded those personnel from the calculations that follow. We also excluded environmental health personnel from the analyses that follow.

RESULTS

The 3 study states had 99 local health departments, serving a population of almost 2 000 000, about half of which lives in rural areas. The average local health department had fewer than 10 in-house employees.

The supply of professional public health personnel, excluding environmental health workers, was virtually identical across states on a per capita basis. Despite different organizational formats across states, local health departments had approximately 31 full-time equivalents for every 100 000 residents, or approximately 1 local health department professional for every 3225 residents. This remarkable uniformity in workforce supply represents the product of convergent evolution, because no joint planning is done across any of these states' boundaries.

Alaska and Wyoming actually had a greater relative supply of public health professionals in the rural compared with the urban areas, as can be seen in Table 1. Rural local health departments tend to have fewer support staff. The major rural–urban differences lie in the size of the populations served and the number of people who work in the local health departments. Rural local health departments in these 3 states are very small, serving on average slightly more than 10 000 people, with about 5 working public health professionals in a typical office. Virtually all local health department professional personnel in rural Alaska are full-time; by contrast, most rural public health personnel in Montana work part-time, with Wyoming between these 2 extremes.

As shown in Table 2, the core of all the local health departments—urban or rural—is the public health nurse. These nurses constitute the majority of the professional workforce in the rural local health departments, largely because few other professionals work in these remote settings. Urban areas also rely heavily on nurses, but other types of professional personnel—from epidemiologists to nutritionists—play a greater role. Managers are also slightly more abundant in rural areas, primarily because Alaska uses more management personnel in these settings. By contrast, every other occupational category within public health is more plentiful in urban as opposed to rural local health departments.

Both state-to-state and rural–urban differences are seen in the extent to which individual local health departments are successful in recruiting professionals with a public health background. Alaska, with its predominantly full-time staff, has a highly professionalized workforce. Montana, with its predominantly

TABLE 1—Rural–Urban Differences in Local Health Department (LHD) Staffing: Alaska, Montana, and Wyoming, 1999–2000

	Alaska		Montana		Wyoming		Total	
	Rural (n = 22)	Urban (n = 2)	Rural (n = 46)	Urban (n = 6)	Rural (n = 21)	Urban (n = 2)	Rural (n = 89)	Urban (n = 10)
Demographics								
Population served	279 106	342 710	369 155	446 140	326 703	138 033	974 965	926 882
Average district population	12 687	171 355	8025	74 357	15 447	69 016	10 483	92 688
Staffing (excludes sanitarians)								
Total professional staff	121	104	213	175	150	40	484	319
Total professional FTEs	116.6	96.7	97.8	139.5	99.9	34.9	314.3	271.1
Professional FTEs/100 000	41.8	28.2	26.5	31.3	30.6	25.3	32.2	29.2
Total support FTEs	63.8	82.0	34.3	37.7	45.5	17.8	143.6	137.5
Support FTEs/100 000	22.9	23.9	9.3	8.5	13.9	12.9	14.7	14.8
Mean total FTEs/100 000	64.6	52.1	35.8	39.7	44.5	38.2	47.0	44.1

Note. FTE = full-time equivalent.

TABLE 2—Composition and Per Capita Supply of Professional Public Health Workforce in Local Health Departments (LHDs): Alaska, Montana, and Wyoming, 1999–2000

	Mean FTEs per 100 000 Population								Percentage of LHDs With Staff in This Category	
	Alaska		Montana		Wyoming		Total			
Professional Categories	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural LHDs	Urban LHDs
Management	5.0	1.8	5.7	6.1	6.3	6.1	5.7	4.5	59	100
Clinical personnel										
Physicians	0.0	0.1	0.4	0.1	0.3	0.3	0.2	0.1	62	80
Nurses	28.5	11.8	15.4	10.3	20.6	12.6	20.8	11.2	96	100
Physician assistants	0.0	0.0	0.03	0.0	0.0	0.7	0.1	0.1	3	10
Nurse practitioners	3.2	1.5	0.9	0.4	1.2	2.5	1.7	1.1	23	60
Nutritionists	0.4	2.6	0.4	2.2	0.3	0.6	0.4	2.1	11	70
Health educators	0.0	0.9	0.7	2.8	0.0	0.7	0.3	1.8	3	70
Social workers	0.0	0.0	1.0	2.2	0.2	0.0	0.5	1.1	8	50
Other direct care providers	4.3	0.9	1.9	3.4	1.3	1.1	2.4	2.1	26	80
Subtotal	36.4	17.7	20.7	21.4	24.0	18.5	26.3	19.6	NA	NA
Other public health[a]	0.4	8.8	0.0	3.9	0.3	0.7	0.2	5.2	1	50
Total	41.8	28.2	26.4	31.4	30.6	25.3	32.2	29.3	NA	NA

Note. FTE = full-time equivalent; NA = not applicable.
[a]Includes epidemiologists, disease investigators, laboratory scientists, communications specialists, planners, and others.

part-time workforce, recruits public health professionals from other delivery settings, often individuals without previous public health experience or training. Wyoming again falls somewhere in between.

In the 3 states we studied, rural local health departments had relatively few vacancies. Whereas 70% of the urban local health departments were recruiting for public health nurses, only 21% of the rural local health departments had a similar vacancy. Nurse practitioners were the most difficult professionals to recruit. Rural–urban differences showed no clear pattern. Where recruitment was difficult, low salaries, difficulty finding qualified local professionals, and problems attracting personnel were reported to be common.

DISCUSSION

The Rural Local Health District

This study found that the core of the rural public health system is the public health nurse. There is approximately 1 full-time equivalent public health nurse for every 6000 people. In many cases, these nurses learn on the job. Many have no specific public health training and no experience in public health, and many of them work part-time.

Rural–Urban Differences

This study showed that the per capita supply of public health personnel was similar in the rural and urban places we studied. The differences were more subtle. Rural public health personnel were less likely to have formal public health training and experience and more likely to be employed part-time. Perhaps more important, rural public health personnel had a much smaller team of people with whom to interact and a much narrower range of public health skills represented in the local office.

Personnel shortages are relatively infrequent, even in the most remote rural areas. Many of the rural public health workers have been in these positions for long periods. The challenges of continuing education and further training can be immense, but rural public health workers tend to stay in their local communities.

Formal input to the rural local health department team from physicians and dentists is virtually nonexistent. Most rural local health departments have a volunteer physician who can sign death certificates or attend an occasional meeting. Our results conform almost exactly with those of the 1 other comprehensive national study that examined small local health departments.[7] These authors concluded that there is a weakness in the "front lines" of public health; our results would certainly support that conclusion.

The rural public health system is small and isolated, but so are many other public functions located in rural communities. For these professionals to be effective—and to survive their often-stressful jobs—they must be connected with other professionals at the local, regional, and state levels. Our impression is that where the state plays a large role in organizing and running the system, local public health workers feel much more to be a part of something larger than themselves. Where state involvement is less pervasive, local health department staff feel much more uncertain and alone. ■

About the Authors

The authors are with the University of Washington, Seattle.

Requests for reprints should be sent to Roger A. Rosenblatt, MD, MPH, Department of Family Medicine, University of Washington, Box 354696, Seattle, WA 98195-4696 (e-mail: rosenb@u.washington.edu).

This brief was accepted November 6, 2001.

Contributors

R.A. Rosenblatt conceived the project with M. Richardson. R.A. Rosenblatt designed the study, guided the survey development and methodology, and wrote the final draft. S. Casey conducted the mail survey, handled follow-up, and analyzed the data. M. Richardson did much of the original contact with state and local public health officials and helped plan the research and survey.

Acknowledgments

This study was performed by the WWAMI (Washington, Wyoming, Alaska, Montana, Idaho) Center for Health Workforce Studies, which is supported by the National Center for Health Workforce Information and Analysis, Bureau of Health Professions, Health Resources and Services Administration.

References

1. Mays GP, Miller CA, Halverson PK. *Local Public Health Practice: Trends and Models.* Washington, DC: American Public Health Association; 2000.

2. Turnock BJ. *Public Health: What It Is and How It Works.* Gaithersburg, Md: Aspen Publishers; 1997.

3. Slifkin RT, Silberman P, Reif S. *The Effect of Market Reform on Rural Public Health Departments.* Chapel Hill: North Carolina Rural Health Research Program, Cecil G. Sheps Center for Health Services Research, University of North Carolina; 2000. Working Paper 65.

4. de la Puente JL. *Determination of Supply of and Requirements for Public/Community Health Personnel.* Final Report. Washington, DC: American Public Health Association; 1984. Contract No. HRA-240-83-0078.

5. de la Puente JL. *Survey of Public Health/Community Health Personnel.* Final Report. Washington, DC: American Public Health Association; 1983. Contract No. HRA-232-81-0056.

6. Kennedy VC, Spears WD, Loe HD Jr, Moore FI. Public health workforce information: a state-level study. *J Public Health Manage Pract.* 1999;5(3):10–19.

7. Gerzoff RB, Brown CK, Baker EL. Full-time employees of US local health departments, 1992–1993. *J Public Health Manage Pract.* 1999;5(3):1–9.

Determinants of Maternal Vitamin A or Beta-Carotene Supplementation Coverage: Village-Based Female Distributors in Nepal

Joanne Katz, ScD, Keith P. West Jr, DrPH, Lee Wu, MSc, Subarna K. Khatry, FRCS, Elizabeth Kimbrough Pradhan, MPH, Parul Christian, DrPH, Steven C. LeClerq, MPH, and Sharada Ram Shrestha, MPH

Interest in community-based micronutrient supplementation programs is increasing as the health impacts of these programs continue to be documented in randomized trials.[1–14] However, if such programs are to have the desired effects, high coverage levels must be sustained over time. Success rates may be enhanced if local residents take primary responsibility for distribution of supplements. Few data are available to provide guidance about characteristics of community-based distributors that best predict coverage. We attempted to identify such characteristics among village-based women hired to deliver nutritional supplements to women of childbearing age in rural Nepal.

METHODS

Between 1994 and 1997, a community-randomized, placebo-controlled trial was conducted in the Sarlahi district of Nepal to assess the impact on maternal and infant health and survival of weekly vitamin A or beta-carotene supplementation among 44 646 women of childbearing age.[15,16] Local women delivered coded weekly supplements and recorded receipt of these supplements.

All women who applied for employment as distributors were interviewed in regard to demographic and socioeconomic characteristics such as age, literacy, years of education, number of hours spent doing housework, occupation of head of household, land ownership, type of dwelling and sanitation, and socioeconomic status (measured in terms of ownership of animals, ox carts, radios, watches, bicycles, furniture, and kitchen utensils). Information was also gathered on whether a weekly market, health post, school, or medicine shop was located in a given ward. Two years into the trial, the hours per week that distributors spent on this work were ascertained via interviews. Number of recipients for whom each distributor was responsible was recorded throughout the trial.

On the basis of analyses of socioeconomic data from the trial population, we developed a Guttman scale to characterize respondents' socioeconomic status as reflected by land, cattle, and ox cart ownership and whether their house had an upper story.[17] Scores could range from 0 (no land, cattle, or ox cart ownership and no upper story) to 4 (land, cattle, and ox cart ownership and upper story). Values below 3 denoted "lower" socioeconomic status.

For each supplement recipient, coverage was defined as the percentage of possible doses received in the trial. The overall coverage rate attained by each distributor was estimated as the mean coverage rate among all recipients for which that distributor was responsible. Low and high coverage rates were defined as less than 50% and 70% or more, respectively. This choice of cutoffs was based on the minimum coverage thought to produce beneficial health outcomes and the maximum coverage that would not lead to any improvements in community health.

Two multiple logistic regression models were fitted to the coverage data, one with low coverage and the other with high coverage as the outcome. Hours per week spent distributing supplements and doing housework and number of women for whom the distributor was responsible were entered into the model as continuous variables. Selection of variables for model inclusion was based on P values below .10 for individual associations with coverage.

RESULTS

Distributors spent an average of 5 hours per week delivering supplements. The mean coverage rate among the 345 distributors was 61% (range: 16% to 86%). The mean number of women to whom distributors were responsible for providing supplements was 105 (SD=36).

In comparison with women from villages without a weekly market, those from villages that contained a weekly market were more likely to have low coverage rates (odds ratio [OR]=2.32; 95% confidence interval [CI]=1.01, 5.28), to do more housework each week (42.5 hours vs 38.2 hours), and to distribute supplements to more women (114 vs 103; Table 1). Also, women with higher coverage rates were more likely to be illiterate (OR=5.07; 95% CI=0.94, 27.44). Finally, distributors with higher coverage rates spent 1 hour less per week distributing supplements, as they had fewer women to supplement.

TABLE 1—Distribution of Predictors of Low (<50%) and High (≥70%) Supplement Coverage and Adjusted Odds Ratios: Sarlahi District, Nepal, 1994–1997

	Coverage Rate, %							
	<50 (n = 55)		50–70 (n = 218)		≥70 (n = 72)			
Predictor	No. (%)	Mean	No. (%)	Mean	No. (%)	Mean	Low Coverage, Adjusted OR[a] (95% CI)	High Coverage, Adjusted OR[a] (95% CI)
Weekly market	10 (18.2)		21 (9.6)		7 (9.7)		2.32 (1.01, 5.28)	1.00 (0.41, 2.45)
Age, y								
<20	6 (9.6)		20 (9.2)		9 (12.7)		1.00	. . .
20–29	42 (69.1)		146 (67.0)		49 (69.0)		0.87 (0.32, 2.32)	0.85 (0.36, 1.99)
≥30	7 (21.3)		51 (23.4)		13 (18.3)		0.42 (0.12, 1.45)	0.88 (0.32, 2.43)
Illiteracy	1 (1.8)		2 (0.9)		5 (5.6)		0.94 (0.10, 8.92)	5.07 (0.94, 27.44)
Low socioeconomic status	27 (49.1)		82 (37.6)		33 (45.8)		1.71 (0.92, 3.19)	1.29 (0.74, 2.25)
Workload								
Supplement distribution, h/wk		5.5		5.6		4.5	1.00 (0.89, 1.12)	0.83 (0.73, 0.95)
Housework, h/wk		42.4		38.2		38.2	1.02 (1.00, 1.03)	1.00 (0.99, 1.02)
No. of women in sector		114		104		101	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)

Note. OR = odds ratio; CI = confidence interval.
[a]From multivariate logistic regression.

DISCUSSION

Our results show that successful supplement distribution can depend on factors external to the distributor (village, household, and recipient characteristics) as well as personal attributes (age and literacy). The weekly market may have made it more difficult for distributors to find recipients at home, distributors themselves may have made more visits to the market, or a weekly market might serve as a marker for other village characteristics that could distract women from compliance.

The literacy rate among the recipients was 14%.[15,16] Illiterate distributors may have related better to women who were more like themselves, leading to increased compliance. A similar result was seen in a vitamin A supplementation program for preschool children in Indonesia, where village-based male distributors with less education had higher coverage rates.[18]

As mentioned, distributors with higher coverage rates spent an average of 1 hour less per week doing their work. Amount of time spent was a function not only of the number of households to be visited but also of the amount of travel time between households. Hence, time spent on work might not mean more time spent with each recipient.

Amount of time spent distributing supplements did have an effect on coverage, but distributors with low coverage rates had an average of 13 more women to supplement than did those with high coverage rates. They also reported an average of 4 more hours of housework per week than did distributors with higher coverage rates. Thus, although there was no evidence that work conditions (hours spent and number of households visited) or competing demands (housework) were associated with high coverage, these factors did appear to be related to low coverage.

Recipient characteristics that predicted higher coverage rates included older age and higher parity, previous history of child deaths, and lower socioeconomic status.[19] This profile of the compliant recipient fits the notion that compliance is higher among distributors who are similar to recipients in terms of education level. If the results of ongoing and planned trials involving antenatal micronutrient supplementation confirm the benefits of such interventions, the success and sustainability of programs may be enhanced through a better understanding of the characteristics of distributors and recipients. ■

About the Authors

Joanne Katz, Keith P. West Jr, Lee Wu, Elizabeth Kimbrough Pradhan, Parul Christian, and Steven C. LeClerq are with the Department of International Health, Center for Human Nutrition, Johns Hopkins School of Hygiene and Public Health, Baltimore, Md. Subarna K. Khatry and Sharada Ram Shrestha are with the Nepal Nutrition Intervention Project, Sarlahi, Kathmandu, Nepal.

Requests for reprints should be sent to Joanne Katz, ScD, Johns Hopkins Bloomberg School of Public Health, 615 N Wolfe St, Room W5515, Baltimore, MD 21205–2103 (e-mail: jkatz@jhsph.edu).

This brief was accepted October 18, 2001.

Contributors

J. Katz conceived the analysis plan and wrote the brief. K.P. West provided scientific input into the design and conduct of the study. L. Wu conducted the analyses and assisted with interpretation of data. S.K. Khatry provided input into the design and conduct of the study. E. Kimbrough Pradhan contributed substantially to data collection and management, quality control, data analysis, and study design. P. Christian contributed to study design and implementation and provided input on the analysis and writing of the brief. S.C. LeClerq assisted with the study design and implementation and the oversight of data collection. S.R. Shrestha helped with the design, implementation, and oversight of the study.

Acknowledgments

This study was carried out by the Center for Human Nutrition and the Sight and Life Research Institute of the Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, through cooperative agreement HRN-A-00-97-00015-00 between Johns Hopkins University and the Office of Health and Nutrition, US Agency for International Development; a grant from the Bill and Melinda Gates Foundation, Seattle, Wash; Task Force Sight and Life, Roche Pharmaceuticals, Basel, Switzerland; and the United Nations Children's Fund, Nepal, in collaboration

with the National Society for the Prevention of Blindness, Kathmandu, Nepal.

References

1. Rahmathullah L, Underwood BA, Thulasiraj RD, et al. Reduced mortality among children in southern India receiving a small weekly dose of vitamin A. *N Engl J Med.* 1990;323:929–935.

2. West KP Jr, Pokhrel RP, Katz J, et al. Efficacy of vitamin A in reducing preschool child mortality in Nepal. *Lancet.* 1991;338:67–71.

3. Daulaire NMP, Starbuck ES, Houston RM, Church MS, Stukel TA, Pandey MR. Childhood mortality after a high dose of vitamin A in a high risk population. *BMJ.* 1992;304:207–210.

4. Ghana VAST Study Team. Vitamin A supplementation in northern Ghana: effects on clinic attendances, hospital admissions and child mortality. *Lancet.* 1993; 342:7–12.

5. Lawless JW, Latham MC, Stephenson LS, Kinoti SN, Pertet AM. Iron supplementation improves appetite and growth in anemic Kenyan primary school children. *J Nutr.* 1994;124:645–654.

6. Menendez C, Kahigwa E, Hirt R, et al. Randomised placebo-controlled trial of iron supplementation and malaria chemoprophylaxis for prevention of severe anemia and malaria in Tanzanian infants. *Lancet.* 1997;350:844–850.

7. Kolsteren P, Rahman SR, Hilderbrand K, Diniz A. Treatment for iron deficiency anaemia with a combined supplementation of iron, vitamin A and zinc in women of Dinajpur, Bangladesh. *Eur J Clin Nutr.* 1999; 53:102–106.

8. Osendarp SJM, van Raaij JMA, Arifee SE, Wahed MA, Baqui AH, Fuchs GJ. A randomized, placebo-controlled trial of the effect of zinc supplementation during pregnancy on pregnancy outcome in Bangladeshi urban poor. *Am J Clin Nutr.* 2000;71:114–119.

9. Smith JC, Makdani D, Hegar A, Rao D, Douglass LW. Vitamin A and zinc supplementation of preschool children. *J Am Coll Nutr.* 1999;18:213–222.

10. Berger J, Dyuck JL, Galan P, et al. Effect of daily iron supplementation on iron status, cell-mediated immunity, and incidence of infections in 6–36 month old Togolese children. *Eur J Clin Nutr.* 2000;54:29–35.

11. Umeta M, West CE, Haidar J, Deurenberg P, Hautvast JG. Zinc supplementation and stunted infants in Ethiopia: a randomised controlled trial. *Lancet.* 2000;355:2021–2026.

12. Sazawal S, Black RE, Bhan MK, et al. Zinc supplementation reduces incidence of persistent diarrhea and dysentery among low socioeconomic children in India. *J Nutr.* 1996;126:443–450.

13. Sazawal S, Black RE, Bhan MK, Jalla S, Sinhu A, Bhandari N. Efficacy of zinc supplementation in reducing the incidence and prevalence of acute diarrhea—a community-based, double-blind, controlled trial. *Am J Clin Nutr.* 1997;66:413–418.

14. Sazawal S, Black RE, Jalla S, Mazumdar S, Sinha A, Bhan MK. Zinc supplementation reduces the incidence of acute lower respiratory infections in infants and preschool children: a double-blind controlled trial. *Pediatrics.* 1998;102:1–5.

15. West KP Jr, Katz J, Khatry SK, et al. Double blind, cluster randomised trial of low dose supplementation with vitamin A or beta-carotene on mortality related to pregnancy in Nepal. *BMJ.* 1999;318:570–575.

16. Katz J, West KP Jr, Khatry SK, et al. Maternal low-dose vitamin A or beta-carotene supplementation has no effect on fetal loss and early infant mortality: a randomized cluster trial in Nepal. *Am J Clin Nutr.* 2000; 71:1570–1576.

17. Christian P, West KP Jr, Khatry SK, et al. Night blindness of pregnancy in rural Nepal—nutritional and health risks. *Int J Epidemiol.* 1998;27:231–237.

18. Tarwotjo I, West KP Jr, Mele L, et al. Determinants of community-based coverage: periodic vitamin A supplementation. *Am J Public Health.* 1989;79: 847–849.

19. Christian P, West KP Jr, Khatry SK, et al. Compliance to vitamin A and beta-carotene supplementation and pregnancy-related mortality in Nepali women—program implications. Paper presented at: International Vitamin A Consultative Group Meetings, February 12–15 2001, Hanoi, Vietnam.

Prevalence of Exclusive Breastfeeding Among US Infants: The Third National Health and Nutrition Examination Survey (Phase II, 1991–1994)

| Ruowei Li, MD, PhD, Cynthia Ogden, PhD, Carol Ballew, PhD, Cathleen Gillespie, BSc, and Laurence Grummer-Strawn, PhD

Because of the recognized benefits of exclusive breastfeeding,[1–7] the American Academy of Pediatrics and the World Health Organization strongly encourage exclusive breastfeeding for the first 6 months of life.[7,8] Unfortunately, national goals for exclusive breastfeeding in the United States are not yet established, at least in part because of a lack of data.

The Third National Health and Nutrition Examination Survey (NHANES III) sample is nationally representative and particularly valuable in providing data on exclusive breastfeeding. We used NHANES III data to examine the prevalence of exclusive breastfeeding among US infants to obtain baseline data for program evaluation and public health policymaking.

METHODS

NHANES, a series of cross-sectional surveys conducted by the National Center for Health Statistics (NCHS), is designed to produce nationally representative data on the civilian, noninstitutionalized US population. Details on the study's design and methods are described elsewhere.[9] NHANES III, conducted between 1988 and 1994 in participants aged 2 months and older, was divided into phases I (1988–1991) and II (1991–1994). Surveys from phases I and II included questions regarding initiation and duration of breastfeeding, whereas only the phase II survey included a question on supplements to breast milk.

NCHS collected the breastfeeding data by means of a household youth questionnaire during a home interview with a parent or other proxy respondent for the child. The current study included only children younger than 6 years. We estimated the proportion of children *ever breastfed* from the question, "Was _____ ever breastfed or fed breastmilk?"; the proportion *exclusively breastfed* at a given age from "How old was _____ when _____ was first fed something other than breastmilk or water, including formula, juice, or solid foods?"; and the proportion *still breastfed* at a given age from "How old was _____ when _____ completely stopped breastfeeding or being fed breastmilk?"

The sociodemographic and environmental factors that we considered in this analysis were those identified in previous studies as important predictors of breastfeeding behaviors.[10,11] NHANES III classified the race and ethnicity of respondents as non-Hispanic White, non-Hispanic Black, Mexican American, or other. The survey defined maternal age as that of the mother at the baby's birth. Education of household head was represented by the highest grade the family reference person completed. Smoking status referred to whether the mother smoked during pregnancy. We used self-reported maternal height and weight at the time of the household interview to calculate body mass index (BMI; weight in kg/height in m^2), which was categorized as normal weight (BMI<25), overweight (BMI=25.0–29.9), and obese (BMI≥30).[12]

NHANES III did not collect data on gestational age, maternal education, or parity, but

TABLE 1—Proportion of Children Exclusively Breastfed (BF), by Sociodemographic and Environmental Factors: NHANES III (Phase II, 1991–1994)

Characteristic	Exclusively BF at 7 Days		Exclusively BF at 2 mo		Exclusively BF at 4 mo		Exclusively BF at 6 mo	
	n	Weighted % (SE)	n	Weighted % (SE)	n	Weighted % (SE)	n	Weighted % (SE)
All infants	3836	47.4 (2.7)	3836	32.2 (2.4)	3651	19.4 (1.4)	3463	9.5 (1.3)
Male infants	1928	48.8 (3.5)	1928	32.8 (3.4)	1836	19.2 (1.6)	1732	9.3 (1.3)
Female infants	1908	46.0 (2.4)	1908	31.6 (2.0)	1815	19.7 (1.7)	1731	9.6 (1.8)
Birthweight								
<2500 g	324	25.8 (3.7)	324	17.5 (3.9)	307	13.2 (3.9)	297	7.3 (3.8)
Normal	3498	49.4 (2.9)	3498	33.5 (2.6)	3330	20.0 (1.5)	3152	9.7 (1.4)
Maturity								
Premature	304	28.4 (5.1)	304	14.7 (3.2)	291	5.2 (1.7)	282	2.7 (1.4)
Full term	2475	47.4 (3.3)	2475	31.9 (3.0)	2349	19.9 (1.8)	2221	9.6 (1.8)
Child race/ethnicity								
Non-Hispanic White	1398	55.0 (3.3)	1398	36.2 (3.2)	1290	22.3 (1.8)	1178	10.9 (1.7)
Non-Hispanic Black	1047	22.8 (2.1)	1047	16.2 (1.7)	1020	8.6 (1.5)	993	4.2 (0.9)
Mexican American	1112	47.7 (2.5)	1112	36.4 (2.1)	1081	21.4 (1.4)	1055	7.1 (0.9)
Other	279	42.6 (5.6)	279	30.5 (5.8)	260	18.2 (4.9)	237	11.5 (4.2)
Maternal age, y								
<20	577	28.6 (3.4)	577	21.0 (3.0)	553	9.4 (2.3)	535	4.1 (1.4)
20–24	1105	37.5 (2.6)	1105	25.1 (2.7)	1059	12.6 (1.6)	993	5.1 (1.3)
25–29	1118	52.0 (2.7)	1118	32.6 (2.6)	1058	18.9 (1.5)	1009	8.1 (1.3)
≥30	1026	58.4 (4.1)	1026	42.0 (3.7)	971	29.3 (2.6)	916	16.3 (2.5)
Maternal education								
Less than high school	887	29.5 (3.6)	887	21.9 (3.2)	844	11.8 (2.9)	815	5.7 (2.6)
High school	1111	38.1 (2.6)	1111	24.9 (2.5)	1048	12.9 (1.7)	982	4.5 (0.6)
Some college	543	55.5 (4.1)	543	34.2 (4.2)	510	24.5 (3.4)	473	12.4 (2.9)
College graduate	157	72.6 (5.8)	157	53.6 (7.7)	147	38.0 (5.1)	132	19.5 (6.6)
Household head education								
Less than high school	1296	29.8 (3.4)	1296	24.5 (3.6)	1256	13.6 (2.8)	1212	7.1 (2.8)
High school	1302	39.2 (3.0)	1302	24.2 (2.5)	1228	11.6 (1.8)	1158	5.4 (1.5)
Some college	613	56.0 (2.3)	613	37.1 (2.7)	574	27.6 (2.7)	538	16.3 (3.1)
College graduate	578	73.2 (4.1)	578	49.9 (4.6)	547	31.8 (2.8)	512	13.1 (2.7)
Parity								
Primiparous	1359	47.0 (2.9)	1359	29.7 (2.5)	1296	17.8 (1.9)	1223	7.3 (1.2)
Multiparous	2008	47.5 (3.5)	2008	33.4 (3.1)	1907	20.1 (1.8)	1816	10.0 (1.9)
Smoking in pregnancy								
Yes	681	32.8 (3.3)	681	17.4 (3.1)	646	8.1 (1.8)	611	3.3 (1.3)
No	3141	51.7 (2.7)	3141	36.4 (2.7)	2991	22.6 (1.7)	2838	11.2 (1.5)
Mother's BMI								
Normal (<25)	2095	52.5 (2.7)	2095	35.4 (2.5)	1989	21.0 (1.5)	1874	9.4 (1.1)
Overweight (25–29)	881	39.5 (3.6)	881	28.2 (3.7)	827	18.3 (3.1)	783	10.6 (2.7)
Obese (≥30)	662	38.7 (5.0)	662	25.9 (4.3)	640	15.1 (2.8)	617	9.2 (3.4)
Residence								
Metropolitan	2108	50.3 (3.2)	2108	35.1 (3.2)	2018	19.8 (1.7)	1912	8.7 (1.4)
Rural	1728	44.1 (4.5)	1728	28.9 (3.7)	1633	18.9 (2.1)	1551	10.3 (2.1)
Region								
Northeast	516	42.4 (4.1)	516	25.5 (3.4)	485	17.1 (2.7)	453	11.7 (3.2)
Midwest	713	48.5 (3.3)	713	33.4 (3.2)	664	22.8 (2.2)	633	9.4 (0.9)
South	1775	37.9 (4.3)	1775	23.3 (2.9)	1703	12.1 (2.0)	1615	4.4 (0.6)
West	832	64.9 (4.3)	832	50.0 (3.8)	799	29.2 (1.7)	762	15.5 (2.5)
Poverty–income ratio								
0–99%	1378	32.2 (4.0)	1378	23.2 (3.5)	1329	13.2 (2.7)	1290	6.8 (2.4)
100–184%	822	42.5 (3.2)	822	29.1 (2.8)	781	17.0 (2.0)	731	10.4 (1.6)
185–349%	844	52.9 (3.4)	844	39.9 (3.6)	790	24.0 (2.4)	746	9.8 (1.5)
>350%	523	66.2 (3.8)	523	37.3 (3.1)	496	23.9 (3.2)	460	10.9 (3.2)
Missing values	269	37.5 (6.6)	269	29.2 (6.0)	255	17.4 (4.4)	236	11.9 (3.8)

Notes. (SE) = standard error. Missing numbers for each factor can be derived from the difference between the total number of infants and the summary number for each factor (see the example given under "Poverty–Income Ratio").

TABLE 2—Ever Breastfed and Breastfeeding Duration, by Sociodemographic and Environmental Factors: NHANES III (Phases I and II, 1988–1994)

Characteristic	Ever Breastfed		Breastfeeding at 6 mo		Breastfeeding at 12 mo	
	n	Weighted % (SE)	n	Weighted % (SE)	n	Weighted % (SE)
All infants	8215	53.6 (1.7)	7363	22.4 (1.2)	6123	8.9 (0.8)
Male infants	4062	53.7 (2.1)	3636	23.0 (1.7)	3009	9.2 (1.2)
Female infants	4153	53.6 (1.7)	3727	21.9 (1.2)	3114	8.5 (0.8)
Birthweight						
<2500 g	669	34.1 (3.0)	617	11.8 (2.2)	533	3.9 (1.5)
Normal	7350	55.5 (1.8)	6557	23.5 (1.3)	5417	9.3 (0.9)
Maturity						
Premature	709	35.7 (4.3)	650	11.1 (2.3)	554	6.5 (2.0)
Full term	5597	54.2 (1.8)	4997	23.4 (1.6)	4158	9.0 (1.0)
Child race/ethnicity						
Non-Hispanic White	3067	60.2 (2.0)	2580	26.6 (1.6)	1863	10.6 (1.3)
Non-Hispanic Black	2172	26.3 (1.5)	2033	8.3 (0.8)	1842	2.7 (0.6)
Mexican American	2479	56.0 (2.0)	2325	22.5 (1.4)	2112	9.9 (0.8)
Others	497	53.6 (3.5)	425	19.0 (2.8)	306	6.5 (1.8)
Maternal age, y						
<20	1265	32.5 (2.7)	1155	9.5 (1.7)	981	4.0 (1.0)
20–24	2321	44.8 (1.9)	2084	13.4 (1.4)	1737	4.4 (0.9)
25–29	2494	58.5 (1.9)	2236	24.8 (1.7)	1877	9.8 (1.4)
≥30	2104	65.3 (2.3)	1857	33.3 (2.0)	1497	13.7 (1.4)
Maternal education						
Less than high school	1584	30.3 (2.1)	1426	10.0 (1.7)	1187	3.9 (1.1)
High school	2201	45.4 (2.3)	1941	16.0 (1.2)	1522	5.7 (1.2)
Some college	994	62.1 (2.6)	865	26.3 (2.8)	661	8.2 (1.9)
College graduate	283	81.8 (3.4)	233	43.6 (5.8)	159	20.7 (5.4)
Household head education						
Less than high school	2839	35.8 (2.1)	2616	14.9 (1.9)	2276	7.2 (1.3)
High school	2690	46.0 (2.1)	2391	15.9 (1.1)	1955	6.5 (1.2)
Some college	1349	60.1 (1.9)	1182	24.9 (1.9)	979	10.0 (1.6)
College graduate	1149	80.2 (2.2)	1002	39.1 (2.4)	770	13.5 (2.0)
Parity						
Primiparous	2895	55.2 (1.7)	2582	19.1 (1.4)	2096	6.0 (0.8)
Multiparous	4355	52.2 (2.0)	3903	24.6 (1.6)	3246	10.2 (1.2)
Smoking in pregnancy						
Yes	1632	35.9 (2.1)	1452	9.4 (1.0)	1161	4.4 (1.0)
No	6551	59.1 (1.7)	5884	26.4 (1.4)	4937	10.2 (1.0)
Mother's BMI						
Normal (<25)	4617	58.1 (1.7)	4107	25.0 (1.3)	3345	10.0 (1.0)
Overweight (25–29)	1833	46.4 (2.4)	1621	17.3 (1.7)	1361	5.7 (0.9)
Obese (≥30)	1262	44.8 (3.6)	1166	16.9 (2.8)	1011	5.6 (1.8)
Residence						
Metropolitan	4370	58.8 (2.2)	3920	24.1 (1.8)	3290	9.8 (1.2)
Rural	3845	48.5 (2.9)	3443	20.8 (2.0)	2833	7.9 (1.2)
Region						
Northeast	1058	47.6 (3.5)	906	18.2 (2.4)	703	8.2 (1.4)
Midwest	1575	56.2 (2.1)	1398	24.8 (1.6)	1081	9.3 (2.0)

Continued

we obtained this information from data linkages with the children's birth certificates. We defined prematurity as a gestational age of less than 37 weeks. Maternal education was classified the same way as for household-head education, and parity was classified as primiparous (i.e., the mother had had no previous live births) or multiparous.

The entire 6-year NHANES III sampled a total of 8765 children younger than 6 years. The overall interview response rate for these children was approximately 94%. Data on ever breastfeeding, exclusive breastfeeding, and breastfeeding duration were available for approximately 99% of the interviewed children. We calculated the weighted percentages and their standard errors with SUDAAN to take into account the complex sample design.[13]

RESULTS

The proportions of children exclusively breastfed were approximately 47% at 7 days after birth, 32% at 2 months, 19% at 4 months, and 10% at 6 months, whereas the proportions of children still being breastfed at these ages were approximately 52%, 40%, 29%, and 22%, respectively. Table 1 shows that at each of these time points, exclusive breastfeeding was least common among low-birthweight, premature, or non-Hispanic Black infants and those of mothers who were younger than 20 years, had lower education or income, smoked during pregnancy, or lived in the South.

The proportions of children ever breastfed and still being breastfed at 6 and 12 months were also stratified by sociodemographic and environmental factors (Table 2). We observed patterns similar to those for the exclusively breastfed children. In addition, we found lower breastfeeding initiation and duration among mothers who were overweight or obese and among families living in rural areas. Although primiparous mothers had a higher rate of initiating breastfeeding than did multiparous mothers, they had a lower rate of continuing breastfeeding throughout the infant's first year.

DISCUSSION

Less than half of the children in NHANES III began exclusive breastfeeding. At age 2

TABLE 2—*Continued*

South	3360	44.1 (1.8)	3024	16.0 (1.7)	2588	5.4 (1.2)
West	2222	69.8 (3.6)	2035	32.7 (2.5)	1751	14.0 (1.6)
Poverty–income ratio						
0–99%	2857	36.5 (2.6)	2636	14.3 (1.7)	2286	7.6 (1.1)
100–184%	1758	50.1 (1.9)	1572	20.3 (1.5)	1307	6.5 (1.2)
185–349%	1847	58.5 (1.9)	1625	27.3 (1.9)	1302	11.9 (1.8)
> 350%	994	75.4 (2.2)	854	30.7 (2.7)	657	9.2 (2.0)
Missing value	759	45.8 (3.6)	676	14.9 (1.9)	571	6.4 (1.6)

Notes. (SE) = standard error. Missing numbers for each factor can be derived from the difference between the total number of infants and the summary number for each factor (see the example given under "Poverty–Income Ratio").

months, the percentage of infants still being exclusively breastfed was considerably lower than the percentage who were receiving *any* breast milk at this point. By the age of 6 months, slightly less than 10% of infants were being exclusively breastfed.

Although the factors that influence the initiation and duration of breastfeeding have been broadly studied,[10,11] previous studies have rarely examined the factors associated with exclusive breastfeeding. Our study indicates that the proportion of infants exclusively breastfed varied by subgroup, with the lowest rate found among non-Hispanic Black and premature infants and the highest rate among infants of mothers who had graduated from college. Our study also suggests that the factors associated with exclusive breastfeeding were similar to those associated with the initiation and duration of any breastfeeding.[10,11]

Our results regarding the initiation and duration of breastfeeding are similar to those from previous Ross Laboratories Mothers' Surveys.[14] Our analysis showed that only 3 subgroups in NHANES III met the Healthy People 2010 goal[15] of 75% for breastfeeding initiation: mothers who had graduated from college (81.8%), families with a household head who had graduated from college (80.2%), and families with an income exceeding 350% of the poverty–income ratio (75.4%). None of the subgroups met the goals for breastfeeding at 6 months (50%) or 12 months (25%).

In summary, this is the first nationally representative study available that indicates that initiation and maintenance of exclusive breastfeeding are low in the United States. Public health efforts are needed to improve the rate of exclusive breastfeeding—and, in particular, the duration of such feeding—among non-Hispanic Blacks and socioeconomically disadvantaged groups. ■

About the Authors

Ruowei Li, Carol Ballew, Cathleen Gillespie, and Laurence Grummer-Strawn are with the Division of Nutrition and Physical Activity, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Ga. Cynthia Ogden is with the Division of Health Examination Statistics, the National Center for Health Statistics, Hyattsville, Md.

Requests for reprints should be sent to Ruowei Li, Division of Nutrition and Physical Activity, Mail Stop K-25, Centers for Disease Control and Prevention, 4770 Buford Hwy NE, Atlanta, GA 30341-3717 (e-mail: ril6@cdc.gov).

This article was accepted May 1, 2001.

Contributors

R. Li planned the study, analyzed the data, and wrote the brief. C. Gillespie extracted the data set and conducted preliminary analysis for this study. C. Ogden linked the NHANES III with children's birth certificates and analyzed data for gestational age, maternal education, and parity. C. Ballew and L. Grummer-Strawn assisted with study design, supervised data analysis, and contributed to the writing of the brief.

References

1. Cushing AH, Samet JM, Lambert WE, et al. Breastfeeding reduces risk of respiratory illness in infants. *Am J Epidemiol.* 1998;147(9):863–870.

2. Scariati PD, Grummer-Strawn LM, Fein SB. A longitudinal analysis of infant morbidity and the extent of breastfeeding in the United States. *Pediatrics.* 1997; 99(6):E5.

3. Duncan B, Ey J, Holberg CJ, et al. Exclusive breast-feeding for at least 4 months protects against otitis media. *Pediatrics.* 1993;91(5):867–872.

4. Raisler J, Alexander C, O'Campo P. Breast-feeding and infant illness: a dose-response relationship? *Am J Public Health.* 1999;89(1):25–30.

5. Dewey KG, Heinig MJ, Nommsen-Rivers LA. Differences in morbidity between breast-fed and formula-fed infants. *J Pediatr.* 1995;126(5 Pt I):696–702.

6. Howie PW, Forsyth JS, Ogston SA, et al. Protective effect of breast feeding against infection. *BMJ.* 1990;300(6716):11–16.

7. Work Group on Breastfeeding, American Academy of Pediatrics. Breastfeeding and the use of human milk. *Pediatrics.* 1997;100(6):1035–1039.

8. World Health Organization: The optimal duration of exclusive breastfeeding. Note for the Press No. 7, April 2, 2001. Available at: www.who.int/inf-pr-2001/er/note2001-07.html. Accessed May 13, 2002.

9. National Center for Health Statistics. Plan and operation of the Third National Health and Nutrition Examination Survey 1988–1994. *Vital Health Stat.* 1994; 1(32).

10. Scott JA, Binns CW. Factors associated with the initiation and duration of breastfeeding: a review of the literature. *Breastfeeding Rev.* 1999;7(1):5–16.

11. Tan JC, Jeffery HE. Factors that influence the choice of infant feeding. *J Paediatr Child Health.* 1995; 31(5):375–378.

12. World Health Organization Expert Committee. Physical status: the use and interpretation of anthropometry. *World Health Organ Tech Rep Ser.* 1995;854: 1–452.

13. Shah BV, Barnwell BG, Bieler GS. *SUDAAN User's Manual, Release 7.5.* Research Triangle Park, NC: Research Triangle Institute; 1997.

14. Ryan AS. The resurgence of breastfeeding in the United States. *Pediatrics.* 1997;99(4):1–5.

15. *Healthy People 2010.* Washington, DC: US Department of Health and Human Services; 2000.

Sexual and Drug Use Behavior Among Women Who Have Sex With Both Women and Men: Results of a Population-Based Survey

Susan Scheer, PhD, MPH, Ingrid Peterson, MPH, Kimberly Page-Shafer, PhD, MPH, Viva Delgado, MPH, Alice Gleghorn, PhD, Juan Ruiz, MD, DrPH, Fred Molitor, PhD, William McFarland, MD, PhD, Jeffrey Klausner, MD, MPH, and the Young Women's Survey Team

Recent HIV/AIDS trends in the United States suggest a relative increase in HIV infections among women attributed to injection drug

use or heterosexual contact.[1] Although the biological risk of female-to-female sexually transmitted HIV is unknown, it is thought to be much lower than the risk of transmission between men and women, including instances in which a condom is used.[2] However, studies focusing on women who have sex with women (WSW) have shown that some subgroups of WSW exhibit high levels of sexual risk behaviors with men as well as unsafe injection drug use.[3,4] Thus, if risk assumptions are based on self-reported or presumed sexual identity, possible risks for HIV infection may be underestimated in some subgroups of WSW.

Few studies have estimated the proportion of WSW or characterized their behavior in samples representative of the population as a whole. Here we describe sexual and drug use behaviors associated with HIV and other sexually transmitted diseases (STDs) among WSW who took part in a door-to-door, population-based survey of women aged 18 to 29 years. The survey was conducted between April 1996 and January 1998 among residents of low-income neighborhoods in Northern California. Study methods have been described in detail in a previous article.[5]

Of 2547 women who completed the study, 2229 (88%) reported sex exclusively with men, 189 (7%) reported sex with both men and women, and 16 (1%) reported sex exclusively with women. Of the 7 HIV-positive women, 4 reported only male partners, 2 reported both male and female partners, and 1 reported only female partners. None of the 16 WSW who reported sex exclusively with women reported any injection drug use. Therefore, analyses of risk were limited to those who reported sex with both men and women and those who reported sex exclusively with men (Table 1).

Compared with women who had sex exclusively with men, women who had sex with both men and women were significantly more likely to report past and recent high-risk sexual behavior, including sex with an HIV-positive man, multiple male sexual partners, sex with a man who has sex with men, sex with an injection drug user, trading of sex for drugs or money, and anal sex. They were also more likely to report past and recent injection drug use, including use of heroin, cocaine, and speed. Finally, they were more likely to have serological markers for both hepatitis B virus (anti-HBc, HbsAG, or both) and hepatitis C virus (anti-HCV). Rates of HIV and other STDs did not significantly differ owing to the small numbers of these infections.

The rates of sexual and injection drug risk activities exhibited by women in this population-based survey who reported sex with both men and women place this group at potentially higher risk of HIV and other STDs than women who were exclusively sexual with either men or women. Prevention efforts should avoid assumptions based on reported sexual identity and should acknowledge that women who report sex with both women and men may be at increased risk for HIV and other STDs. ■

TABLE 1—Prevalence of Sexual Behaviors, Injection Drug Use, and STD/HIV Infections: Population-Based Survey, Northern California

Risk Behavior or Marker	Women Who Reported Sex Exclusively With Men[a]	Women Who Reported Sex With Men and Women[a]	*P*
Sexual risk			
Ever had sex with MSM, %	3	30	<.001
Sex with MSM within 6 months, %	<1	10	<.001
Ever had sex with an IDU, %	8	38	<.001
Sex with IDU within 6 months, %	2	19	<.001
Ever had sex with an HIV-positive man, %	1	5	<.001
Sex with HIV-positive man within 6 months, %	<1	2	<.001
Mean No. of lifetime male partners	16	307	<.001
Mean No. of male partners within 6 months	2	9	<.001
Ever traded sex for drugs or money, %	8	40	<.001
Traded sex for drugs or money within 6 months, %	4	24	<.001
Ever had anal sex, %	18	56	<.001
Injection drug use, %			
Ever injected cocaine	<1	10	<.001
Injected cocaine within 6 months	<1	5	<.001
Ever injected heroin	1	17	<.001
Injected heroin within 6 months	<1	9	<.001
Ever injected speed	1	13	<.001
Injected speed within 6 months	<1	6	<.001
Ever shared needles	43	64	.08
Shared needles within 6 months	46	27	.18
STD/HIV prevalence, %			
HIV positive	<1	<1	.27
Syphilis	<1	1	.33
Chlamydia	2	<1	.08
Gonorrhea	<1	1	.64
Hepatitis B	5	9	.03
Hepatitis C	1	8	<.001

Note. Women who reported no sex, refused to report sexual activity, or reported sex with women only were excluded.
STD = sexually transmitted disease; MSM = men who have sex with men; IDU = injection drug user.
[a]Population prevalence adjusted for the survey design.

About the Authors

Susan Scheer, Ingrid Peterson, Viva Delgado, Alice Gleghorn, William McFarland, and Jeffrey Klausner are with the San Francisco Department of Public Health, San Francisco, Calif. Kimberly Page-Shafer is with the University of California Center for AIDS Prevention Studies, San Francisco. Juan Ruiz and Fred Molitor are with the California Department of Health Services, Office of AIDS, Sacramento.

Requests for reprints should be sent to Susan Scheer, PhD, MPH, San Francisco Department of Public Health,

AIDS Office, 25 Van Ness Ave, Suite 500, San Francisco, CA 94102 (e-mail: susan.scheer@SFDPH.org).

This brief was accepted June 6, 2001.

Contributors

S. Scheer contributed to the conception, design, analysis, and writing of the brief. I. Peterson and K. Page-Shafer assisted with management and analysis of the data. V. Delgado, J. Ruiz, F. Molitor, W. McFarland, and J. Klausner contributed to coordination of and data collection for the study; they also contributed to the interpretation of the data. A. Gleghorn contributed to the conception and design of the study. All of the authors reviewed and edited the brief.

Acknowledgments

The Young Women's Survey Team includes Geneva Bell-Sanford, San Joaquin County, California, Department of Public Health; Gail Bolan, San Francisco Department of Public Health; Cynthia Cossin, Viral and Rickettsial Disease Laboratory, Berkeley, California; Carla Dillard Smith, California Prevention Education Program, Oakland; Maria Hernandez, San Francisco Department of Public Health; Tanya Holmes, Alameda County, California, Department of Public Health; Martin Lynch and Juan Reardon, Contra Costa County, California, Department of Public Health; Charlotte Smith, San Mateo County, California, Department of Public Health; Hypolitta Villa, San Joaquin County, California, Department of Public Health; and Francis Wiser, San Mateo County, California, Department of Public Health.

This work was supported in part by cooperative agreements U62/CCU0200, U62/CCU906250-06, U62/CU902019-12, and U61/CCU902019-13 from the Centers for Disease Control and Prevention. Additional funding was provided by the San Francisco AIDS Office and STD Prevention and Control Section.

We thank the public health laboratory directors in the counties of Alameda, Contra Costa, San Francisco, San Mateo, and San Joaquin and the staff from the Viral and Rickettsial Disease Laboratory who tested specimens. Also, we thank Dr Gail Bolan and Harold Rasmussen for their support.

References

1. *HIV/AIDS Surveillance Report.* Atlanta, Ga: Centers for Disease Control and Prevention; 1999;11(2).

2. Vernazza PL, Eron JJ, Fiscus SA, Cohen MS. Sexual transmission of HIV: infectiousness and prevention. *AIDS.* 1999;13:155–166.

3. Gonzales V, Washienko KM, Krone MR, et al. Sexual and drug-use risk factors for HIV and STDs: a comparison of women with and without bisexual experiences. *Am J Public Health.* 1999;89:1841–1846.

4. Bevier PJ, Chiasson MA, Heffernan RT, Castro KG. Women at a sexually transmitted disease clinic who reported same-sex contact: their HIV seroprevalence and risk behaviors. *Am J Public Health.* 1995;85:1366–1371.

5. Ruiz JD, Molitor F, McFarland W, et al. Prevalence of HIV infection, sexually transmitted diseases, and hepatitis and related risk behavior in young women living in low-income neighborhoods of northern California. *West J Med.* 2000;172:368–373.

A Tailored Intervention to Aid Decisionmaking About Hormone Replacement Therapy

Colleen M. McBride, PhD, Lori A. Bastian, MD, Susan Halabi, PhD, Laura Fish, MPH, Isaac M. Lipkus, PhD, Hayden B. Bosworth, PhD, Barbara K. Rimer, DrPH, and Ilene C. Siegler, PhD, MPH

Decision aids related to hormone replacement therapy (HRT), whether delivered in written form,[1,2] along with audiotapes,[3] or as part of discussion groups,[4] have outperformed generic brochures in increasing knowledge and accuracy of risk assessments. However, these decision aids have provided women with population-based estimates of average risk, not individual risk levels that may have bearing on their decisions about HRT. Decision aids individually customized or "tailored" to include only the most relevant information could make it easier for women to consider HRT's risks and benefits.[5,6] Tailored interventions have yet to be evaluated for HRT decisions. We describe the effect of a tailored decision aid on women's accuracy of perceived risk for breast cancer, confidence to decide about HRT, and satisfaction with the decision.

METHODS

Study Design

Between October 1998 and February 1999, interviewers called households from a purchased list to identify women aged 45 to 54 years who were willing to receive written materials about HRT and who did not have a history of breast cancer. Eligible women stratified by baseline HRT use were randomized to either a delayed or an active intervention arm. Women in the active arm received materials 2 weeks after the baseline survey; those in the delayed arm received materials after completing the study. Telephone surveys were conducted at 1 and 9 months. Study protocols were approved by the institutional review board.

Intervention

The trifold decision aid[7] included (1) "Step 1 The Facts" (19 pages), which was tailored to baseline perceived menopausal status, hysterectomy status (no or yes), prior HRT use, and accuracy of perceived risk for breast cancer[8]; (2) "Step 2 What's Important to You," a worksheet to record preferences; "Step 3 Next Steps" (13 pages), which included vignettes of women at decision points similar to those of women receiving the intervention and a checklist of questions for the health care providers of women receiving the intervention.

Outcome Measures

Accuracy was the agreement between women's perceived and objective 10-year risk for breast cancer as measured by the Gail score.[8] Perceived risk was assessed on a 0 (certain not to happen) to 100 (certain to happen) scale. Breast cancer risk factors were used to calculate a Gail score.[8] Accuracy was computed as the difference between the woman's perceived and objective risk score. The woman's perception was accurate unless the absolute value of the difference score exceeded 10%.[9]

Level of confidence in ability to understand the risks and benefits of HRT, make a decision about HRT, and discuss HRT with a health care provider was rated (0=low to 10=high; Cronbach α=.78). Items were summed to yield an average level of confidence.

Women's satisfaction was assessed by agreement (1=strongly disagree to 5=strongly agree) with 6 statements related to being informed about HRT, whether the decision (for those who made a decision) was consistent with their personal values, and overall satisfaction with the decision among those who had made a decision (Cronbach α=.78).[10]

Statistical Analysis

Logistic regression models were tested to predict dichotomized confidence (based on the median baseline value), accuracy of perceived breast cancer risk, and satisfaction outcomes at each follow-up. Covariates were intervention arm, baseline value, race/ethnicity, education, marital status, working for pay, perceived menopausal status, ever use of HRT, hysterectomy status, decision status, and numeracy (for the accuracy outcome).

RESULTS

Recruitment and Follow-Up

Of the 2388 telephone numbers called, 158 (7%) numbers were not working, 844 (35%) people were ineligible, 444 (19%) calls were never answered, and 361 (15%) people refused to participate. Of the 581 women who were randomized, 557 (96%) and 541 (93%) completed the 1- and 9-month surveys, respectively. Complete data are available for 520 (90%) of the women.

Study Outcomes

Women in the active intervention arm were significantly less likely than those in the delayed arm to be working for pay (P=.01) and to have had a hysterectomy (P=.02) (Table 1).

Confidence in ability to decide about HRT. Women in the active arm were more likely than those in the delayed arm to be confident about making a decision at both follow-ups (1 month: odds ratio [OR]=2.5; 95% confidence interval [CI]=1.6, 3.9; 9 months: OR=2.8; 95% CI=1.8, 4.5, respectively) (Table 2). Women in the active arm who were confident at 1 month were more likely to remain confident in their decision at 9 months than were comparable women in the delayed arm (OR=2.5; 95% CI=1.6, 4.0).

Accuracy of perceived risk for breast cancer. At both follow-ups, women in the active arm were more likely to accurately perceive their level of risk for breast cancer than were those in the delayed arm (1 month: OR=1.9; 95% CI=1.3, 2.9; 9 months: OR=1.9; 95% CI=1.2, 2.8, respectively). Among those with accurate risk perceptions at 1 month, women in the active arm were more likely than those in the delayed arm to retain those perceptions at 9 months (OR=2.2; 95% CI=1.3, 3.7).

Satisfaction with decision. At 1 month only, women in the active arm were more likely than those in the delayed arm to report that they were very satisfied with their HRT decision (OR=2.5; 95% CI=1.5, 4.3). However, among those who reported being satisfied at 1 month, women in the active arm were more likely than women in the delayed arm to remain satisfied with their decision at 9 months (OR=2.8; 95% CI=1.5, 5.3).

DISCUSSION

The decision aid improved the accuracy of women's perceptions of breast cancer risk, confidence to make decisions about HRT, and satisfaction with decisions. These intervention effects were sustained between 1- and 9-month follow-ups. HRT decisions might benefit by additional customization or brief telephone counseling calls, which have been effective for other health-related outcomes.[11]

More than 40% of the undecided women in the active arm made a decision by the 9-month follow-up; in addition, among those decided at baseline and 1 month, women in the active arm reported greater satisfaction at 9-month follow-up than did those in the delayed arm. Providing such decision aids to women before clinic appointments could enable them to make better use of limited visit time.

Although most of the participants were more educated and health oriented than the general population, our use of purchased lists resulted in recruitment of a broader cross-section of women than have been included in prior research on this topic; 24% of our par-

TABLE 1—Baseline Characteristics, by Intervention Arm

	Total (N=581)	Delayed Intervention (n=292)	Active Intervention (n=289)	P
Demographics				
Age, %				.98
45–49	55	55	55	
50–54	45	45	45	
Some college or more, %	75	75	76	.83
Married or living as married, %	78	77	78	.82
African American, %	24	24	23	.73
Work for pay, %	85	89	81	.01
Money for special things, %	67	67	67	.95
Health services				
Had health insurance, %	96	96	96	.70
Had regular doctor, %	94	94	94	.85
Menopause related				
Perceive menopause beginning or begun, %	71	71	70	.49
Clinically postmenopausal, %	50	51	49	.65
Symptoms present	90	91	89	.47
Hormone replacement therapy related				
Ever used hormone replacement therapy, %	41	40	41	.89
Decided about hormone replacement therapy, %	47	46	48	.68
Decided to use hormone replacement therapy[a]	80	79	80	.81
Satisfaction with decision, mean (SD)[a]	4.2 (0.7)	4.1 (0.7)	4.2 (0.7)	.17
Currently using hormone replacement therapy, %	29	30	28	.59
Confidence to decide about hormone replacement therapy, mean (SD)	6.6 (2.2)	6.6 (2.2)	6.7 (2.2)	.47
Health-related variables				
Health good or excellent, %	90	90	90	.97
Currently smoking, %	12	10	13	.23
Perceived likely to get breast cancer in lifetime, %	49	50	47	.42
Perceived breast cancer risk accurately, %	32	34	29	.19
Had hysterectomy, %	24	28	20	.02
Has family history of breast cancer, %	11	13	10	.31

[a]Assessed among the 269 women who had decided about hormone replacement therapy.

TABLE 2—Odds Ratios (ORs) and 95% Confidence Intervals (CIs) for Study Outcomes at 1 Month, 9 Months, and Sustained Between 1 and 9 Months

	n	Active vs Delayed Intervention, OR (95% CI)	*P*
Confidence[a]			
1 mo	531	2.5 (1.6, 3.9)	<.0001
9 mo	514	2.8 (1.8, 4.5)	<.0001
1 and 9 mo	507	2.5 (1.6, 4.0)	<.0001
Accuracy of perceived risk for breast cancer[b]			
1 mo	500	1.9 (1.3, 2.9)	.002
9 mo	471	1.9 (1.2, 2.8)	.004
1 and 9 mo	464	2.2 (1.3, 3.7)	.004
Satisfied with decision about hormone replacement therapy[c]			
1 mo	256	2.5 (1.5, 4.3)	.0005
9 mo	310	1.5 (0.9, 2.4)	.09
1 and 9 mo	212	2.8 (1.5, 5.3)	.001

[a]Dichotomous outcome, proportion above the cutpoint (baseline sample median) at each follow-up. The confidence models were adjusted by race/ethnicity, education, age, marital status, work for pay, perceived menopausal status, ever use of hormone replacement therapy, hysterectomy status, baseline confidence, and hormone replacement therapy decision status.
[b]Accurate if the difference between the 10-year perceived risk and the 10-year objective risk is less than or equal to 10% (based on the Gail score[8]). The accuracy-of-perceived-risk models were adjusted by race/ethnicity, education, age, marital status, work for pay, perceived menopausal status, ever use of hormone replacement therapy, baseline accuracy of perceived risk, and numeracy.
[c]Dichotomous outcome, proportion above the cutpoint (baseline sample median) at each follow-up. The satisfaction models were adjusted by race/ethnicity, education, age, marital status, work for pay, perceived menopausal status, ever use of hormone replacement therapy, and hysterectomy status.

ticipants (vs 10% of those in previous studies[2,3]) were African Americans. Involving community groups should be considered as a means of further expanding intervention reach.

As ever-increasing numbers of women enter menopause, rapidly changing knowledge about HRT requires innovative and flexible communication strategies to meet their information needs. ■

About the Authors

Colleen M. McBride, Susan Halabi, Laura Fish, Isaac M. Lipkus, Barbara K. Rimer, and Ilene C. Siegler are with The Comprehensive Cancer Center, Durham, NC. Colleen M. McBride, Lori A. Bastian, Susan Halabi, Isaac M. Lipkus, Hayden B. Bosworth, Barbara K. Rimer, and Ilene C. Siegler are with Duke University Medical Center, Durham, NC. Lori A. Bastian and Hayden B. Bosworth are also with Durham Veteran's Administration Medical Center, Durham, NC. Barbara K. Rimer is also with the National Cancer Institute, Division of Cancer Control and Population Sciences, Bethesda, Md.

Requests for reprints should be sent to Colleen M. McBride, PhD, Duke Cancer Prevention, Detection, and Control Research, DUMC Box 2949, Durham, NC 27710-2949 (e-mail: mcbri002@mc.duke.edu).

This brief was accepted May 23, 2001.

Contributors

All authors contributed to the conception, analysis, interpretation, and writing of the brief. C.M. McBride, L.A. Bastian, L. Fish, I.M. Lipkus, H.B. Bosworth, B.K. Rimer, and I.C. Siegler were key to the development, implementation, and evaluation of the intervention. C.M. McBride, L.A. Bastian, S. Halabi, and I.C. Siegler oversaw data collection, analysis, and interpretation.

Acknowledgments

This work was supported by a grant from the National Cancer Institute (PO1-CA-72099-05).

The authors would like to acknowledge Dr Celette Sugg Skinner for helpful comments on earlier drafts. The authors also thank Maragatha Kuchibhatla and Pauline Lyna, who assisted in data analyses, and Pamela Harris for her assistance in preparing the manuscript.

References

1. O'Connor AM, Fiset V, DeGrasse C, et al. Decision aids for patients considering options affecting cancer outcomes: evidence of efficacy and policy implications. *J Natl Cancer Inst Monogr.* 1999;25:67–80.

2. Rothert ML, Holmes-Rovner M, Rovner D, et al. An educational intervention as decision support for menopausal women. *Res Nurs Health.* 1997;20: 377–387.

3. O'Connor AM, Tugwell P, Wells GA, et al. Randomized trial of a portable, self-administered decision aid for postmenopausal women considering long-term preventive hormone therapy. *Med Decis Making.* 1998; 18:295–303.

4. Hampson SE, Hibbard JH. Cross-talk about the menopause: enhancing provider-patient interactions about the menopause and hormone therapy. *Patient Educ Couns.* 1996;27:177–184.

5. Rimer BK, Conaway M, Lyna P, et al. The impact of tailored interventions on a community health center population. *Patient Educ Counseling.* 1999;37:125–140.

6. Kreuter MW. Dealing with competing and conflicting risks in cancer communication. *J Natl Cancer Inst.* 1999;25:27–35.

7. Bastian LA, McBride CM, Fish L, Lipkus IM, Rimer BK, Siegler I. Evaluating participants' use of an HRT decision-making intervention. *Patient Educ Couns.* In press.

8. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst.* 1989;81:1879–1886.

9. Lerman C, Lustbader E, Rimer B, et al. Effect of individualized breast cancer risk counseling: a randomized trial. *J Natl Cancer Inst.* 1995;87:286–292.

10. Holmes-Rovner M, Kroll J, Rovner DR, et al. Patient decision support intervention: increased consistency with decision analytic models. *Med Care.* 1999; 37:270–284.

11. McBride CM, Rimer BK. Using the telephone to improve health behavior and health service delivery. *Patient Educ Couns.* 1999;37:3–18.

Effects of Hospital Staffing and Organizational Climate on Needlestick Injuries to Nurses

Sean P. Clarke, PhD, RN, Douglas M. Sloane, PhD, and Linda H. Aiken, PhD, RN

Objectives. This study determined the effects of nurse staffing and nursing organization on the likelihood of needlestick injuries in hospital nurses.

Methods. We analyzed retrospective data from 732 and prospective data from 960 nurses on needlestick exposures and near misses over different 1-month periods in 1990 and 1991. Staffing levels and survey data about working climate and risk factors for needlestick injuries were collected on 40 units in 20 hospitals.

Results. Nurses from units with low staffing and poor organizational climates were generally twice as likely as nurses on well-staffed and better-organized units to report risk factors, needlestick injuries, and near misses.

Conclusions. Staffing and organizational climate influence hospital nurses' likelihood of sustaining needlestick injuries. Remedying problems with understaffing, inadequate administrative support, and poor morale could reduce needlestick injuries. (*Am J Public Health.* 2002;92:1115–1119)

Exposures of health care workers to bloodborne pathogens through accidental contact with sharp instruments have been widely publicized, and the prevention and control of exposure to sharp instruments is a high-profile issue. Estimates from the University of Virginia's Exposure Prevention Information Network (EPINet) surveillance system for 1996 placed the number of the percutaneous injuries to US health workers in that year at almost 600 000.[1] In the largest study of needlesticks to date based on nurse reports (as opposed to institutional surveillance), we reported a startlingly high rate of nearly 1 injury per nurse-year using data from a national nurse survey in 1991.[2] Because the potential consequences of hepatitis B and C and HIV and AIDS infection are so severe, the relatively low rates of seroconversion after percutaneous injuries—estimated at less than 0.5% for HIV—are not particularly reassuring.[3–5] Moreover, because the personal and professional consequences of needlestick injuries can be devastating even when they do not result in infections,[5] needlestick and related injuries remain a very serious occupational health concern for nurses and other health care workers.

The dominant perspective in the literature and in most agency guidelines is that the transmission of bloodborne pathogens from patients to health care workers is largely preventable through the use of universal precautions and special equipment (primarily systems that resheathe needles after use and needleless access devices). Exclusive reliance on these strategies is inadequate, however, for several reasons. First, the adoption of universal precautions to date has been far from universal. Studies have shown, for example, that nurse compliance with universal precautions is affected by the availability of protective equipment, the perceived commitment of management to safety, and perceptions regarding the interference of precautions with job performance.[6,7] Second, the adoption of needleless technology has been widespread, but it is unlikely that any technology can ever entirely remove the need for health professionals to handle bare needles and sharps. Third, awareness is increasing that needlestick accidents, like medical errors, complications, and other reportable incidents in hospitals, may be related to organizational factors such as staffing and the nurse practice environment as well as staff education and the types of equipment used.

Although many aspects of sharps injuries and body fluid exposures have been extensively studied, Hanrahan and Reutter[8] noted in their review of the literature that an organizational perspective on this issue is needed. To our knowledge, little research has been conducted to determine what factors produce variations in needlestick injury rates across hospitals or hospital units and whether nurse staffing and organizational climate are important determinants. Examining the organizational context of needlestick injuries is particularly timely, given recent state and national initiatives to reduce bloodborne pathogen exposures by requiring the use of specific types of devices in hospitals and separate broader state initiatives mandating minimum staffing levels in hospitals.

In our previous study of AIDS care provided in 20 hospitals across the United States, 1990–1991, we estimated the frequency of needlestick injuries to hospital nurses based on data from various sources.[2] In addition to retrospective reports from surveyed nurses regarding the number of times they were injured with a blood-contaminated needle in the prior month, we asked the same nurses to report needlesticks at the end of every shift they worked for 30 days (i.e., prospectively). On the basis of the prospective shift-based reports, we estimated that the rate of injuries to staff nurses was 0.8 per nurse per year. Prospective and retrospective rates were statistically indistinguishable. Our data also showed that only about 1 in 4 needlestick injuries were reported to hospital authorities. We also found that nurses who reported recapping needles were at heightened risk for injury and that nurses working in magnet hospitals (3 of the 20 hospitals were known for having an especially positive working climate for nurses[9]) were at significantly reduced risk for injury. The results reported in this article extend the work of that study by exploring how risk factors associated with needlestick injuries and the relative frequency of needlestick injuries among hospital nurses are related to the staffing levels and organizational cli-

mates on the hospital units on which nurses work.

METHODS

Sample

We analyzed data from a survey conducted in 1991 of nurses working on 40 inpatient units in 20 general hospitals located in 11 cities with high AIDS prevalence: New York, NY; Baltimore, Md; Boston, Mass; Chicago, Ill; Miami, Fla; Dallas, Tex; New Orleans, La; Atlanta, Ga; Philadelphia, Pa; San Francisco, Calif; and Los Angeles, Calif.[10,11] In the parent study of hospital organization and inpatient AIDS care, a purposive sample of 10 hospitals was chosen from a national master list of institutions that have specialized AIDS units with at least 10 beds. A matched group of 10 hospitals in the same geographic areas but without specialized AIDS units was drawn on the basis of characteristics such as bed size, governance, and clientele served. Details are discussed in an earlier publication.[10]

In addition to instruments measuring the working climate in the study hospitals and hospital units, the confidential, self-administered questionnaire filled out by the nurses included items dealing with exposures to sharps over the previous 1-month and 1-year periods.[2] Of the 865 staff nurses permanently assigned to the study units who received the questionnaire, 762 returned it, and 732 questionnaires were usable. Additional prospective data dealing with exposures to sharps and near-miss injuries were collected from all nurses working on the study units (regular staff nurses and temporary nurses) at the end of each shift over two 1-month periods in late 1990 and early 1991. Reports were obtained for 12 349 (86%) of the total 14 379 shifts worked by 960 regular and temporary staff nurses. Because the retrospective survey was conducted 2 months before the prospective portion of the study began, the periods of time and the injuries that occurred during the prospective and retrospective data collection did not overlap.

Measures

Exposures to contaminated sharps. The nurses who completed our retrospective survey were asked whether they had ever been stuck with a needle or sharp object contaminated with blood. Those who responded affirmatively were then asked how many times this had occurred and how many of the incidents had occurred in the past month.

The prospective portion of our data collection involved the use of booklets containing a sufficient number of reporting forms (coupons) for a month of shifts. These coupon booklets were distributed to staff and nonstaff nurses (both registered nurses and licensed practical nurses) on each of the study units. One coupon was filled out by each nurse at the end of each shift worked and placed in a secure box on the study unit. Nurses indicated on each coupon whether they had incurred a needle or sharp injury, defined as "a puncture with a needle or sharp instrument that is contaminated with blood," and whether they had had a "near-miss with a used needle or sharp" on that shift.

Staffing data. The number of full-time-equivalent registered nurse positions and the average daily patient census on each of the units for each day of the first month of the study period were determined from administrative data provided by the managers on the nursing units, including payroll sheets and patient assignment worksheets. Ratios of registered nurse positions to average daily patient census on each unit were calculated, cross-checked, and used in the analyses later in this article. In the results presented here, lower-staffed hospital units had registered nurse–to–average daily patient census levels reflecting ratios of approximately 1 nurse for every 10 or more patients on average.

Resource adequacy and nurse manager leadership. Resource adequacy and nurse manager leadership were drawn from the Revised Nursing Work Index, a battery of items that gauge nurses' perceptions of the presence of selected organizational characteristics in their work setting. Details about the development of this tool, its psychometric properties, and its validation in successive studies by our team are available in another recent publication.[12] The Revised Nursing Work Index contains 49 items that asks nurses to indicate, on a 4-point scale from "strongly agree" to "strongly disagree," whether various features are present in their practice setting.

Three theoretically derived summary measures were initially constructed from these items, and 6 empirically derived subscales were subsequently isolated with factor analytic techniques.[13] The resource availability and nurse manager leadership subscales (r= 0.63) used in the current analyses deal with unit-level organizational support for nursing practice and were therefore believed to be the most plausible correlates of needlestick risk. Other subscales address factors such as the influence of senior nurse executives and the quality of nurse–physician relationships. Resource adequacy was derived from 4 items that dealt with nurses' perceptions of whether staffing was sufficient to accomplish the work to be done and to provide quality patient care and whether they had enough time and opportunity to discuss patient care problems with other nurses. Nurse manager leadership was derived from 5 items that dealt with perceptions of the nursing unit manager's leadership and support of nurses' initiative and decisionmaking. Cronbach αs for these 2 subscales in the current data were .83 and .81, respectively. Scores on these subscales provided by each nurse on a given nursing unit were considered independent judgments or evaluations of that unit's organizational climate. Mean subscale scores for all of the nurses on the same units were calculated and used in the analyses described later in this article as aggregate indicators.[12]

To avoid confounding our measures of these 2 organizational characteristics with our measures of needlestick exposures (the occurrence of which influenced nurses' perceptions of resource adequacy and nurse leadership), we calculated all mean unit scores both with and without the evaluations of the nurses who had experienced an injury. Although the results of our analyses were the same regardless of whether the assessments of injured nurses were included in estimating these characteristics, we present the more conservative results (i.e., excluding the evaluations of nurses who were injured).

Emotional exhaustion. The emotional exhaustion subscale of the Maslach Burnout Inventory[14] measures the extent to which nurses feel emotionally overwhelmed by their work. In the current data, this subscale had a Cronbach α coefficient of .89. In our re-

search, we have found that this is a psychometrically valid index that, when aggregated to the level of nursing units, gauges the extent to which working conditions of various types have led to a generalized sense of frustration, strain, and weariness among a particular unit's nursing staff. As in the case of the organizational climate measures, we calculated mean scores for each unit with and without data from the small number of nurses who were injured. There were no differences in the results obtained in the analyses with either approach.

Risk factors. Our survey instrument also asked nurses a series of questions about how often they recapped used needles when they cared for patients with known and unknown HIV status (with responses ranging across 4-point scales from "never" to "always"). A further series of questions asked nurses whether certain factors were present on their units that created a significant risk of exposure to bloodborne infections, including carelessness and inexperience of other staff and uncooperativeness of patients. Last, nurses were asked to estimate, on a 4-point scale ranging from "not very good" to "excellent", how good a job they thought their hospital had done in providing them with adequate knowledge about AIDS and with the supplies and equipment needed to protect themselves.

Data Analysis

We first examined whether variation across hospital units in staffing and organizational climate was a significant predictor of nurses' reports of the presence of specific risk factors associated with needlestick injuries on their units. We then estimated the effects of unit staffing levels and organizational characteristics on the odds of nurses experiencing needlestick injuries or near misses with a sharp over the prospective and retrospective surveillance periods. The organizational climate and staffing measures were examined both as continuous variables and as dichotomous variables to test the possibility that nurses working on units where conditions were poorest experienced needlesticks disproportionately. In the analyses in which these variables were dichotomized, nurses from the 10 units of the 40 that had the lowest levels of resource adequacy, nurse leadership, and nurse-to-patient ratios, and the highest levels of emotional exhaustion, were compared with nurses from the remaining units. Because little difference was seen in the results obtained with the dichotomous (bottom or top quartile vs all others) and continuous approaches, and because of the ease in interpreting the odds ratios computed for dichotomized variables, only the dichotomized results are presented here.

Because the nurses surveyed were grouped by units, their characteristics and their outcomes were not independent, and conventional logistic regression modeling would not have been an appropriate statistical technique. Consequently, in all cases, logistic regression modeling employing generalized estimating equations, with nursing unit as the clustering variable, was used to estimate odds ratios and 95% confidence intervals associated with the effects of the different factors on them.[15] In the case of the analyses of the prospective data, the number of shifts worked by each nurse was used as a control variable because the time at risk for injury in our analyses was directly related to the number of shifts that nurses worked. All analyses were conducted with SAS (Version 6.12; SAS Institute Inc, Cary, NC).

RESULTS

In the retrospective portion of our study, 34 (4.3%) of the 789 nurses who responded to the questionnaires reported a needlestick injury in the previous month. Of the 962 nurses who filled out at least 1 coupon during the prospective survey, 53 (5.5%) reported an injury involving a needlestick or sharp containing blood, and 228 (23.7%) reported an incident involving a near miss.

Table 1 shows that nurses working on hospital units with poorer work climates and lower staffing levels were substantially more likely to report the presence of risk factors associated with needlestick injuries. Although there is some variability in our estimates of the associations between the 4 organizational characteristics and 6 risk factors, 21 of 24 of the associations were significant, and most were substantial. Nurses on units with less adequate resources, lower staffing and less nurse leadership, and higher levels of emotional exhaustion were typically twice as likely to report the presence of risks due to staff carelessness and inexperience, patient uncooperativeness, frequent recapping of needles, and inadequate knowledge or supplies.

Table 2 shows that these same 4 organizational characteristics of hospital units also were related to the likelihood of incurring needlestick injuries and reporting incidents involving near misses. The likelihood of experiencing needlestick injuries in the month before our survey was 3 times higher, or nearly 3 times higher, among nurses on units with less adequate resources, less nurse leadership and support, lower staffing, and higher levels of emotional exhaustion. The likelihood of experiencing needlesticks and near misses during the period of our prospective (shift-to-shift) data collection was similarly affected by these adverse unit characteristics; also, odds ratios were somewhat smaller when the prospective data were used but often involved a doubling, or near doubling, of the odds of needlesticks and near misses. Some of these estimates were not significant at the 95% confidence level, but virtually all were nearly so, and the importance of staffing and organization in affecting these adverse events is indicated by the considerable consistency in the effects estimated across the 3 separate indicators of exposure.

DISCUSSION

The analyses presented here suggest that hospital nurses' exposures to bloodborne pathogens were associated with the organizational characteristics and staffing levels on the hospital units where they worked. Individual nurses' risks of sustaining percutaneous injuries with used sharps were related to aggregate-level characteristics of their hospital units such that working on units characterized by poor working climates was associated with increased risks of injuries and near misses.

The differences in the odds ratios presented in Tables 1 and 2 indicate that slightly different nursing units were designated as having high-risk conditions when different unit characteristics were used, and there were slight differences in our estimates of the effects of these characteristics on the likelihood of being injured. However, some of the units clearly had uniformly poor climates, whether

TABLE 1—Odds Ratios (ORs) and 95% Confidence Intervals (CIs) Estimating the Effects of Various Organizational Characteristics on the Odds of Nurse Reports of Different Needlestick Risk Factors on Hospital Units

	Low Nurse Staffing OR (95% CI)	Low Resource Adequacy OR (95% CI)	Low Nurse Manager Leadership OR (95% CI)	High Emotional Exhaustion OR (95% CI)
Significant risk due to staff carelessness	1.92 (1.31, 2.82)	1.88 (1.25, 2.83)	1.65 (1.05, 2.58)	2.16 (1.33, 3.50)
Significant risk due to staff inexperience	1.74 (1.04, 2.92)	2.18 (1.47, 3.24)	1.80 (1.11, 2.93)	2.06 (1.31, 3.23)
Significant risk due to patient uncooperativeness	2.11 (1.32, 3.38)	2.13 (1.34, 3.40)	1.71 (1.06, 2.76)	1.32 (0.74, 2.35)
Often recaps needles used on patients with unknown HIV status	2.40 (1.29, 4.46)	3.30 (2.08, 5.23)	2.16 (1.22, 3.16)	1.78 (0.87, 3.62)
Feels hospital has not done a good job providing knowledge to protect workers	2.76 (1.72, 4.42)	2.44 (1.48, 4.00)	1.54 (0.88, 2.71)	1.94 (1.15, 3.27)
Feels hospital has not done a good job providing supplies to protect workers	3.56 (2.18, 5.81)	2.94 (1.64, 5.17)	1.86 (1.07, 3.26)	1.86 (1.02, 3.37)

Note. Odds ratios were computed with generalized estimating equations to allow for clustering by hospital unit.

TABLE 2—Odds Ratios (ORs) and 95% Confidence Intervals (CIs) Estimating the Effects of Various Organizational Characteristics on the Odds of Nurses Sustaining Percutaneous Injuries and Incurring Near Misses

	Retrospectively Reported Needlesticks OR (95% CI)	Prospectively Reported Needlesticks OR (95% CI)	Prospectively Reported Near Misses OR (95% CI)
Low nurse staffing	3.03 (1.22, 7.51)	2.06 (1.00, 4.25)	1.95 (1.02, 3.73)
Low resource adequacy	2.69 (1.08, 6.70)	1.73 (0.82, 3.66)	2.04 (1.08, 3.88)
Low nurse manager leadership	2.84 (1.14, 7.08)	1.56 (0.70, 3.49)	1.89 (1.06, 3.40)
High emotional exhaustion	2.54 (0.90, 7.26)	2.08 (1.03, 4.19)	1.57 (0.80, 3.10)

Note. Odds ratios were computed with generalized estimating equations to allow for clustering by hospital unit. Estimates involving prospectively reported needlesticks and near misses were computed after controlling for the number of shifts worked (time at risk).

assessed by our survey-based measures of organizational climate and nurse burnout or measured with institutional reports of nurse staffing. Nurses did not experience needlestick and related injuries at random. Injuries and near misses were clustered in specific units and were significantly more common in units with poor working climates and less staffing.

Although needlestick injuries may be reduced by staff education and the use of safer equipment, managers and policymakers trying to alleviate this problem ultimately must address the effect of staffing levels and work environments on these injuries. Previous discussions of this problem have suggested that clinical nurse specialists and nurse managers are well positioned to influence compliance with safer practices by teaching and modeling appropriate behavior, as well as by helping staff to better evaluate the risks and benefits of their decisions.[16] Although this observation is undoubtedly accurate and provides some concrete guidelines for frontline managers and leaders, the research reported here suggests the need to consider the broader context of nursing care on the units and in hospitals where needlestick injuries occur. The resource adequacy and nurse manager leadership measures in this study not only are a reflection of managerial decisions by frontline nurse leaders but also are markers of the extent to which the top managers in hospitals pay attention to and invest in safe environments for staff and patients alike.

Because the hospitals and nurses in the primary study were sampled to evaluate the effect of specialized AIDS units on patient and nurse outcomes (and not to evaluate needlestick risk in hospital nurses in different settings) and because the sample of nurses in this study was rather small, the results of these analyses must be interpreted cautiously. The results point to a possible effect of staffing and organization on hospital nurses' needlestick risk, but the data presented here do not permit commentary on specific staffing levels that are potentially safe or unsafe or on the nature of the causal relationship involved, if there is one. We are replicating and extending these findings with more recent and detailed data in more representative samples of hospitals. Currently, we are examining survey data from nurses working in a second nationwide sample of 22 hospitals in 1998. Our most recent data come from surveys completed in 1999 as part of an ongoing international study of nursing organization and hospital outcomes in which 43 000 nurses representing all hospitals in Pennsylvania and 3 Canadian provinces, as well as a sizable number of institutions in the United Kingdom and Germany, provided reports similar to those analyzed here.[17]

The recent resurgence of interest in errors and accidents in health care settings heralded by the Institute of Medicine's 1999 report *To Err Is Human*[18] has been characterized by dismay regarding the apparent pervasiveness of quality problems in medical care but also by an optimism that the incidence of misadventures in health care can be reduced by designing better systems to prevent, detect, and minimize hazards. Although needlestick injuries are not medical errors in the strictest sense (as discussed in the Institute of Medicine's report, for instance), they are, like medical errors, adverse events that occur in medical settings, and they have been viewed by

clinicians and administrators and examined by researchers similarly. Because needlestick injuries may serve as a proxy for a broad range of safety and quality issues, understanding the organizational context in which they occur is potentially very important. Remedying problems with understaffing, inadequate administrative support, and poor morale in hospitals may turn out to be among the most important steps in building a safer health care system. ■

About the Authors

The authors are with the Center for Health Outcomes and Policy Research, University of Pennsylvania School of Nursing, Philadelphia. Douglas M. Sloane is also with the Life Cycle Institute and Department of Sociology, Catholic University of America, Washington, DC. Linda H. Aiken is also with the Department of Sociology, University of Pennsylvania, Philadelphia.

Requests for reprints should be sent to Sean P. Clarke, PhD, RN, Center for Health Outcomes and Policy Research, University of Pennsylvania School of Nursing, 420 Guardian Dr, Philadelphia, PA 19104-6096 (e-mail: sclarke@nursing.upenn.edu).

This article was accepted April 6, 2001.

Contributors

S.P. Clarke planned and performed the data analysis and wrote the article. D.M. Sloane assisted in the data analysis and contributed to the writing of the article. L.H. Aiken was the principal investigator on the original study, assisted in the planning of the analysis, and contributed to the writing of the article.

Acknowledgments

This study was supported in part by grant R01-NR02280, "Outcomes of Inpatient AIDS Care," from the National Institute of Nursing Research, National Institutes of Health, and by an institutional postdoctoral fellowship (T32-NR07104), National Institute of Nursing Research, held by Sean P. Clarke, PhD, RN, at the Center for Health Outcomes and Policy Research, University of Pennsylvania.

References

1. International Health Care Worker Safety Center. Annual number of occupational percutaneous injuries and mucocutaneous exposures to blood or potentially infective biological substances [revised June 15, 1998]. Available at: http://www.med.virginia.edu/medcntr/centers/epinet/estimates.html. Accessed January 7, 2001.

2. Aiken LH, Sloane DM, Klocinski JL. Hospital nurses' occupational exposure to blood: prospective, retrospective, and institutional reports. *Am J Public Health.* 1997;87:103–107.

3. Bell DM. Occupational risk of human immunodeficiency virus infection in healthcare workers: an overview. *Am J Med.* 1997;102(5B):9–15.

4. Ippolito G, Puro V, Heptonstall J, Jagger J, De Carli G, Petrosillo N. Occupational human immunodeficiency virus infection in health care workers: worldwide cases through September 1997. *Clin Infect Dis.* 1999;28:365–383.

5. *NIOSH Alert: Preventing Needlestick Injuries in Health Care Settings.* Cincinnati, Ohio: National Institute for Occupational Safety and Health; 1999. DHHS publication 2000-108.

6. Murphy LR, Gershon RM, DeJoy D. Stress and occupational exposure to HIV/AIDS. In: Cooper CL, ed. *Handbook of Stress, Medicine and Health.* Boca Raton, Fla: CRC Press; 1996:177–190.

7. Grosch JW, Gershon RRM, Murphy LR, DeJoy DM. Safety climate dimensions associated with occupational exposure to blood-borne pathogens in nurses. *Am J Ind Med.* 1999;1(suppl 1):122–124.

8. Hanrahan A, Reutter L. A critical review of the literature on sharps injuries: epidemiology, management of exposures and prevention. *J Adv Nurs.* 1997; 25:144–154.

9. Havens DS, Aiken LH. Shaping systems to promote desired outcomes: the Magnet Hospital Model. *J Nurs Adm.* 1999;29:14–20.

10. Aiken LH, Lake ET, Sochalski J, Sloane DM. Design of an outcomes study of the organization of hospital AIDS care. *Res Sociol Health Care.* 1997;14:3–26.

11. Aiken LH, Sloane DM. Effects of organizational innovations in AIDS care on burnout among urban hospital nurses. *Work Occup.* 1997;24:453–477.

12. Aiken LH, Patrician PA. Measuring organizational traits of hospitals: the Revised Nursing Work Index. *Nurs Res.* 2000;49:146–153.

13. Lake E, Sochalski J. Measuring the social organization of nurses' work. Paper presented at: 70th Annual Meeting of the Eastern Sociological Society; March 2000; Baltimore, Md.

14. Maslach C, Jackson SE. *Maslach Burnout Inventory Manual.* 2nd ed. Palo Alto, Calif: Consulting Psychologists Press; 1986.

15. Allison PD. *Logistic Regression Using the SAS System.* Cary, NC: SAS Institute Inc; 1999.

16. Sulzbach-Hoke LM. Risk taking by health care workers. *Clin Nurs Specialist.* 1996;10:30–37.

17. Sochalski J, Aiken LH. Accounting for variation in hospital outcomes: a cross-national study. *Health Aff (Millwood).* 1999;18:256–259.

18. Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err Is Human: Building a Safer Health Care System.* Washington, DC: National Academy of Sciences; 1999.

Prescription Noncompliance due to Cost Among Adults With Disabilities in the United States

Jae Kennedy, PhD, and Christopher Erb, BA

Objectives. This study estimated national prevalence rates of medication noncompliance due to cost and resulting health problems among adults with disabilities.

Methods. Analyses involved 25 805 respondents to the Disability Follow-Back Survey, a supplement to the 1994 and 1995 National Health Interview Surveys.

Results. Findings showed that about 1.3 million adults with disabilities did not take their medications as prescribed because of cost, and more than half reported health problems as a result. Severe disability, poor health, low income, lack of insurance, and a high number of prescriptions increased the odds of being noncompliant as a result of cost.

Conclusions. Prescription noncompliance due to cost is a serious problem for many adults with chronic disease or disability. Most would not be helped by any of the current proposals to expand Medicare drug coverage. (*Am J Public Health.* 2002;92:1120–1124)

Medicare prescription drug insurance is a recurrent focus of American health policy,[1] and a combination of rapidly escalating drug costs[2] and insurance industry trends[3,4] have again thrust the issue to center stage. One of the more compelling rationales offered for expanding drug coverage is that affordability problems have clinical as well as economic consequences; that is, patients who have difficulty paying for medications are less likely to take them and can suffer adverse health effects as a result of noncompliance.[5,6] Although this argument has intuitive appeal, no national data are available on cost-associated noncompliance, leading commentators to question both the scope of affordability problems and the remedies proposed to address them.[7] In the present study, we sought to illuminate a critical aspect of the policy debate by developing the first national prevalence estimates of prescription noncompliance due to cost and resulting health problems among adults with disabilities, a population known to be heavy users of health care,[8,9] including prescription drugs.[9–11]

Medicare recipients with drug coverage are more likely to fill their prescriptions than those without coverage.[12–14] Total and out-of-pocket drug costs are heavily skewed toward individuals with poor health or chronic conditions, even among recipients with drug coverage.[15] Noncompliance with prescription regimens is a widely recognized clinical problem,[16] particularly in the case of treatment of chronic illnesses such as hypertension,[17] and it has been identified as an important predictor of emergency room visits[18] and hospital admissions.[19,20] Numerous studies have linked rates of noncompliance to (1) sociodemographic factors, including age,[21–23] sex,[17] and race/ethnicity[24]; (2) socioeconomic factors, including insurance coverage[25] and out-of-pocket costs[18,19]; and (3) treatment factors, including type[26] and number of drugs prescribed[27] and complexity of drug regimen.[21] We examined the relative influences of these factors on self-reported noncompliance due to cost.

METHODS

Data Source

The Disability Supplement and the Disability Follow-Back Survey (DFS) are special supplements to the National Health Interview Survey (NHIS), a continuing probability survey of households representative of the civilian noninstitutionalized population of the United States.[28] The Disability Supplement was administered to all respondents at the same time they completed the 1994 and 1995 NHIS core surveys. The DFS was administered 6 to 18 months later to respondents who reported impairments, functional limitations, chronic conditions, or receipt of disability benefits in the core NHIS surveys or the Disability Supplement.[29] We used data from the adult supplement, which was administered to 25 805 respondents 18 years or older with disabilities (about 17% of the NHIS sample).

Adults selected for the DFS differed from the general population selected for the NHIS in predictable ways. They were older (according to weighted estimates, 35% of DFS adult respondents were 65 years or older, compared with 13% of NHIS adult respondents), had lower incomes (19% of DFS respondents had incomes at or below the poverty level, compared with 12% of NHIS respondents), and were in worse health (69% of DFS adult respondents rated their health as fair or poor, compared with 34% of NHIS respondents).

Data Analysis

We used a case–control design to examine risk factors associated with prescription noncompliance due to cost. We weighted all data so that they would be generalizable to the overall US population. SUDAAN statistical software was used to account for the clustered sample design of the NHIS and the lack of independence in the error terms.[30] Unadjusted and adjusted odds ratios (ORs) were calculated for demographic (age, sex, race/ethnicity), socioeconomic (income, health insurance coverage), and health and disability (self-assessed health status, severity of activity limitations, number of prescriptions) factors. Respondents who were not prescribed any medications and those who reported that they did not take their medications as prescribed for reasons other than cost were omitted from comparisons.

RESULTS

Almost 70% of the disabled adult population—about 28 million people—reported having been prescribed 1 or more medications

TABLE 1—Self-Reported Number of Prescriptions and Compliance Rates Among US Adults With Disabilities

	Estimated No. (%)
Medications prescribed	
None	12 161 (30.0)
1–2	12 400 (30.6)
3–5	10 846 (26.8)
6–9	3 961 (9.8)
≥10	1 139 (2.8)
Take medicine(s) as prescribed[a]	
All of the time	24 762 (86.8)
Most of the time	2 576 (9.0)
Some of the time	793 (2.8)
Rarely	210 (0.7)
Never	194 (0.7)

Note. Data are population estimates in 1000s derived from the National Center for Health Statistics.[29]
[a]About 2% of people who had one or more prescriptions (317 respondents) selected the "don't know" response and were omitted from subsequent analyses.

(Table 1), and more than 85% of this group indicated that they always used their medications as prescribed. However, an estimated 3.8 million adults reported that they did not always use their medications as prescribed. These respondents were asked to select 1 or more of 8 reasons for their noncompliance (Table 2). About 1.3 million adults with disabilities cited 1 or more concerns related to cost (i.e., they did not get their prescription filled, did not fill their prescription completely, did not refill their prescription, or used their medicine less often than prescribed because of cost). This subset of noncompliant respondents was the focus of all subsequent analyses.

Table 3 identifies specific factors associated with prescription noncompliance due to cost. Persons with incomes below the poverty level were at higher risk for cost-associated noncompliance than were persons with incomes above the poverty level (OR=1.6; 95% confidence interval [CI]=1.3, 2.0). Uninsured adults were nearly 4 times as likely to be noncompliant owing to cost as their counterparts with private insurance (OR=3.9; 95% CI=3.0, 5.1). Adults with private and public health insurance (i.e., supplemental Medicare coverage) exhibited relatively low rates of cost-associated noncompliance (OR=0.7; 95% CI=0.5, 1.0).

Individuals who described their health as fair or poor were more likely to be noncompliant than were those who rated their health as good, very good, or excellent (OR=1.4; 95% CI=1.1, 1.7). The relationship between severity of disability and cost-associated noncompliance appeared to be curvilinear, with the highest level of noncompliance among moderately impaired adults who were limited in, but did not require assistance with, 1 or more activities of daily living (OR=1.9; 95% CI=1.5, 2.5). Disabled adults who were prescribed 3 or more medications were more likely than those who were prescribed 1 or 2 medications to report cost-associated noncompliance (3–5 medications: OR=1.4; 95% CI=1.1, 1.7; 6 or more medications: OR=1.6; 95% CI=1.2, 2.1).

Sex and race/ethnicity appeared to be only modestly related to cost-associated noncompliance, but there was a strong negative relationship between age and noncompliance: younger adults (those aged 18–34 years) were nearly 10 times more likely to be noncompliant as a result of cost than were members of the oldest cohort (those 75 years or older; OR=0.1; 95% CI=0.1, 0.2). However, members of younger cohorts were also less likely to be prescribed medications.

To clarify this relationship, we plotted the number of adults with prescriptions and the proportions reporting cost-associated noncompliance according to age group. Figure 1 shows that prescription rates increased with age: only 35% of disabled adults aged 18 to 24 years were prescribed medications, in comparison with 86% of adults aged 65 years or older. Among disabled adults with prescriptions, rates of cost-associated noncompliance peaked at about 10% for those aged 25 to 44 years and declined sharply in older age cohorts. Only about 2% of adults aged 65 to 74 years reported cost-associated noncompliance, and this rate dropped to below 1% among adults 75 years or older.

All noncompliant respondents were asked whether they had experienced any adverse health consequences (Table 4). Among those who were noncompliant owing to cost, more than half identified 1 or more resulting health problems. The most common problems involved exacerbation of conditions or symptoms; for example, nearly 350 000 people reported pain or discomfort resulting from cost-associated noncompliance. A relatively small proportion of respondents reported that noncompliance problems led directly to additional health care use: an estimated 66 000 people had to visit a doctor's office or emergency room, and about 46 000 had to be hospitalized.

TABLE 2—Reasons Given by Adults With Disabilities for Noncompliance

	Estimated No. (%)
Affordability	1280 (33.9)
Did not refill when ran out owing to cost	898 (23.7)
Use less often than prescribed to stretch out owing to cost	853 (22.6)
Did not get when first prescribed owing to cost	767 (20.2)
Did not get entire prescription filled owing to cost	741 (19.5)
Other	2483 (65.8)
Sometimes forget to use	1789 (47.4)
Don't use as prescribed because of side effects	926 (24.5)
Don't use because of perceived lack of need	826 (22.0)
Cannot pick up from drug store or get delivered	125 (3.3)
Total noncompliant[a]	3798 (100)

Note. Data are population estimates in 1000s derived from the National Center for Health Statistics.[29]
[a]Currently prescribed one or more medications and does not always take as prescribed.

DISCUSSION

Our study showed that about 1.3 million adults with disabilities reported that the cost

TABLE 3—Factors Associated With Prescription Noncompliance due to Cost Among Adults With Disabilities

	Estimated No.[a] (n = 25 836)	Noncompliant due to Cost, %	Unadjusted		Adjusted	
			OR	95% CI	OR	95% CI
Age, y						
18–34	2 787	20.8	1.00		1.00	
35–44	3 342	10.8	0.99	0.77, 1.27	0.95	0.73, 1.25
45–54	4 022	7.1	0.62	0.49, 0.80	0.59	0.45, 0.78
55–64	4 413	3.6	0.30	0.23, 0.40	0.28	0.20, 0.38
65–74	5 733	2.1	0.17	0.13, 0.24	0.21	0.14, 0.32
≥75	5 539	0.9	0.08	0.05, 0.12	0.10	0.07, 0.16
Sex						
Male	10 118	4.2	1.00		1.00	
Female	15 718	5.4	1.30	1.10, 1.53	1.20	0.99, 1.44
Race						
White	20 481	4.9	1.00		1.00	
Hispanic	1 695	3.7	0.75	0.57, 0.99	0.45	0.32, 0.64
Black	2 999	6.4	1.32	1.05, 1.67	0.87	0.65, 1.15
Other	662	4.1	0.84	0.47, 1.50	0.80	0.43, 1.47
Income at or below poverty level[b]						
No	19 137	4.0	1.00		1.00	
Yes	4 167	8.8	2.34	1.94, 2.82	1.59	1.26, 2.02
Health insurance[c]						
Private only	6 993	5.2	1.00		1.00	
Public only	7 787	5.3	1.02	0.83, 1.26	1.02	0.77, 1.34
Mix of private and public	9 409	1.6	0.30	0.23, 0.39	0.70	0.49, 0.98
Uninsured	1 548	21.3	4.90	3.87, 6.20	3.90	3.02, 5.05
Health status						
Excellent–good	13 863	7.9	1.00		1.00	
Fair–poor	11 782	6.0	1.53	1.31, 1.80	1.39	1.13, 1.72
Severity of disability[d]						
No activity limitations	11 282	4.2	1.00		1.00	
Activity limit only	3 416	7.7	1.91	1.54, 2.37	1.93	1.48, 2.50
Assistance needed	11 138	4.9	1.18	0.98, 1.43	1.28	1.01, 1.63
No. of prescriptions						
1–2	11 112	4.9	1.00		1.00	
3–5	9 970	4.8	0.98	0.82, 1.17	1.37	1.11, 1.71
≥6	4 751	10.6	1.03	0.81, 1.32	1.62	1.19, 2.19

Note. Data are population estimates in 1000s derived from the National Center for Health Statistics (NCHS).[29] OR = odds ratio; CI = confidence interval.

[a]Total includes adults with disabilities who always took their medications as prescribed, plus adults with disabilities who did not take their medications as prescribed owing to cost concerns (respondents who were noncompliant solely for other reasons, as well as those who were not prescribed any medications, were omitted from this analysis).

[b]Family income data were missing for an estimated 9.8% of respondents and deleted in the multivariate model.

[c]Health insurance status based on NCHS recode; ambiguous categories ("private/unknown if public" and "public/unknown if private") coded as "private only" or "public only."

[d]Activity limitation was assessed in 15 domains: bathing, dressing, eating, toileting, transferring, walking, getting outside, light housework, heavy housework, transportation, meal preparation, shopping for groceries, managing medications, managing money, and using the telephone.

of the medicine(s) they were prescribed was so high that they could not afford to fill or refill their prescriptions or to take their medication as prescribed. More than half of this group identified 1 or more potentially serious and costly health problems that they attributed to noncompliance. These prevalence figures are impressive; for several reasons, however, they probably underestimate the true scope of drug affordability problems among people with disabilities.

First, our data did not allow us to estimate the number of people who take their medications as prescribed but do so at great personal cost. For some people with disabilities or chronic illnesses, limited incomes force a monthly choice between medication and food.[31] Second, although only recently released to the research community, the surveys we examined are somewhat dated. Drug costs have skyrocketed in the period since the data were collected,[2] potentially threatening the health and economic security of many more adults with and without disabilities.[6,32,33] Third, all compliance data were self-reported and thus subject to biases associated with such survey methods.[34] Indeed, underreporting of noncompliance is such a widely recognized problem[26] that many researchers use independent verification strategies such as pill counts[35] and electronic monitoring.[36]

Despite these limitations, our analysis raises some provocative research and policy questions. As might be expected, income and insurance status were strong predictors of noncompliance due to cost. The magnitude of the insurance differences, however, was striking; after other risk factors had been controlled, disabled adults without insurance were nearly 4 times more likely than those with private insurance to report medication noncompliance due to cost.

The finding that people who were in poorer health or who took more medications were also at higher risk of cost-associated noncompliance was consistent with previous research. However, the curvilinear relationship found between severity of disability and noncompliance due to cost merits further investigation. The relatively lower rate of noncompliance among Hispanic adults with disabilities was unanticipated, and additional research is clearly needed to verify this relationship.

Our most remarkable finding from a public policy perspective was that cost-associated noncompliance was concentrated primarily in younger cohorts. This result seems to contradict much of the recent political commentary on drug affordability, although other studies

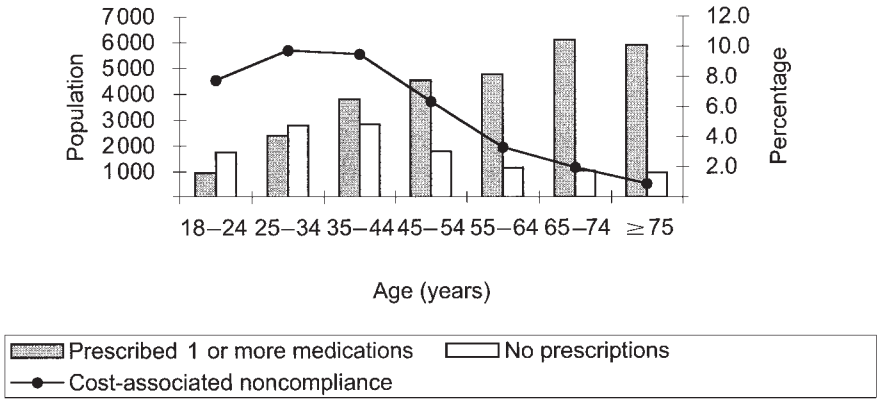


FIGURE 1—Number of disabled adults who had prescriptions and proportions noncompliant owing to cost, by age group.

TABLE 4—Reported Health Problems Attributed to Medication Noncompliance due to Cost Among Adults With Disabilities

	Estimated No. (%)
Experienced one or more problems owing to noncompliance	672 (52.5)
Pain or discomfort	349 (27.2)
Condition for which medicine prescribed got worse	267 (20.9)
Dizziness or fainting	159 (12.4)
Change in blood pressure, breathing, or other vital signs	154 (12.0)
Disorientation	93 (7.3)
Had to go to the doctor or emergency room	66 (5.2)
Other condition(s) got worse	64 (5.0)
Had to be admitted to the hospital	46 (3.6)
Overdose or withdrawal	37 (2.9)
Drug reaction	35 (2.8)
Other	152 (11.9)
Total noncompliant owing to cost	1280 (100)

Note. Data are population estimates in 1000s derived from the National Center for Health Statistics.[29]

have also revealed a negative relationship between age and compliance.[21,22] Additional research is needed, however, before we would concur with the conclusion of Park et al. that, in terms of medication compliance, "older is wiser."[23]

Indeed, at least for the population of adults with disabilities, the more appropriate adage might be "younger is poorer" (or, at least, "younger is less likely to be insured"). Most of the 1.3 million disabled adults identified in this study would not be helped by any of the current proposals to expand Medicare drug coverage, because only 27% received Medicare. If this population were included in the policy debate and ways were found to increase prescription drug coverage for all adults with chronic illnesses and disabilities, much of the exacerbation of symptoms and conditions found in this study—and many of the associated health care expenditures—could be avoided. ■

About the Authors

At the time of this study, Jae Kennedy was with the Department of Community Health, University of Illinois at Urbana–Champaign. Christopher Erb is with the Medical Scholars Program, University of Illinois at Urbana–Champaign.

Requests for reprints should be sent to Jae Kennedy, PhD, Department of Health Policy and Administration, Washington State University at Spokane, 310 N Riverpoint Blvd, Box H, Spokane, WA 99202-1675 (e-mail: jjkenned@mail.wsu.edu).

This article was accepted June 7, 2001.

Note. The analyses, interpretations, and conclusions are those of the authors and do not necessarily reflect the views of the National Center for Health Statistics, the National Institute of Disability and Rehabilitation Research, Washington State University, or the University of Illinois.

Contributors

J. Kennedy designed the study, conducted the analyses, and wrote the article. C. Erb conducted the literature review, collaborated in analysis design and interpretation, and contributed to the writing of the article.

Acknowledgments

Preparation of this article was supported in part by a fellowship from the National Institute of Disability and Rehabilitation Research (grant H133F990001) and by a grant from the Mary Jane Neer Fund at the University of Illinois. The Disability Follow-Back Survey adult data were provided by the National Center for Health Statistics, and prevalence data on prescription drug use among adults with disabilities were first presented by John Hough (of the Division of Birth Defects, Child Development, Disability and Health, Centers for Disease Control and Prevention) at the 1999 Annual Meeting of the American Public Health Association.

References

1. Long SH. Prescription drugs and the elderly: issues and options. *Health Aff.* 1994;13(2):157–174.

2. *Hard to Swallow: Rising Prices for America's Seniors.* Washington, DC: Families USA; 1999.

3. Moran DW. Prescription drugs and managed care: can 'free-market detente' hold? *Health Aff.* 2000;19(2):63–77.

4. McClellan M, Spatz ID, Carney S. Designing a Medicare prescription drug benefit: issues, obstacles, and opportunities. *Health Aff.* 2000;19(2):26–41.

5. *Disturbing Truths and Dangerous Trends: The Facts About Medicare Beneficiaries and Prescription Drug Coverage.* Washington, DC: National Economic Council; 1999.

6. Soumerai SB, Ross-Degnan D. Inadequate prescription-drug coverage for Medicare enrollees—a call to action. *N Engl J Med.* 1999;340:722–728.

7. Frogue J. *How to Provide Prescription Drug Coverage Under Medicare.* Washington, DC: Heritage Foundation; 1999.

8. Chrischilles EA, Foley DJ, Wallace RB, et al. Use of medications by persons 65 and over: data from the established populations for epidemiologic studies of the elderly. *J Gerontol.* 1992;47:M137–M144.

9. LaPlante MP, Rice DP, Wenger BL. *Medical Care Use, Health Insurance, and Disability in the United States.* Washington, DC: National Institute of Disability and Rehabilitation Research; 1995.

10. Max W, Rice DP, Trupin L. *Medical Expenditures for People With Disabilities.* Washington, DC: National Institute of Disability and Rehabilitation Research; 1995.

11. Mueller C, Schur C, O'Connell J. Prescription drug

spending: the impact of age and chronic disease status. *Am J Public Health.* 1997;87:1626–1629.

12. Davis M, Poisal J, Chulis G, Zarabozo C, Cooper B. Prescription drug coverage, utilization, and spending among Medicare beneficiaries. *Health Aff.* 1999;18(1): 231–243.

13. Poisal JA, Murray LA, Chulis GS, Cooper BS. Prescription drug coverage and spending for Medicare beneficiaries. *Health Care Financ Rev.* 1999;20(3): 15–27.

14. Stuart B, Grana J. Ability to pay and the decision to medicate. *Med Care.* 1998;36:202–211.

15. Steinberg EP, Gutierrez B, Momani A, Boscarino JA, Neuman P, Deverka P. Beyond survey data: a claims-based analysis of drug use and spending by the elderly. *Health Aff.* 2000;19(2):198–211.

16. Sackett DL, Haynes RB, eds. *Compliance With Therapeutic Regimens.* Baltimore, Md: Johns Hopkins University Press; 1976.

17. Costa FV. Compliance with antihypertensive treatment. *Clin Exp Hypertens.* 1996;18:463–472.

18. Olshaker JS, Barish RA, Naradzay JF, Jerrard DA, Safir E, Campbell L. Prescription noncompliance: contribution to emergency department visits and cost. *J Emerg Med.* 1999;17:909–912.

19. Col N, Fanale JE, Kronholm P. The role of medication noncompliance and adverse drug reactions in hospitalizations of the elderly. *Arch Intern Med.* 1990; 150:841–845.

20. Maronde RF, Chan LS, Larsen FJ, Strandberg LR, Laventurier MF, Sullivan SR. Underutilization of antihypertensive drugs and associated hospitalization. *Med Care.* 1989;27:1159–1166.

21. Bailey JE, Lee MD, Somes GW, Graham RL. Risk factors for antihypertensive medication refill failure by patients under Medicaid managed care. *Clin Ther.* 1996;18:1252–1262.

22. Faulkner DL, Young C, Hutchins D, McCollam JS. Patient noncompliance with hormone replacement therapy: a nationwide estimate using a large prescription claims database. *Menopause.* 1998;5:226–229.

23. Park DC, Hertzog C, Leventhal H, et al. Medication adherence in rheumatoid arthritis patients: older is wiser. *J Am Geriatr Soc.* 1999;47:172–183.

24. Monane M, Bohn RL, Gurwitz JH, Glynn RJ, Levin R, Avorn J. Compliance with antihypertensive therapy among elderly Medicaid enrollees: the roles of age, gender, and race. *Am J Public Health.* 1996;86: 1805–1808.

25. Bazargan M, Barbre AR, Hamm V. Failure to have prescriptions filled among black elderly. *J Aging Health.* 1993;5:264–282.

26. McElnay JC, McCallion CR, al-Deagi F, Scott M. Self-reported medication non-compliance in the elderly. *Eur J Clin Pharmacol.* 1997;53:171–178.

27. Botelho RJ, Dudrak RD. Home assessment of adherence to long-term medication in the elderly. *J Fam Pract.* 1992;35:61–65.

28. Massey JT, Moore TF, Parsons VL, Tadros W. Design and estimation for the National Health Survey, 1985–94. *Vital Health Stat 2.* 1989;No. 110.

29. *Data File Documentation, National Health Interview Survey on Disability, Phase 1 and Phase 2, 1994.* Hyattsville, Md: National Center for Health Statistics; 1998. CD-ROM series 10, No. 9.

30. *Software for the Statistical Analysis of Correlated Data (SUDAAN), Version 7.5.* Research Triangle Park, NC: Research Triangle Institute; 1998.

31. Chubon SJ, Schulz RM, Lingle EW Jr, Coster-Schulz MA. Too many medications, too little money: how do patients cope? *Public Health Nurs.* 1994;11: 412–415.

32. Thornton D. By the numbers: the rising costs of prescription drugs. *Healthplan.* 1999;40(5):71–72.

33. Rogowski J, Lillard LA, Kington R. The financial burden of prescription drug use among elderly persons. *Gerontologist.* 1997;37:475–482.

34. Babbie E, Wagenaar T. *Practicing Social Research.* 3rd ed. Belmont, Calif: Wadsworth Publishing Co; 1983.

35. Larrat EP, Taubman AH, Willey C. Compliance-related problems in the ambulatory population. *Am Pharmacy.* 1990;30(2):18–23.

36. Park DC, Morrell RW, Frieske D, Kincaid D. Medication adherence behaviors in older adults: effects of external cognitive supports. *Psychol Aging.* 1992;7: 252–256.

Twenty Years of Public Health Research: Inclusion of Lesbian, Gay, Bisexual, and Transgender Populations

Ulrike Boehmer, PhD

Objectives. This study determined to what extent lesbian, gay, bisexual, and transgender (LGBT) populations have been studied over the past 20 years of public health research.

Methods. From MEDLINE English-language articles on human subjects published between 1980 and 1999, I identified articles that included LGBT individuals. The abstracts were analyzed with a coding procedure that categorized the content by topic, sexual orientation, and race/ethnicity.

Results. LGBT issues were addressed by 3777 articles, or 0.1% of all Medline articles; 61% of the articles were disease-specific, and 85% omitted reference to race/ethnicity. Research unrelated to sexually transmitted diseases addressed lesbians and gay men with similar frequency, whereas bisexual persons were less frequently considered, and the least amount of research focused on transgender individuals.

Conclusions. Findings supported that LGBT issues have been neglected by public health research and that research unrelated to sexually transmitted diseases is lacking. (*Am J Public Health.* 2002;92:1125–1130)

Despite increased interest in lesbian, gay, bisexual, and transgender (LGBT) populations' health as a legitimate focus of scientific study, public health research has not been systematically reviewed to identify the extent to which LGBT issues have been addressed. Historically, public health researchers have not recognized LGBT persons as a population with distinct health issues outside a framework of sexual deviance or sexually transmitted diseases (STDs). Before the 1980s, most studies that addressed LGBT populations focused on the etiology of homosexuality[1] or on mental disorders because homosexuality was classified as such until 1973.[2–4] In the 1980s, a new virus, later known as HIV, was initially recognized to occur among gay men.[5,6] Research on HIV and AIDS incorporated measures of same-sex behaviors to assess "risk behaviors."[7,8] In the 1990s, influenced by social movement advocacy, debate arose over whether lesbians' risk factors for developing breast cancer might be higher compared with those of heterosexual women.[1,9]

This focus on diseases that are of relevance to LGBT individuals developed into the recognition that LGBT populations are diverse communities with disparate health concerns. Hence, a systematic review of LGBT research also must identify gaps in knowledge regarding diverse sectors of LGBT populations in relation to behaviors, race/ethnicity, and socioeconomic position. This review is crucial to inform adequately the goals set for LGBT health in documents such as *Healthy People 2010: Understanding and Improving Health*[10] and its companion document on LGBT health, which addresses these populations more fully.[11]

Documents that strive to identify solutions to LGBT populations' health concerns, such as *Healthy People 2010* and the 1999 Institute of Medicine report on lesbian health,[1] regularly point to the lack of representative, population-based data on LGBT individuals. In the absence of adequate data, LGBT research is frequently hindered by methodological issues, such as defining, measuring, and sampling of LGBT populations.[1,2] The lack of uniform definitions and measures for LGBT persons, such as behavior, identity, and desire,[12,13] and of different sampling strategies severely limits the generalizability of study results.[14] Recognizing that existing LGBT studies may suffer from these limitations, one might wonder how often researchers have actually been able to collect data on LGBT populations and publish the results. Accordingly, the objective of this study was to determine the extent to which LGBT issues have been addressed in public health research and the contexts in which LGBT populations have been studied over the last 20 years.

METHODS

The National Library of Medicine contains 3 822 822 citations of articles based on studies with human subjects and published in English between 1980 and 1999. The goal was to determine how many of these studies specified the inclusion of LGBT individuals. With the same restrictions, a MEDLINE keyword search for the years 1980 to 1999 was conducted on Ovid on January 19, 2001. "Gay," "lesbian," "bisexual," "sexual orientation," "sexual identity," "bisexuality," "homosexuality, male," "homosexuality, female," "transgender," "transsexual," "transvestite," "cross-dresser," and "hermaphrodite" were used as key words. The abstracts of articles that fulfilled these requirements were downloaded and their content analyzed with a 3-step coding procedure:

1. The first coding step distinguished between disease-specific and non-disease-specific articles. The abstracts' contents were then summarized into broad areas that described a cluster of research topics.
2. The second coding step determined whether the abstract mentioned lesbian, gay, bisexual, or transgender individuals.
3. The third coding step focused on the race/ethnicity of the study population. Socioeconomic position of the study population was not coded, because data on this variable typically were either missing or poorly delineated.

To maximize the available data, all types of studies were considered, with the exception of letters that commented on research previously published. The first level of coding distinguished between articles that focused on specific medical conditions and those that did not. All articles that related their content to a specific disease were coded as "disease focused," even if their subject matter was psychosocial rather than biomedical. The implication of this coding rule was that caregiving was coded as disease-specific if the investigation was in the context of taking care of people with HIV or AIDS. An article on the caregiving burden of LGBT persons was non-disease-specific if it was not investigated in the context of a specific disease.

Mental disorders listed in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV),*[15] were coded as disease-specific, with 1 exception. Diverging from the *DSM-IV* classification of "gender identity disorder," articles on transgender identity were coded as non-disease-specific research on identity rather than a mental disorder. Continued examination of the data, reevaluation of codes, and refinement of coding rules led to the development of themes or broad areas that grouped non-disease-specific studies into clusters of research topics.

The second level of coding on sexual orientation and transgender identity led to the exclusion of articles that did not mention LGBT individuals in their abstracts. Because no agreed-on measure of sexual orientation exists and researchers often do not define sexual orientation,[14] coding of sexual orientation was not limited to specific definitions such as behavior, identity, or desire. Similarly, "transgender" had been added only recently to the lesbian, gay, and bisexual terminology, and the definition of transgender identity is evolving.[2,16] Over the years, different terms have been used; therefore, abstracts that mentioned transsexuals, transvestites, cross-dressers, or hermaphrodites were coded as "transgender."

Other coding rules evolved empirically. Abstracts referring to homosexual persons or homosexuality were coded as "gay" and "lesbian." Those that mentioned sexual orientation without specifying gender were coded as "lesbian," "gay," "bisexual women," and "bisexual men." Gay patients were coded as "gay." Furthermore, articles in which LGBT populations were only marginally addressed were included. This rule led, for instance, to the inclusion of articles that focused on heterosexual women but mentioned bisexual men in the context of heterosexual women's sexual risk behavior.

Because many abstracts provided incomplete descriptions of the study population with respect to racial/ethnic categories, coding of race/ethnicity was limited to either "race omitted" or "race reported." Research mostly lacked a definition of race as a biological, socioeconomic, or sociocultural construct.[17] Thus, the code "race reported" was not restricted to a specific definition of race/ethnicity and did not distinguish between self-reported and observed race/ethnicity.

RESULTS

The search of the MEDLINE database generated 4537 citations published between 1980 and 1999, which represented 0.1% of all articles recorded in MEDLINE. From the initial 4537 citations, 760 were excluded on the basis of the following criteria: (1) an abstract was not available, despite extensive efforts by the author to locate and review articles at various libraries; (2) they were letters in response to previously published articles; and (3) abstracts did not mention LGBT populations. The remaining abstracts of 3777 articles (0.1% of MEDLINE articles) were suitable for analysis with the 3-step coding procedure.

Table 1 provides summary information on the 3 categories by which the 3777 articles were coded. The majority of articles—2285 (61%)—were coded as disease-specific. Most articles focused on gay men (80%), followed by bisexual men (39%). Lesbians, bisexual women, and transgender persons, as a group, were included in 46% of the articles. Most articles (85%) omitted the racial/ethnic background of LGBT individuals.

Table 2 lists the topics of the 3777 articles organized into broad research areas. The vast majority of the 2285 disease-specific articles focused on STDs, particularly HIV and AIDS. Disease-specific research also occurred in the broad areas mental disorders and cancer. Articles on mental disorders focused mostly on addiction, followed by various psychiatric disorders, including schizophrenia and bipolar and mood disorders. Fifteen articles on sexual dysfunction included erectile dysfunction, dyspareunia, and compulsive sexual behavior. A cluster of 36 articles focused on various cancers, whereas 23 articles covered various diseases, subdivided into 8 infectious and 15 noninfectious diseases.

The content analysis of 1492 non-disease-specific articles was summarized into 7 broad areas. Most articles addressed identity, sexual behavior, and desire of LGBT populations. These articles considered the development of LGBT identity, including sexual choices and behaviors, relationships, coming-out issues, and aging, as well as body image and physiology, including gender-related surgery. The second largest research area consisted of 312 articles that focused on the etiology of sexual orientation and transgender identity, including research on the sexual orientation of siblings and twins and comparisons of hormonal levels between LGBT individuals and heterosexual individuals. The 207 articles on the health care of LGBT persons addressed providers' attitudes and cultural competency; the

TABLE 1—Distribution of Articles That Considered Lesbian, Gay, Bisexual, and Transgender Populations (N = 3777), by Content, Sexual Orientation, and Race/Ethnicity

	n	%
Type of content		
Disease	2285	60.5
Nondisease	1492	39.5
Sexual orientation[a]		
Lesbian	1043	27.6
Gay	3027	80.1
Bisexual women	352	9.3
Bisexual men	1486	39.3
Transgender	346	9.2
Race/ethnicity		
Omitted	3212	85.0
Included	565	15.0

[a]More than 1 sexual orientation can be addressed in a single article. Therefore, the number of lesbian, gay, bisexual, and transgender articles adds up to more than 3777 and more than 100%.

TABLE 2—Articles That Considered Lesbian, Gay, Bisexual, and Transgender Populations, by Content Areas (N = 3777; 100%)

Disease-Specific Content Areas		Non-Disease-Specific Content Areas	
Sexually transmitted diseases	2108 (55.8%)	Identity, sexual behavior, or desire	490 (13.0%)
AIDS, including opportunistic infections	1958	Identity, including gender or sexual identity	245
Sexually transmitted diseases other than AIDS	150	Relationship	66
Mental disorders	118 (3.1%)	Sexual behavior	51
Addiction, alcohol or drug	53	Coming out	39
Mental illness, including personality disorder	32	Aging	21
Sexual dysfunction	15	Physiology and body image	68
Pedophilia	10	Etiology	312 (8.3%)
Eating disorder	8	Neuroscience or endocrinology	141
Cancer	36 (1.0%)	Genetics and birth order	94
Anal cancer	16	History of homosexuality	47
Breast cancer	6	Various etiologic theories	30
Cervical cancer	3	Health care	207 (5.5%)
Ovarian cancer	3	Provider attitudes or patient-provider	126
Kaposi sarcoma (not AIDS related)	2	Health needs and utilization	81
Squamous cell cancer	2	Family	87 (2.3%)
Unspecified cancer	2	Parenting	61
Sertoli cell tumor	1	Reproduction	26
Nephroblastoma	1	Attitudes	75 (2.0%)
Various diseases	23 (0.6%)	Measuring attitudes toward lesbian, gay, bisexual, and transgender populations	29
Infectious diseases	8	Attitudes in educational institutions	15
Viral or bacterial infections	3	Images or stereotypes of lesbian, gay, bisexual, and transgender persons	18
Fungal infection	1	Change in attitudes toward lesbian, gay, bisexual, and transgender persons	7
Gastrointestinal diseases	4	Attitudes held by lesbian, gay, bisexual, and transgender individuals	6
Noninfectious diseases	15	Risk factors	70 (1.9%)
Heart disease	3	Risky sexual behavior	26
Menstrual disorders	2	Risky health behaviors	15
Respiratory disease	2	Suicide	17
Immune disorders	2	Prostitution	8
Osteoporosis	1	Stress due to minority status	4
Diabetes	1	Violence	52 (1.4%)
Breast disease	1	Assault	20
Benign tumor	1	Sexual abuse	16
Ovarian disease	1	Domestic violence	13
Urethritis	1	Various forms of violence	3
		Miscellaneous topics	199 (5.3%)
		Arts, literature	62
		Rights, legal aspects	27
		Community and lesbian, gay, bisexual, and transgender culture	27
		Methodology	21
		Freudian theory	21
		Religion or church	15
		Gay and lesbian studies	6
		Conversion therapy	5
		Career	4
		Social network	2
		Caregiving	2
		Articles on singular topic	7

TABLE 3—Distribution of Articles That Considered Lesbian, Gay, Bisexual, and Transgender Populations, by Specific Content Areas, Year of Publication, and Sexual Orientation[a]

	1980–1984					1985–1989					1990–1994					1995–1999				
	L	G	BW	BM	T	L	G	BW	BM	T	L	G	BW	BM	T	L	G	BW	BM	T
Sexually transmitted disease	2	42	2	24	0	13	262	11	213	1	37	715	24	475	10	64	955	47	453	9
Mental disorder	7	15	0	2	4	5	15	2	5	0	19	21	2	9	2	25	28	7	11	2
Cancer	0	0	0	0	4	0	3	0	3	2	0	3	0	2	3	9	11	4	9	0
Identity	33	47	10	9	33	51	60	11	24	35	59	72	16	18	40	85	97	28	30	32
Etiology	16	19	4	5	25	19	19	3	4	32	34	44	16	15	30	70	93	31	31	39
Health care	11	11	1	1	1	21	18	3	2	2	52	38	11	11	4	86	72	46	42	7
Family	10	5	0	0	0	16	9	1	0	1	10	6	0	0	0	38	13	3	3	1
Attitudes	4	6	0	0	1	5	6	1	1	0	18	17	2	2	0	36	37	8	7	0
Risk factors	0	1	0	0	1	1	9	0	2	4	6	17	4	8	0	17	34	9	13	3
Violence	0	1	0	0	0	2	4	0	0	0	11	14	3	5	1	19	26	7	8	0

Note. L = lesbian; G = gay; BW = bisexual women; BM = bisexual men; T = transgender.
[a]More than 1 sexual orientation can be addressed in a single article. Therefore, the number of lesbian, gay, bisexual, and transgender articles adds up to more than 3777.

patient–provider relationship; health needs of LGBT individuals; and their experiences with the delivery of health services. Family issues included parenting of and by LGBT individuals; alternative insemination; and adoption. "Attitudes" as a topic included measuring attitudes toward LGBT populations; attitudes in educational institutions; images and stereotypes of LGBT individuals; identification of people as LGBT; change in attitudes toward LGBT persons; and attitudes held by LGBT persons. Seventy articles focused on various risk factors, including risky sexual behaviors such as the lack of condom use. "Risky health behaviors" included research on smoking. Suicide, prostitution, and stress due to minority status among LGBT people are other topics among the category "risk factors." "Violence" was another area of research, including hate crimes, rape, incest, and domestic violence; 52 articles addressed violence.

Table 3 provides more detail on the 10 content areas of research, separating them by sexual orientation and years of publication. Although the number of articles in all areas increased with time, articles on STDs showed the biggest increase and focused predominantly on gay and bisexual men. Initially, research on mental disorders and cancer focused more often on gay men than on lesbians, but both were addressed with somewhat equal frequency in the 1990s. Among non-disease-focused articles, lesbian, gay, and transgender individuals were addressed in articles on identity and etiology more frequently than bisexual persons were. Research on health care focused predominantly on lesbians and gay men and less on bisexual and transgender persons. Family is the only content area that addressed lesbians more frequently than any other group. Research on attitudes, risk factors, and violence tended to focus more frequently on lesbians and gay men than on bisexual and transgender persons.

Table 4 presents all articles by sexual orientation and years of publication. It indicates differences in the proportion of research on each group and also after the exclusion of 2108 articles on STDs that focused predominantly on gay and bisexual men, as presented in Table 3. Comparisons of each group's proportion of articles between the first 5 years (1980–1984) and the last 5 years (1995–1999) of the study show the change in research attention on each group. The proportion of articles on lesbians and on transgender persons decreased by 5% and 22%, respectively. In contrast, the proportion of articles on gay men increased by 25%, on bisexual men increased by 20%, and on bisexual women increased by 5%. After research on STDs was excluded, the largest gain in research attention was among lesbians; articles on lesbians increased by 20%. Research attention on gay men, on bisexual women, and on bisexual men increased by 16%, 15%, and 15%, respectively, whereas research attention on transgender persons decreased by 21%.

The same calculations were performed with regard to race/ethnicity (data not shown). Between 1980 and 1984, 20 (7.8%) of 256 articles specified race/ethnicity, and between 1995 and 1999, 274 (16.1%) of 1707 articles specified LGBT persons' race/ethnicity. After articles on STDs were excluded, 16 (7.7%) of 207 articles specified race/ethnicity between 1980 and 1984, and 74 (10.4%) of 709 articles did so between 1995 and 1999. Thus, reporting of study subjects' race/ethnicity between the first 5 years (1980–1984) and the last 5 years (1995–1999) increased by a factor of 2, whereas reporting of race/ethnicity increased by only a factor of 1.3 among articles on non-STD-related topics. Therefore, the increase in recognition of subjects' race/ethnicity was driven mostly by STD-focused research.

DISCUSSION

LGBT persons are estimated to constitute between 1% and 10% of the total population.[12] Only 3777 articles, the equivalent of 0.1% of the MEDLINE database, focused on LGBT individuals over the past 20 years. This indicates that public health research neglected these populations and that LGBT persons are underrepresented as explicit research subjects.

TABLE 4—Distribution of Articles That Considered Lesbian, Gay, Bisexual, and Transgender Populations, by Sexual Orientation and Years of Publication[a]

	1980–1984 n (%)	1985–1989 n (%)	1990–1994 n (%)	1995–1999 n (%)	Total All Years N (%)
Lesbian	88 (34.4)	148 (25.2)	301 (24.6)	506 (29.6)	1043 (27.6)
Gay	152 (59.4)	419 (71.3)	1012 (82.5)	1444 (84.6)	3027 (80.1)
Bisexual women	18 (7.0)	38 (6.5)	90 (7.3)	206 (12.1)	352 (9.3)
Bisexual men	43 (16.8)	263 (44.7)	557 (45.4)	623 (36.5)	1486 (39.3)
Transgender	71 (27.7)	82 (13.9)	92 (7.5)	101 (5.9)	346 (9.2)
Total	256 (100)	588 (100)	1226 (100)	1707 (100)	3777 (100)
After exclusion of 2108 articles on sexually transmitted diseases					
Lesbian	86 (41.5)	135 (45.6)	264 (57.8)	442 (62.3)	927 (55.5)
Gay	110 (53.1)	157 (53.0)	297 (65.0)	489 (69.0)	1059 (63.5)
Bisexual women	16 (7.7)	27 (9.1)	66 (14.4)	159 (22.4)	265 (15.9)
Bisexual men	19 (9.2)	50 (16.9)	82 (17.9)	171 (24.1)	324 (19.4)
Transgender	71 (34.3)	81 (27.4)	82 (17.9)	92 (13.0)	326 (19.5)
Total	207 (100)	296 (100)	457 (100)	709 (100)	1669 (100)

[a]More than 1 sexual orientation can be addressed in a single article. Therefore, the number and percentage of lesbian, gay, bisexual, and transgender articles add up to more than the total.

The classification of articles by topic showed that 56% of the research on LGBT persons was disease-specific in the context of STDs. In particular, 52% of the research was disease-specific in the context of HIV or AIDS. Subsequently, the coding of articles by sexual orientation appears to indicate a gender gap, in that most articles focused on gay and bisexual men (80% and 39%, respectively). After the exclusion of STD-focused articles, this gender gap was greatly reduced.

Emerging in the 1980s, AIDS brought visibility to LGBT persons as a population with specific health concerns. One would expect that thereafter LGBT health concerns other than STDs would be examined. For example, some evidence suggests that LGBT individuals have a higher prevalence of known risk factors for heart disease,[1,2] the leading cause of mortality in the United States.[10] Behavior changes are targeted as primary prevention of heart disease in the general population.[10] Without public health research that explores heart-healthy lifestyle interventions in the context of sexual orientation and transgender identity, cardiovascular risk factors may not be reduced among LGBT persons.

It is unknown how many researchers attempted to move the focus of LGBT research to non-STD-related research topics, but funding was 1 major barrier that prevented a broadening of LGBT health research. A review of lesbian-, gay-, and bisexual-related National Institutes of Health funding indicated that, since 1982, $20 million annually was spent on HIV-focused research, compared with an average of $532 000 annually between 1974 and 1992 on lesbian, gay, and bisexual research unrelated to HIV.[18]

In the 1990s, the LGBT communities publicized AIDS and breast cancer as 2 major health threats to LGBT populations.[9] However, the review of the MEDLINE database did not reflect such notoriety, because only 6 articles on breast cancer were found—fewer than expected, considering the attention on breast cancer by LGBT communities and the national media. Without restricting MEDLINE to LGBT persons, using only "breast cancer" as a key word, breast cancer citations indexed in MEDLINE more than doubled from the 1980s to the 1990s—from 10 258 to 26 554 articles—which coincided with an increase in funding for breast cancer research in the 1990s.[19] Yet the expansion of breast cancer research neglected lesbians and bisexual women. Requiring the inclusion—or at least a justification of the exclusion—of sexual orientation and transgender identity, as the National Institutes of Health's Revitalization Act mandated for women and minorities,[20–22] would raise researchers' awareness and may be necessary to remedy the underrepresentation of LGBT populations in research.

The lack of knowledge about health problems of LGBT persons is further compounded with regard to race/ethnicity: 85% of the studies did not report race/ethnicity. Race/ethnicity has been shown to affect incidence, likelihood of survival, and mortality rates of many health problems. The omission of race/ethnicity is an important shortcoming within LGBT research because dual minority status—sexual and racial/ethnic—affects illness, health needs, and behaviors.

In addition, the duality of socioeconomic and LGBT status affects health and mediates its perceptions. The reviewed LGBT population–related abstracts provided only sparse, inconsistent information that did not allow even a basic coding of abstracts as reporting vs not reporting socioeconomic position. This lack of complete, uniform data is a major limitation of this review but is consistent with public health data, which mostly lack conceptually valid measures of socioeconomic status.[23,24] Another limitation of this analysis is its restriction to abstracts rather than articles to handle the sheer volume of studies. Some abstracts could potentially misrepresent the content of studies, but publishers generally urge authors to write abstracts that summarize their studies comprehensively, which made the coding of abstracts a reasonable approach. Another study limitation is the restriction to MEDLINE. Several journals that either exclusively or frequently focus on LGBT health are excluded from the MEDLINE database, suggesting the possibility that this review represents LGBT health only as reflected in 1 mainstream database and not the state of LGBT health research.

The finding that most LGBT research is related to STDs raises questions about the framework of lesbian-, gay-, bisexual-, and transgender-related public health research and points to the dominance of a biomedical paradigm that narrowly understands LGBT health in relation to sexual behavior. This reduction of LGBT individuals to their "different" sexuality needs to be expanded to include the recognition of sexual orientation and transgender identity as cultural and social

categories that shape all health experiences. A new framework that incorporates social and cultural categories will advance public health knowledge and research practice related to LGBT populations and will encompass the racial/ethnic and socioeconomic diversity of LGBT populations as well. ■

About the Author

Ulrike Boehmer is with Boston University, School of Public Health, Boston, Mass, and Center for Health Quality, Outcomes, and Economic Research, Bedford, Mass.

Requests for reprints should be sent to Ulrike Boehmer, PhD, Center for Health Quality, Outcomes, and Economic Research, 200 Springs Rd (152), Bedford, MA 01730 (e-mail: boehmer@bu.edu).

This article was accepted May 23, 2001.

Acknowledgments

I would like to thank Dr Dan Berlowitz for his help with the classification of physical diseases, Dr Angela Radan for her help with the classification of mental diseases, and an anonymous reviewer for valuable suggestions.

References

1. Solarz AL, ed. *Lesbian Health: Current Assessment and Directions for the Future.* Washington, DC: Institute of Medicine, Committee on Lesbian Health Research Priorities, Neuroscience and Behavioral Health Program, Health Sciences Policy Program, Health Sciences Section; 1999.

2. *Lesbian, Gay, Bisexual, and Transgender Health: Findings and Concerns.* San Francisco, Calif: Gay and Lesbian Medical Association and Center for Lesbian, Gay, Bisexual, and Transgender Health, Mailman School of Public Health, Columbia University; 2000.

3. Lamberg L. Gay is okay with APA—forum honors landmark 1973 events. American Psychiatric Association [news]. *JAMA.* 1998;280:497–499.

4. Spitzer RL. The diagnostic status of homosexuality in DSM-III: a reformulation of the issues. *Am J Psychiatry.* 1981;138:210–215.

5. Centers for Disease Control. Pneumocystis pneumonia—Los Angeles. *MMWR Morbid Mortal Wkly Rep.* 1981;30:250–252.

6. Altman LK. Rare cancer seen in 41 homosexuals. *New York Times.* July 3, 1981;sect 20:1.

7. Treichler PA. AIDS, gender, and biomedical discourse: current contests for meaning. In: Fee E, Fox DM, eds. *AIDS: The Burdens of History.* Berkeley: University of California Press; 1988:190–266.

8. Schiller NG, Crystal S, Lewellen D. Risky business: the cultural construction of AIDS risk groups. *Soc Sci Med.* 1994;38:1337–1346.

9. Boehmer U. *The Personal and the Political: Women's Activism in Response to the Breast Cancer and AIDS Epidemics.* Albany: State University of New York Press; 2000.

10. *Healthy People 2010: Understanding and Improving Health.* Washington, DC: US Dept of Health and Human Services; 2001. Also available at: http://www.health.gov/healthypeople/default.htm. Accessed August 15, 2000.

11. Gay and Lesbian Medical Association. *Healthy People 2010 Companion Document for Lesbian, Gay, Bisexual and Transgender (LGBT) Health.* Available at: http://www.glma.org/policy/hp2010/. Accessed April 12, 2001.

12. Laumann EO, Gagnon JH, Michael RT, Michaels S. *The Social Organization of Sexuality: Sexual Practices in the United States.* Chicago, Ill: University of Chicago Press; 1994.

13. Bradford J, White JC. Lesbian health research. In: Goldman MB, Hatch MC, eds. *Women and Health.* San Diego, Calif: Academic Press; 2000:64–78.

14. Sell RL, Petrulio C. Sampling homosexuals, bisexuals, gays, and lesbians for public health research: a review of the literature from 1990 to 1992. *J Homosex.* 1996;30(4):31–47.

15. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.* Washington, DC: American Psychiatric Association; 1994.

16. Currah P, Minter S, Green J. *Transgender Equality: A Handbook for Activists and Policymakers.* Washington, DC: Policy Institute of the National Gay and Lesbian Task Force; 2000.

17. Williams DR. Race and health: basic questions, emerging directions. *Ann Epidemiol.* 1997;7:322–333.

18. Silvestre AJ. Gay male, lesbian and bisexual health-related research funded by the National Institutes of Health between 1974 and 1992. *J Homosex.* 1999;37(1):81–94.

19. Altman R. *Waking Up/Fighting Back: The Politics of Breast Cancer.* Boston, Mass: Little Brown & Co; 1996.

20. Haynes SG, Hatch M. State-of-the-art methods for women's health research. In: Goldman MB, Hatch MC, eds. *Women and Health.* San Diego, Calif: Academic Press; 2000:37–49.

21. Haynes SG. The role of women in health care and research. In: Goldman MB, Hatch MC, eds. *Women and Health.* San Diego, Calif: Academic Press; 2000: 25–26.

22. Seaman B, Wood SF. Role of advocacy groups in research on women's health. In: Goldman MB, Hatch MC, eds. *Women and Health.* San Diego, Calif: Academic Press; 2000:27–36.

23. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health.* 1997;18:341–378.

24. Krieger N, Fee E. Social class: the missing link in US health data. *Int J Health Serv.* 1994;24:25–44.

Prediction of Depressive Distress in a Community Sample of Women: The Role of Sexual Orientation

Alicia K. Matthews, PhD, Tonda L. Hughes, PhD, RN, Timothy Johnson, PhD, Lisa A. Razzano, PhD, and Roberta Cassidy, MS, RN

Objectives. This study compared factors known or hypothesized to influence depressive symptomatology in a community sample of lesbians and heterosexual women.

Methods. Data were collected in a multisite survey of lesbians' physical and mental health.

Results. Findings confirmed earlier reports suggesting that traumatic life events such as physical and sexual abuse, and individual traits and coping styles are risk factors for depressive distress. However, findings of higher rates of suicidal behavior and of several risk factors for depressive distress among lesbians suggest that risk for depression may differ among lesbians and heterosexual women.

Conclusions. Sexual orientation may represent an important but poorly understood risk factor for depressive distress as well as suicidal ideation and behavior. (*Am J Public Health.* 2002;92:1131–1139)

Studies have consistently shown that rates of depression among women are twice as high as those among men.[1–5] Reported rates of lifetime risk for depression among women in the general US population range between 10% and 25%, with point prevalence rates in community samples ranging from 5% to 9%.[6–8] Risk factors for depression include genetic, biochemical, and hormonal factors[9,10]; family history[11,12]; previous depressive episodes[8]; chronic general medical conditions[13]; personality style or coping strategies[14]; negative life events associated with poverty[5]; psychosocial events or stressors[15]; and substance dependence.[8]

Numerous studies suggest that risk factors unique to women contribute substantially to sex differences in depression.[5] These factors include women's roles and status,[16] female sex role socialization,[17,18] presence of dependent children,[19] lower socioeconomic status relative to men,[20] and victimization experiences (e.g., childhood sexual abuse,[21] physical or sexual violence, battering by an intimate partner, marital and acquaintance rape, sexual harassment[22–28]). Although considerable knowledge about the sex correlates of depression in the general population has been amassed,[3,4,12] the applicability of this knowledge to lesbians is unknown.[29,30]

LESBIANS AND DEPRESSION

Despite a paucity of data on depression among lesbians, this group generally is thought to be at greater risk for depression than are heterosexual women.[30–34] In addition to risk factors shared with heterosexual women (e.g., relationship status and satisfaction, divorce or dissolution of an intimate relationship, perceived lack of or low social support from friends and family), lesbians are believed to be affected by additional, unique risk factors, including the coming out process, level of disclosure of sexual orientation, discrimination experiences, and chronic stress associated with being a member of a stigmatized minority group.[30,35–38]

The aims of the present study were to (1) compare indicators of depressive distress among lesbians and a demographically matched sample of heterosexual women and (2) examine the relationships of several hypothesized or known predictors with depressive distress in lesbians and heterosexual women. Such information is important in understanding risk factors for depression among lesbians and for identifying interventions that address these risks.

METHODS

Study Design and Data Collection

Data were collected as part of a study initiated by the Chicago Lesbian Community Cancer Project (LCCP) in 1992. The study began in Chicago and was replicated in Minneapolis–St Paul, Minn, and New York City during 1994 through 1996. The goal was to gather information on the general health status as well as behavioral and environmental health risks of lesbians. The study was designed to obtain a diverse sample of women who relate sexually or affectionally (or both) with women.

As a means of reaching the broadest possible range of women, the survey instrument was distributed in a variety of formal and informal lesbian venues (e.g., potluck dinners; discussion groups; bookstores; softball and bowling leagues; coffee houses; college social, support, therapeutic, musical, and political groups and organizations). In addition, lesbian participants were recruited through numerous informal social networks. Individual and group settings were used to collect survey data.

Each lesbian who completed the questionnaire was instructed to give a second, color-coded copy to a female friend, acquaintance, or colleague whose work role (including student, homemaker, or retiree) was as similar as possible to her own. In the Chicago survey, we did not specify that the "work-role counterpart" selected be heterosexual. Given the less-than-optimal results (a sample consisting of only about half as many heterosexual women as lesbians), instructions provided in the Minnesota and New York surveys specified that the work-role counterpart be a woman whom the lesbian knew or presumed to be heterosexual.

Unlike respondents at the other 2 sites, lesbian respondents in New York City were given a small incentive of $15 for completing the survey and for recruiting a work-role counterpart; heterosexual work-role counterparts were given $10 for completing the survey. In addition, representatives of harder-to-reach lesbian groups who agreed to act as

TABLE 1—Data Collection Sites, Study Years, and Sample Sizes

Site	Years	No. of Lesbians	No. of Heterosexual Women	Total
Chicago	1994–1995	273	134	407
Minneapolis–St Paul	1994–1995	160	67	227
New York	1995–1996	117	78	195
Total		550	279	829

distributors of the survey were paid $5 for each completed survey they returned.

These strategies resulted in a more racially diverse sample and a larger proportion of heterosexual respondents in New York than at the other 2 survey sites. Lesbians and work-role counterparts were given instructions to complete and return the survey in person or by mail in a preaddressed, postage-paid envelope. As a means of preserving anonymity, no code numbers or other identifying data were included on the questionnaires. Because we wished the participants to remain anonymous, we were unable to calculate precise response rates. However, on the basis of the number of questionnaires distributed and returned at each of the sites, we estimated that the overall response rate was approximately 48%.

The data presented here were derived from the combined Chicago, New York, and Minneapolis–St Paul samples. Table 1 displays the numbers of lesbian and heterosexual respondents at each of the sites. It should be noted that the samples differed according to location of data collection on several important characteristics, including age, race, education, and income level (for a more detailed description of the sample and site differences, see Hughes et al.[39]). For example, the New York sample was older and more racially diverse, had lower levels of formal education, and had a higher median income level than the Chicago and Minnesota samples ($Ps \leq .05$).

The site differences just described constituted both a limitation and a strength of the data set. On the one hand, these differences, as well as their influence on other variables, might be obscured by combining the samples. On the other hand, the combined sample was more heterogeneous and thus more representative of lesbians and heterosexual women overall. Other advantages of combining the sample included greater variability in responses, a larger sample size, and the ability to detect smaller differences.

Measures

The study questionnaire addressed a broad range of areas that influence women's mental and physical health, including personal health history (e.g., general, menstrual, gynecologic health), health-related practices (e.g., diet, health screening, alternative health practices such as acupuncture), mental health (e.g., physical and sexual abuse, use of psychotherapeutic medications), use of legal and illegal substances, access to and use of health and mental health services, relationships and supports, and demographic characteristics. In addition, the survey instrument included questions about sexual attraction and sexual behavior, on the basis of which orientation was defined.

The definition of sexual orientation used in the present analyses was based on responses to 2 survey questions regarding (1) current sexual interest or attraction and (2) sexual behavior in the year preceding the survey. Both questions included the following response categories: "only men," "mostly men," "equally men and women," "mostly women," and "only women." The question concerning sexual behavior also included the category "I have not had sex in the past year." We created 3 categories of sexual orientation—lesbian, bisexual, and heterosexual—by summarizing the combinations of responses to these 2 questions.

Approximately two thirds of the women (n=550; 62%) were categorized as lesbians; 279 (32%) were categorized as heterosexuals, and 33 (4%) were classified as bisexuals. In the remaining cases, respondents were not classified because responses to the attraction and behavior questions were missing or inconsistent (n=19; 2%). These individuals, along with women categorized as bisexual, were omitted from the analyses presented here.

Dependent Variables

Because the study questionnaire did not include a standardized measure of depression, 4 indicators of depressive distress were explored: history of therapy for an emotional or mental health problem, history of therapy or counseling for depression or use of an antidepressant medication, suicidal ideation, and suicide attempts.

History of therapy/counseling. As a means of assessing history of mental health service use, respondents were asked whether they had "ever received therapy or counseling for an emotional or mental health problem." Respondents also were asked to indicate reasons for seeking mental health services.

Treatment for depression. Past history of depression was assessed via participants' responses to 2 questions focusing on depression as a reason for seeking counseling and past use of antidepressant medication. Responses to these questions were combined and coded (0=response of no to both questions, 1=response of yes to at least one of the questions).

Suicidal ideation. Another potential indicator of past depression or other mental health problems is suicidal thoughts. Respondents were asked to report whether they "had ever seriously considered committing suicide at some time in the past."

Suicide attempts. The final indicator of depressive distress used in the study was history of suicide attempts. Respondents were asked "Have you ever tried to kill yourself?" Those who reported having made at least 1 suicide attempt were asked how old they were when the attempt(s) occurred.

Predictors of Depressive Distress

Potential predictors of depressive distress were selected on the basis of known or hypothesized risk factors for depression and included history of physical violence and sexual abuse, level of stress, social support, and coping strategies.

Physical violence. As a means of assessing history of physical violence, participants were asked "Have you ever been the victim of nonsexual physical violence?"

Childhood sexual abuse. Rates of childhood sexual abuse were calculated via participants' responses to the question, "Has anyone ever forced you to engage in any form of sexual

activity that you didn't want?" A subsequent question asked how old the respondent was when this incident occurred. Although the more common definition of childhood sexual abuse is forced sexual activity before the age of 18 years,[40] questions included in the survey did not permit separation of women who reported that they had experienced forced sex between 15 and 19 years of age. Thus, only women who indicated that they had experienced unwanted sex before the age of 15 years were scored positively on the childhood sexual abuse measure (1=yes, 0=no).

Perceived level of stress. Level of current stress was assessed on the basis of responses to a 4-point Likert scale item (0=no stress, 3=high stress). Percentage agreement scores were calculated.

Global stress index. A composite stress index was created through summing responses to 18 different potential sources of stress (e.g., money, work, family) and the perceived intensity ratings (0=none, 3=extreme) associated with each source. The resultant global stress index (GSI) was a continuous variable ranging from 0 to 54 (α=0.83). Higher scores on the index reflected a greater number as well as a greater severity of stressors.

Lack of social support. Participants were asked to indicate "significant" sources of social support. Those who indicated no sources of such support were assigned a code of 0; all others were assigned a code of 1, indicating that they perceived themselves as having at least 1 source of social support.

Coping strategies. Studies suggest that use of a combination of emotion-focused and problem-focused coping strategies in response to stress is associated with improved psychological adjustment.[41,42] As a means of exploring the relationship between the use of several coping responses and depressive distress, respondents were asked to rate their use of the following positive coping strategies on a 4-point Likert scale (0=never, 3=often): talking about a problem, doing something fun, confronting the problem, and exercising.

A continuous measure ranging from 0 to 12 (mean=3.4, SD=1.9; α=0.51) was created through summing responses to the 4 coping strategies. Higher scores were associated with more frequent use of positive coping strategies in response to stressful life events. Finally, respondents were asked how often they become "overly emotional" in response to stress. Responses to the 4-point Likert scale (0=never, 3=often) were calculated as a percentage agreement score.

Data Analyses

Univariate statistical techniques were used to generate frequency distributions (measures of central tendency and dispersion); *t* tests were used to test for differences between continuous variables. Chi-square analyses and Pearson product–moment correlation coefficients revealed simple bivariate associations between indicators of depressive distress and selected predictor variables. A series of logistic regression analyses was conducted to explore the unique associations of the independent variables with the 4 indicators of depressive distress. In logistic regression models, history of therapy, history of treatment for depression, history of suicidal ideation, and history of suicide attempts measures were coded 1 to indicate a positive response and 0 to indicate the absence of such a response.

With the exception of 2 variables, all predictors included in the regression analyses were coded as either continuous or dummy variables. Perceived levels of current stress and emotionality during stress were measured with 4-point Likert-type scales. Analyses were conducted initially with the ordinally scaled measures and then with sets of dummy coded variables. The second set of analyses produced monotonic trends across the dummy variables, suggesting that the findings were consistent with the first set of analyses. Therefore, to conserve degrees of freedom and reduce the number of odds ratios (ORs) reported, we included these 2 variables in the analyses as ordinal measures. All significant differences reported exceeded the .05 probability level.

RESULTS

Sample

Table 2 presents demographic characteristics of the 829 (550 lesbian and 279 heterosexual) respondents included in the analyses. The average age of participants was 43 years (SD=10.8). The majority of the women taking part were White (74%), were married or involved in a committed relationship (66%), had more than a high school education (84%), and were employed full time for pay (73%). The median household income range for both lesbians and heterosexual women was $36 000 to $50 999.

TABLE 2—Respondent Demographic and Socioeconomic Characteristics

	Lesbians (n = 550), No. (%)	Heterosexuals (n = 279), No. (%)
Age, y		
<30	53 (10)	44 (16)
31–40	187 (34)	89 (32)
41–50	186 (34)	82 (30)
51–60	88 (16)	36 (13)
>60	31 (6)	25 (9)
Education		
High school or less	79 (15)	46 (17)
Some college	263 (48)	132 (48)
Advanced degree	204 (37)	98 (36)
Annual income, $		
<10 000	35 (6)	20 (7)
10 000–20 999	77 (14)	27 (10)
30 000–35 999	160 (29)	72 (26)
36 000–50 999	121 (22)	55 (20)
51 000–75 999	89 (16)	47 (17)
>76 000	66 (12)	56 (20)
Ethnicity		
African American	75 (14)	42 (15)
Caucasian	418 (76)	200 (72)
Other	57 (10)	29 (10)
Relationship status		
Single	102 (24)	55 (25)
In a committed relationship	280 (66)	39 (18)
Married	37 (9)	118 (53)
Employment status		
Full time	409 (75)	202 (72)
Part time	88 (16)	46 (17)
Unemployed	37 (7)	16 (6)
Retired	14 (3)	8 (3)
Disabled	11 (2)	4 (1)

Note. Numbers on which percentages were based vary because of missing data on some variables.

Indicators of Depressive Distress

Past use of therapy or counseling. The majority of lesbians (78%) reported that they

TABLE 3—Study Predictor Variables, by Sexual Orientation

	Lesbians (n = 550)	Heterosexuals (n = 279)
Dependent variable, No. (%)		
Ever received therapy	429 (78)**	157 (56)
Ever treated for depression	284 (58)	109 (52)
Suicidal ideation	280 (51)**	104 (38)
Suicide attempts	91 (22)*	22 (13)
Predictor variable		
Childhood sexual abuse, No. (%)	165 (30)**	45 (16)
Physical abuse, No. (%)	246 (45)	114 (41)
Moderate or extreme stress level, No. (%)	461 (85)	229 (83)
Emotionality in response to stress (sometimes or often), No. (%)	367 (68%)	207 (75%)
Global stress index, mean (SD)	16 (6.8)	17 (7.0)
Positive coping strategies, mean (SD)	3.3 (1.8)	3.5 (1.9)

Note. Numbers on which percentages were based vary because of missing data on some variables. Scores on the GSI range from 0–54, with higher scores reflecting both greater number and severity of life stressors. Positive coping strategies scores range from 0–12, with higher scores representing more frequent use of a variety of positive coping strategies in response to stress.
*$P \leq .01$; **$P \leq .001$.

had "received therapy or counseling for an emotional or mental health problem" at some point in their life (Table 3). This rate was significantly higher than that among heterosexual women (56%; $\chi^2_1 = 43.3$, $P \leq .001$). However, rates of current therapy or counseling did not differ for lesbians (38%) and heterosexual women (30%; $\chi^2_1 = 2.98$, not significant).

Treatment for depression. Similar percentages of lesbians (56%) and heterosexual women (49%) reported that they had sought therapy or counseling for depression. Twenty-six percent of lesbians and 20% of heterosexual women reported that they had been prescribed medication for a mental or emotional problem. Among those who had received medication, 68% of lesbians and 75% of heterosexual women reported taking an antidepressant medication at some point. The majority of women who reported seeking help for sadness or depression (83%) also reported receiving antidepressants. More than half of the total sample of lesbians (58%) and heterosexual women (52%) reported at least one of these 2 indicators of past treatment for depression.

Suicidal ideation and suicide attempts. Significant differences were found between lesbians and heterosexual women in regard to reports of whether they had seriously considered committing suicide and whether they had actually attempted suicide in the past. Fifty-one percent of lesbians and 38% of heterosexual women reported seriously considering suicide at some point in the past ($P \leq .001$). Most suicide attempts among women in this study occurred between the ages of 15 and 29 years. More than twice as many lesbians as heterosexual women in this age group reported suicide attempts ($P \leq .01$).

Predictors of Depressive Distress

Physical and sexual abuse. Although lesbians and heterosexual women were equally likely to report that they had been victims of nonsexual physical violence (45% and 41%, respectively), significantly more lesbians (30%) than heterosexual women (16%) reported experiencing childhood sexual abuse ($P \leq .001$). Because our analyses included only women who had experienced forced sex before the age of 15 years (rates were 45% and 41%, respectively, of all lesbians and heterosexual women who reported any unwanted sex), these rates probably underestimate the number of lesbians and heterosexual women who actually experienced childhood sexual abuse.

Global stress. Overall mean scores on the global stress index were in the lower range and did not differ according to sexual orientation. Mean stress index scores were 17 (SD = 7.0) for heterosexual women and 16 (SD = 6.8) for lesbians ($t_{823} = 0.573$, NS).

Perceived stress. No differences were found between the lesbians and heterosexual women in terms of level of perceived stress. The majority of lesbians (85%) and heterosexual women (83%) reported moderate to extreme levels of perceived stress ($\chi^2_3 = 1.1$, NS). The only statistically significant differences in sources of stress for lesbians and heterosexual women involved children ($\chi^2_4 = 58$, $P < .001$) and sexual identity ($\chi^2_4 = 72$, $P < .001$); more heterosexual women rated children as moderately or extremely stressful, and more lesbians rated sexual identity as moderately or extremely stressful.

Perceived support. Differences were observed between lesbians and heterosexual women in terms of perceived lack of support. More heterosexual women (6%) than lesbians (3%) reported an absence of social support ($\chi^2_1 = 5.5$, $P \leq .01$).

Coping strategies and response to stress. Overall, use of positive coping strategies was low among both lesbians and heterosexual women (mean of 3.4 on the 12-point measure, SD = 1.9). Whereas similar (and relatively low) percentages of lesbians (7%) and heterosexual women (6%) reported talking or reasoning out feelings during times of stress, lesbians were more likely to report never using talking as a coping strategy (46% vs 37%; $\chi^2_3 = 9.6$, $P < .05$).

Moreover, fewer lesbians (19%) than heterosexual women (25%) reported doing something fun when they were stressed ($\chi^2_3 = 7.4$, $P = .06$) or using exercise as a coping strategy (36% vs 43%). Only 7% of lesbians and 8% of heterosexuals reported confronting situations directly. A higher percentage of heterosexual women (75%) than lesbians (68%) reported becoming overly emotional in response to stress, but this difference was not statistically significant ($\chi^2_3 = 4.6$, $P > .05$).

Multivariate Predictor Models of Depressive Distress

Only variables significantly related to at least 1 of the indicators of depressive distress in the bivariate analyses were included in the multivariate analyses. Demographic characteristics retained included education level (0 =

TABLE 4—Pooled Logistic Regression Analyses: Correlates of Depressive Distress

	Ever Received Therapy (n = 795)		Treated for Depression (n = 778)		Suicidal Ideation (n = 775)		Suicide Attempts (n = 778)	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Education level	2.05	1.05, 2.09***	1.41	1.11, 1.79**	1.00	0.708, 1.14	0.622	0.447, 0.866*
Race	1.41	1.13, 1.75**	1.83	1.42, 2.38***	0.810	0.596, 1.10	0.900	0.708, 1.14
Sexual orientation	2.93	2.04, 4.21***	1.74	1.24, 2.45***	1.78	1.27, 2.50***	2.15	1.25, 3.69**
Childhood sexual abuse	1.87	1.18, 2.98**	1.52	1.04, 2.22*	2.48	1.71, 3.61***	2.41	1.52, 3.84***
Physical abuse	1.89	1.30, 2.75**	1.65	1.19, 2.29**	2.52	1.83, 3.47***	2.60	1.61, 4.19***
Global stress index	1.05	1.02, 1.08**	1.07	1.04, 1.09*	1.05	1.02, 1.08***	1.00	0.969, 1.04
Current stress	1.48	1.05, 2.09*	1.49	1.08, 2.05*	1.11	0.820, 1.52	0.995	0.653, 1.51
Positive coping skills	1.06	0.98, 1.15**	1.04	0.987, 1.10	0.996	0.940, 1.05	0.952	0.877, 1.03
Emotionality	0.58	0.48, 0.74***	1.70	1.36, 2.14***	1.38	1.11, 1.71**	1.52	1.11, 2.09**

Note. OR = odds ratio; CI = confidence interval.
*P < .05; **P < .01; ***P < .001.

high school or less, 1=some college or bachelor's degree, 2=professional or advanced degree), ethnicity (0=non-White, 1=White), and sexual orientation (0=heterosexual, 1=lesbian). Other variables retained were history of physical and sexual abuse (0=no, 1=yes), global stress index score (0–54), emotionality during stress (0=never, 3=often), perceived level of current stress (0=none, 3=extreme), and use of positive coping strategies (12-point measure).

Table 4 presents the odds ratios and corresponding confidence intervals for each of the independent predictors of the 4 indicators of depressive distress. The first model tested the relationship between history of therapy or counseling for a mental health problem and the 9 predictor variables. An evaluation of the overall model against a constant-only model produced a statistically reliable result (χ^2_9=173.70, P<.001). As a set, the predictors reliably distinguished between women who had participated in counseling for a mental health problem and those who had not (log-likelihood=–784.52, df=9).

Overall prediction rates were modest (75%); the model successfully identified more of the women who had participated in counseling for a mental health problem than of those who had not. Significant independent predictors of lifetime use of therapy or counseling were higher education level, White race/ethnicity, lesbian sexual orientation, higher global stress index score, history of childhood sexual abuse or physical abuse, high current stress levels, greater use of positive coping strategies, and more frequent emotionality during stress.

The second model tested the relationship between the treatment for depression measure and our study predictors. An evaluation of the overall depression model with all 9 predictors against a constant-only model produced a statistically reliable result (χ^2_9=154.35, P≤.001). As a set, the predictors reliably distinguished between women who had received treatment for depression and those who had not (log-likelihood=–921.46, df=9).

Although overall prediction rates were modest (69%), the model successfully identified more of the women who had received treatment for depression than of those who had not. Significant independent predictors of having received treatment for depression were higher education level, White race/ethnicity, lesbian sexual orientation, higher global stress index score, history of childhood sexual abuse or physical abuse, high current stress levels, and more frequent emotionality during stress.

The next model tested the overall goodness of fit between the predictor variables and history of suicidal ideation. As a set, the predictors reliably distinguished between women with and without histories of suicidal ideation (χ^2_9=141.59, P≤.001; log-likelihood=–928.28, df=9). Again, overall prediction rates for this model were modest (69%). On the basis of the 9 predictor variables, success rates were 61% for identifying women with a history of suicidal ideation and 75% for identifying women with no such history. Lesbian sexual orientation, history of physical violence or of childhood sexual abuse, higher global stress index score, and more frequent emotionality in response to stress were the strongest predictors of suicidal ideation.

The final model tested predictors of history of suicide attempts. A test of the full model was statistically reliable (χ^2_9=82.05, P≤.001), indicating that the predictors, as a set, reliably distinguished between women with a history of suicide attempts and those without such a history (log-likelihood=–526.19, df=9). Prediction was better for nonattempters (99%); only 10% of attempters were correctly predicted, resulting in an overall success rate of 88%. As can be seen in Table 4, lower education level, lesbian sexual orientation, history of physical violence or of childhood sexual abuse, and more frequent emotionality in response to stress were independent predictors of past suicide attempts.

To investigate the possibility that one or more of the variables may have confounded or otherwise influenced the association between sexual orientation and our measures of depressive distress, we conducted additional logistic regression analyses using block entry of variables. The first block included sexual orientation; the second block included education level and ethnicity; the third block included perceptions of high current stress, history of physical violence or of childhood

TABLE 5—Logistic Regression Analyses: Separate Models for Lesbians and Heterosexual Women

	Ever Received Therapy				Treated for Depression				Suicidal Ideation				Suicide Attempts			
	Lesbians (n = 528)		Heterosexuals (n = 267)		Lesbians (n = 528)		Heterosexuals (n = 267)		Lesbians (n = 528)		Heterosexuals (n = 267)		Lesbians (n = 525)		Heterosexuals (n = 267)	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Education	2.0	1.4, 2.8***	2.1	1.4, 3.3***	1.3	1.0, 1.7*	1.7	1.1, 2.6**	0.92	0.69, 1.2	1.3	0.87, 1.9	0.67	0.46, 0.99*	0.55	0.28, 1.0
Race	1.7	1.3, 2.3***	1.0	0.77, 1.5	1.4	1.1, 1.9**	1.2	0.86, 1.7	0.82	0.63, 1.0	0.95	0.67, 1.3	0.80	0.59, 1.0	1.0	0.60, 1.9
Childhood sexual abuse	2.3	1.2, 4.2**	1.1	0.51, 2.5	1.4	0.94, 2.2	1.7	0.84, 3.8	2.5	1.6, 3.8***	2.5	1.2, 5.4**	2.4	1.4, 4.1***	1.9	0.66, 5.7
Physical abuse	2.3	1.4, 3.9***	1.3	0.76, 2.4	1.7	1.1, 2.5**	1.3	0.77, 2.4	2.1	1.4, 3.1***	3.5	1.9, 6.2***	2.9	1.7, 5.1***	2.0	0.75, 5.6
Global stress index	1.1	1.0, 1.1**	1.0	1.0, 1.1*	1.0	1.0, 1.1***	1.1	1.0, 1.1***	1.0	1.0, 1.1**	1.0	1.0, 1.1**	1.0	0.96, 1.0	0.98	0.92, 1.0
Current stress	1.0	0.69, 1.6	2.5	1.4, 4.5***	1.3	0.94, 2.0	1.8	1.0, 3.2*	1.2	0.86, 1.8	0.88	0.51, 1.5	0.96	0.59, 1.5	1.5	0.64, 3.8
Positive coping	1.0	0.93, 1.1	1.1	1.0, 1.2*	1.0	0.95, 1.1	0.98	0.87, 1.1	1.0	0.92, 1.1	0.90	0.79, 1.0	0.89	0.80, 1.0	0.97	0.79, 1.1
Emotionality	0.61	0.44, 0.81***	0.53	0.35, 0.79**	0.54	0.41, 0.70***	0.72	0.48, 1.0	0.65	0.50, 0.85***	0.95	0.63, 1.4	0.59	0.42, 0.85**	1.1	0.57, 2.2

Note. OR = Odds ratio; CI = confidence interval.
*P < .05; **P < .01; ***P < .001.

sexual abuse, and global stress index score; and the final block included coping responses and emotionality in response to stress. Blocks 2, 3, and 4 represented potential confounders, mediators, and moderators, respectively. Inclusion of covariates had no impact on the direction or magnitude of the main effects of sexual orientation on any of the dependent variables (data not shown). Given these results, we used simultaneous entry of predictor variables in all subsequent analyses.

To determine the relative influence of each of the predictor variables on the indicators of depressive distress for lesbians in comparison with heterosexual women, we conducted separate regression analyses for the 2 groups. Predictor variables were entered simultaneously for each of the models. As shown in Table 5, similar patterns were found for the predictors in our pooled analyses and in the analyses run separately by sexual orientation. However, significant independent predictors and the strength of the association between the predictors and the dependent variables differed in the lesbian and heterosexual models.

The model predicting suicide attempts showed the greatest variability. Consistent with the pooled model shown in Table 4, lower education level, history of childhood sexual abuse or of physical abuse, and more frequent emotionality in response to stress were all significant predictors of suicide attempts among lesbians. However, none of the independent variables tested were significant predictors of suicidal attempts among heterosexual women ($\chi^2_8 = 10.0$, NS).

Finally, we estimated interaction effects using the pooled data for each of our 4 models of depressive distress. Only the interaction between ethnicity and sexual orientation predicting lifetime use of therapy or counseling was significant: White lesbians were more likely to have ever participated in therapy (OR = 0.42, $P < .05$). In addition, an interaction was observed between perceived stress level and sexual orientation, indicating that perception of stress is a stronger predictor of lifetime use of therapy or counseling for heterosexual women than for lesbians (OR = 1.6, $P < .05$).

DISCUSSION

Studies of lesbians' mental health have historically been characterized by methodological limitations such as small, homogeneous samples; inconsistent or absent definitions of sexual orientation; and lack of comparison groups of heterosexual women. The present study addressed several of these limitations by collecting data from lesbians in 3 geographic locations and by asking lesbian participants to assist in the recruitment of a heterosexual comparison group. These methods produced a large and relatively diverse group of lesbians as well as a heterosexual comparison group that was demographically very similar to the lesbians.

Obtaining large probability samples of lesbians that can be compared with women from the general population is both difficult and expensive. In the absence of such samples, the method used in this study provides the most rigorous comparison because it "controlled" for many of the demographic and life experiences—other than sexual orientation—that might have influenced the dependent variables.[43] As a result, differences found between lesbians and heterosexual women in this study can more confidently be attributed to sexual orientation.

The differences in rates of suicidal ideation and in suicide attempts between lesbians and heterosexual women are particularly striking. Among our respondents, 51% of lesbians, in comparison with 38% of heterosexual women, reported that they had seriously considering committing suicide at some time in the past. Almost all of the suicide attempts occurred between the ages of 15 and 29 years, and lesbians in this age group were twice as likely as heterosexual women to have attempted suicide. Similar findings were reported in the National Lesbian Health Care Survey.[31]

Earlier analyses of the multisite data described here[39] revealed that although the majority of both lesbians and heterosexual women who had sought therapy or counseling had done so in their 20s or 30s, more of the lesbians than of the heterosexual women had sought therapy or counseling during

these years ($P \leq .001$). There were no differences between lesbians and heterosexual women in terms of therapy use rates during their 40s, 50s, and 60s. Furthermore, lesbians were significantly more likely than heterosexual women to report problems related to sexual identity, suicidal feelings, sexual abuse, and alcohol and other drugs as reasons for seeking therapy or counseling.[39]

Younger lesbians may be at increased risk for using alcohol and other drugs to self-medicate against the anxiety and depression associated with accepting a stigmatized identity. In addition, because many lesbians begin to question their sexual identity or "come out" in their teens or 20s,[44] these age groups may be at a particularly high risk for depression, suicide, or both.

Women who have experienced sexual or physical violence are more likely than women who have not been abused to suffer from psychological problems, including suicide attempts, major depression, dissociative disorders, and alcohol and other drug abuse.[45–47] Epidemiological studies have shown that 15% to 33% of adult women report childhood sexual abuse.[47,48] Rates of childhood sexual abuse in our combined sample reached the upper limits of rates reported in one of these studies.[48]

The rates of childhood sexual abuse among lesbians found in our study were significantly higher than those found among heterosexual women and about the same as rates reported for women in the general population. Suicidal ideation and suicide attempts (as well as the depression measure included in the study) were strongly associated with childhood sexual and physical abuse. Women with histories of sexual or physical abuse were 2 to 3 times more likely to have had thoughts of killing themselves and to have acted on these thoughts. Past traumatic experiences, such as sexual or physical abuse, may add to the vulnerability of young lesbians, who may be grappling with issues related to coming out, and may increase the risk of suicide.

Because this was a convenience sample, the high rates of childhood sexual abuse may reflect sample bias. The lesbians who chose to participate in our study may have been more likely to have had difficult or traumatic experiences, including childhood sexual abuse, or may have been more likely to report such experiences. The higher rates of childhood sexual abuse among lesbians may also be attributable to lesbians' greater willingness to acknowledge and report this experience.

Studies have consistently shown that a majority of lesbians report past use of therapy or counseling.[31,49,50] Therapy experiences may increase lesbians' comfort with acknowledging both their sexual identity and other stigmatized statuses or experiences such as childhood sexual abuse (T.L. Hughes, T. Johnson, and S.C. Wilsnack, unpublished data, 2002). Nevertheless, the high rates observed of childhood sexual abuse, suicidal ideation, and suicide attempts have important implications for clinicians who treat lesbians, particularly younger lesbians and those in the early stages of coming out.

The regression analyses conducted in this study showed a clear, direct, and independent association of sexual orientation with the 4 indicators of depressive distress. However, with the exception of suicide attempts, the pattern of associations between predictor variables and depressive distress did not vary by sexual orientation. These findings suggest that although the models tested included factors that are indeed associated with depressive distress, other factors not included in our models may also in part account for the association between sexual orientation and depressive distress. For example, self-esteem, internalized homophobia, level of social support, and religious attitudes and beliefs may be important variables that moderate or mediate the relationship between sexual orientation and depressive distress.

Sexual orientation appears to be an important potential risk factor in women's experiences of depressive distress. However, whether this risk is conferred through the long-term or chronic stress associated with membership in a stigmatized minority group or through more time-limited stressful life circumstances, such as those associated with coming out as lesbian, is not yet clear. More research is needed if there is to be a better understanding of risk factors for depression among lesbians. It is also important to examine factors unique to lesbians that may be protective, such as higher levels of education[51] and use of therapy.[31,50] Lesbians in this study, as has been true of those in a number of other investigations, reported high rates of therapy and counseling. Use of mental health services may moderate the psychological consequences of coming out as a member of a stigmatized minority group or the consequences of other traumatic life events such as childhood sexual abuse.

Limitations

This study addressed some of the limitations of previous research on lesbians' mental health by exploring the role of sexual orientation in depressive distress, more systematically defining sexual orientation, and including a more appropriate heterosexual comparison group.[43] Despite these strengths, several important limitations must be noted.

First, because the findings were derived from a nonrandom, convenience sample, their generalizability is limited. In addition, although efforts were made to increase minority participation in the study, only limited success was achieved in recruiting women of color. Provision of monetary incentives for completing the survey and participation of African American and Hispanic lesbians in assisting with recruitment substantially improved the representation of these groups at the New York site, suggesting that such methods are important for increasing the participation of racial/ethnic minorities in other studies of lesbian health.

Second, the indicators of depressive distress included in this study did not assess duration or intensity of symptoms. History of therapy or counseling as an indicator of psychological distress may be confounded by issues related to access to mental health services and may underestimate actual rates of depressive distress among the women in our study. This may be especially true among African Americans, a group showing a consistent pattern of underuse of formal mental health services.[52,53]

In previous analyses of data from the African American lesbians in the present sample, we found rates of suicidal ideation and suicide attempts that were similar to those reported for lesbians in the overall sample, but somewhat lower rates of therapy or counseling (T.L. Hughes, A. Matthews, L. Razzano, F. Aranda, and A. Haas, unpublished data,

2002). Larger studies involving standard definitions and sufficient numbers of racial/ethnic minority women are needed to more fully explore depressive distress among lesbians.

Finally, because data for this study were collected at one point in time, we were unable to assess the temporal order of the dependent and independent variables. Longitudinal studies are needed to more accurately identify the particular variables that predict depressive distress.

Conclusions

The findings of this study support earlier reports suggesting that traumatic life events, such as physical and sexual abuse, and life stress are risk factors for depressive distress in women. In addition, our findings suggest that sexual orientation may represent an important but poorly understood risk factor for depressive distress, especially suicidal behavior. Consistent with recent recommendations of the Institute of Medicine,[54] findings presented here contribute to a better understanding of the full range of female experience and to greater knowledge about associations between sexual orientation and mental health outcomes. Such information is important if treatment planning and service provision for women are to be effective. ■

About the Authors

Alicia K. Matthews is with the Department of Psychiatry, University of Chicago, Chicago, Ill. Tonda L. Hughes and Roberta Cassidy are with the Department of Public Health, Mental Health and Administrative Nursing, University of Illinois at Chicago. Timothy Johnson is with the Survey Research Laboratory, University of Illinois at Chicago. Lisa A. Razzano is with the Mental Health Services Research Program, University of Illinois at Chicago.

Requests for reprints should be sent to Alicia K. Matthews, PhD, University of Chicago, Department of Psychiatry, 5841 S Maryland Ave (MC3077), Chicago, IL 60637 (e-mail: Amatthew@yoda.bsd.uchicago.edu).

This article was accepted June 6, 2001.

Contributors

Each of the authors provided substantial contributions to the article, including conceptualization and design of the analyses (A.K. Matthews, T.L. Hughes, T. Johnson, L.A. Razzano), data analysis and interpretation (A.K. Matthews, L.A. Razzano, T.L. Hughes, T. Johnson), literature review and write-up (A.K. Matthews and R. Cassidy), and drafting and revision of the article.

Acknowledgments

Merging of the three data sets, data analysis, and preparation of the article were supported by the Lesbian Health Fund; a research grant from the National Institute of Mental Health (1R24 MH54212); the Department of Psychiatry, University of Illinois at Chicago; and an internal research support grant from the University of Illinois at Chicago College of Nursing. The Chicago survey was supported by the Chicago Board of Health and the Chicago Foundation for Women.

We would like to acknowledge the Lesbian Community Cancer Project, as well as Ellie Emanuel, PhD, principal investigator for the Minneapolis–St Paul survey, and Ann Pollinger Haas, PhD, principal investigator for the New York City survey.

References

1. Radloff LS. Sex differences in depression: the effects of occupation and marital status. *Sex Roles.* 1975; 1:249–265.

2. Kessler RC, McGonagle KA, Zhoa S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry.* 1994; 51:8–19.

3. Nolen-Hoeksema S. Sex differences in unipolar depression: evidence and theory. *Psychol Bull.* 1987; 101:259–282.

4. Nolen-Hoeksema S. *Sex Differences in Depression.* Stanford, Calif: Stanford University Press; 1990.

5. McGrath E, Keita GP, Strickland BR, Russo NF. *Women and Depression: Risk Factors and Treatment Issues.* Washington, DC: American Psychological Association; 1990.

6. Maxmen JS, Ward NG. *Essential Psychopathology and Its Treatment: Second Edition Revised for DSM-IV.* New York, NY: WW Norton & Co; 1995.

7. Romanoski AJ, Folstein MF, Nestadt G, et al. The epidemiology of psychiatric-ascertained depression and *DSM-III* depressive disorders: results from the Eastern Baltimore Mental Health Survey clinical reappraisal. *Psychol Med.* 1992;22:629–655.

8. *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Washington, DC: American Psychiatric Association; 1994.

9. Gershon ES. The genetics of affective disorders. In: Grinspoon L, ed. *Psychiatric Update.* Vol. 2. Washington, DC: American Psychiatric Association; 1983: 434–457.

10. Gershon ES, Berrettini W, Nurnberger J, Goldin LR. Genetics of affective illness. In: Meltzer HY, ed. *Psychopharmacology: The Third Generation of Progress.* New York: Raven Press; 1987:481–491.

11. Allen MG. Twin studies of affective illness. *Arch Gen Psychiatry.* 1976;33:1476–1478.

12. Klerman GL, Weissman MM. Depressions among women: their nature and causes. In: Gutentag M, Salasin S, Belle D, eds. *The Mental Health of Women.* New York, NY: Academic Press Inc; 1985:57–92.

13. Strickland BR. Sex-related differences in health and illness. *Psychol Women Q.* 1988;12:381–399.

14. Lazarus R. Coping theory and research: past, present, and future. *Psychosom Med.* 1993;55:234–247.

15. Kaplan G, Roberts R, Camacho T, Coyne J. Psychosocial predictors of depression: prospective evidence from the Human Population Laboratory Studies. *Am J Epidemiol.* 1987;125:206–220.

16. Kandel DB, Davies M, Ravels V. The stressfulness of daily social roles for women: marital, occupational and household roles. *J Health Soc Behav.* 1985;26: 64–78.

17. Belle D. Social ties and social support. In: Belle D, ed. *Lives in Stress: Women and Depression.* Beverly Hills, Calif: Sage Publications; 1982:147–167.

18. Aneshensel C. Marital and employment role-strain, social support, and depression among women. In: Hobfoll S, ed. *Stress, Social Support, and Women.* Washington, DC: Hemisphere Publishing; 1986: 99–114.

19. Throits PA. Multiple identities: examining gender and marital status differences in distress. *Am Sociol Rev.* 1986;51:259–272.

20. Golding JM. Gender differences in depressive symptoms: statistical considerations. *Psychol Women Q.* 1988;12:61–74.

21. Rosenthal HG. The learned helplessness syndrome: specific strategies for crisis intervention with the suicidal sufferer: emotional first aid. *Am J Crisis Intervention.* 1986;3:5–8.

22. Hilberman E, Munson K. Sixty battered women. *Victimology Int J.* 1977–1978;2:460–470.

23. Roth S, Liebowitz L. The experiences of sexual trauma. *J Trauma Stress.* 1988;1:79–108.

24. Bagley C, Ramsey R. Sexual abuse in childhood: psychological outcomes and implications for social work practice. *J Soc Work Hum Sexuality.* 1986;5: 33–47.

25. Murphy SM, Kilpatrick DG, Amick-Mcmullan A, et al. Current psychological functioning of child sexual assault survivors. *J Interpersonal Violence.* 1988;3: 55–79.

26. Gidycz CA, Koss MP. The impact of adolescent sexual victimization: standardized measures of anxiety, depression, and behavioral deviancy. *Violence Vict.* 1989;4:139–149.

27. Koss MP. The women's mental health research agenda: violence against women. *Am Psychol.* 1990; 45:374–380.

28. Sorenson SB, Golding JM. Depressive sequelae of recent criminal victimization. *J Trauma Stress.* 1990;3: 337–350.

29. Green B, Boyd-Franklin N. African American lesbian couples: ethnocultural considerations in psychotherapy. In: Hill M, Rothblum E, eds. *Couples in Therapy: Feminist Perspectives.* New York, NY: Haworth Press; 1996:49–60.

30. Rothblum ED. Depression among lesbians: an invisible and unresearched phenomenon. *J Gay Lesbian Psychother.* 1990;1(3):67–87.

31. Bradford J, Ryan C, Rothblum E. National Lesbian Health Care Survey: implications for mental health care. *J Consult Clin Psychol.* 1994;62:228–242.

32. Cochran SD, Mays VM. Depressive distress among homosexually active African American men and women. *Am J Psychiatry.* 1994;151:524–529.

33. Hughes TL. Sexual identity and alcohol use: a comparison of lesbians' and heterosexual women's patterns of drinking. Paper presented at: New Approaches to Research on Sexual Orientation, Mental Health, and Substance Abuse (conference), September 1999, Rockville, Md.

34. Trippet SE. Lesbians' mental health concerns. *Health Care Women Int.* 1994;15:317–323.

35. Ayala J, Coleman H. Predictors of depression

among lesbian women. *J Lesbian Stud.* 2000;4(3):71–86.

36. Oetjen H, Rothblum ED. When lesbians aren't gay: factors affecting depression among lesbians. *J Homosex.* 2000;39;49–73.

37. Meyer IH. Minority stress and mental health in gay men. *J Health Soc Behav.* 1995;36:38–56.

38. Tait D. Stress and social support networks among lesbian and heterosexual women: a comparison study. *Smith Coll Stud Soc Work.* 1997;67:213–224.

39. Hughes TL, Haas A, Razzano L, Matthews AK, Cassidy R. Comparing lesbians' and heterosexual women's mental health: findings from a multi-site study. *J Gay Lesbian Soc Serv.* 2000;11:57–76.

40. Wyatt GE. The sexual abuse of Afro-American and White American women in childhood. *Child Abuse Negl.* 1985;9:507–519.

41. Maes S, Leventhal H, de Ridder DTD. Coping with chronic diseases. In: Zeidner M, Endler NS, eds. *Handbook of Coping: Theory, Research, Application.* New York, NY: John Wiley & Sons Inc; 1996:221–251.

42. Wineman NM, Durand EJ, Steiner RP. A comparative analysis of coping behaviors in persons with multiple sclerosis and spinal cord injury. *Res Nurs Health.* 1994;17:185–194.

43. Hughes TL, Wilsnack SC, Johnson T. Conducting research on lesbians' mental health: what is an appropriate comparison group? In: Omoto A, Kurtzman H, eds. *Recent Research on Sexual Orientation, Mental Health, and Substance Abuse.* Washington, DC: American Psychological Association. In press.

44. Ryan C, Futterman D. *Lesbian and Gay Youth: Care and Counseling.* New York, NY: Columbia University Press; 1998.

45. Blumenthal SJ. *Violence Against Women: Fact Sheet.* Washington, DC: Office of Women's Health Research, US Dept of Health and Human Services; 1995.

46. Langeland W, Hartgers C. Child sexual and physical abuse and alcoholism. *J Stud Alcohol.* 1998;59:336–348.

47. Wilsnack SC, Vogeltanz ND, Klassen AD, Harris TR. Childhood sexual abuse and women's substance abuse: national survey findings. *J Stud Alcohol.* 1997;58:264–271.

48. Polusny MA, Follette VM. Long-term correlates of child sexual abuse: theory and review of the empirical literature. *Appl Prev Psychol.* 1995;4:143–166.

49. Morgan KS. Caucasian lesbians' use of psychotherapy: a matter of attitude. *Psychol Women Q.* 1992;16:127–130.

50. Sorensen L, Roberts SJ. Lesbian uses of and satisfaction with mental health services: results from the Boston Lesbian Health Project. *J Homosex.* 1997;33:35–49.

51. Rothblum ED, Factor R. Lesbians and their sisters as a control group: demographic and mental health factors. *J Psychol Sci.* 2001;12:63–69.

52. Sue DW, Sue D. *Counseling the Culturally Different: Theory and Practice.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1990.

53. Sussman LK, Robin LN, Earls F. Treatment-seeking for depression by black and white Americans. *Soc Sci Med.* 1987;24:187–196.

54. Solarz A, ed. *Lesbian Health: Current Assessment and Directions for the Future.* Washington, DC: National Academy Press; 1999.

Risk Behaviors, Medical Care, and Chlamydial Infection Among Young Men in the United States

Leighton Ku, PhD, MPH, Michael St. Louis, MD, Carol Farshy, BS, Sevgi Aral, PhD, Charles F. Turner, PhD, Laura D. Lindberg, PhD, and Freya Sonenstein, PhD

Objectives. This study assessed factors related to chlamydial infection among young men in the United States.

Methods. Data were from interviews of nationally representative samples of 470 men aged 18 to 19 years (teenagers) and 995 men aged 22 to 26 years (young adults) and from urine specimens tested by means of polymerase chain reaction.

Results. Although a majority of the men reported occasional unprotected intercourse, only a minority perceived themselves to be at risk for contracting a sexually transmitted disease (STD). Chlamydial infection was detected in 3.1% of the teenagers and 4.5% of the young adults. A minority of those infected had symptoms or had been tested for STDs; very few had been diagnosed with STDs.

Conclusions. Chlamydial infection is common but usually asymptomatic and undiagnosed. Primary and secondary prevention efforts should be increased, particularly among young adult men. (*Am J Public Health.* 2002;92:1140–1143)

Primary prevention of sexually transmitted diseases (STDs) involves adopting safer sexual behaviors, whereas secondary prevention involves detecting and treating infected persons to reduce transmission to sexual partners. Our study examined *Chlamydia trachomatis* (referred to here as "chlamydial infection") among young men for the purpose of making recommendations for improving primary and secondary prevention of the disease. We based our examination on the National Surveys of Adolescent Males (NSAM).

METHODS

NSAM included two surveys: a 1995 nationally representative survey of 1729 men aged 15 to 19 years (which had a 75% response rate)[1] and a 1988 nationally representative survey of 1880 men aged 15 to 19 years (which had a 74% response rate).[2] Seventy-five percent of the respondents to the 1988 survey were reinterviewed in 1995.

Data for the teenaged sample were collected from 470 men aged 18 to 19 years who were part of the 1995 survey; data for the young adult sample were collected from 995 men aged 22 to 26 years who were part of the 1988 survey and who were reinterviewed in 1995. Both the teenaged and the young adult sample weights depend on the original sample probabilities and on poststratification adjustments made to correspond with census data. The young adult sample weights also include longitudinal adjustments to compensate for attrition between 1988 and 1995.

Respondents were interviewed about their background, behaviors, and attitudes. They then completed self-administered questionnaires about sensitive topics, including STD symptoms. After the interview and questionnaire were completed, respondents older than 18 years were asked for a urine specimen. They were informed that the specimen would be tested for STDs and not for drugs and that positive cases would be reported to health departments where legally required.[3] Respondents received $10 to $20 for the interview and an additional $10 to $20 for the specimen. Some interviews were conducted by telephone; in these cases, urine specimens were not collected.

After urine specimens were collected, they were packed in ice, frozen, and shipped overnight (still packed in ice) to the Centers for Disease Control and Prevention (CDC) for analysis. Commercial polymerase chain reaction (Amplicor, Roche Diagnostic Systems) was used to test the specimens for chlamydial infection.[4] All positive cases were confirmed by ligase chain reaction; there were no discordant positives.

Laboratory results were not available for all respondents. Results were unavailable for 382 of the 1377 young adult respondents (28%); 14% were unavailable for logistical reasons (primarily because interviews were conducted by telephone, but also because of shipping damage, etc.) and 14% were unavailable because of respondent refusal. Results were unavailable for 108 of the 578 teenaged respondents (18%), 6% for logistical reasons and 12% because of respondent refusal.

We conducted extensive analyses to determine whether the missing data caused nonresponse biases.[5] Respondents for whom laboratory results were missing were not at higher risk for chlamydial infection than were those for whom results were present. Using multiple imputation methods,[6,7] we determined that the nonresponse bias was negligible. We used actual laboratory results in our analyses of the young adult respondents and teenaged respondents. All analyses were weighted and adjusted to take into account the complex sampling design.

RESULTS

STD Risk Behaviors and Perceived Risk

Three fourths of the teenagers and nine tenths of the young adults surveyed were sexually active (Table 1). Three fourths of the sexually active teenagers had engaged in at least 1 act of unprotected sexual intercourse during the past year. Among the young adults, three fourths of those who were single and sexually active and almost all of those who were married or cohabiting had engaged in unprotected intercourse during the past year.

Only 15% of the teenagers reported having 3 or more sexual partners during the past year; the rate was two times higher for Black teenagers than for non-Black teenagers. One third of the single, sexually active young adults, but few of the married or cohabiting young adults, reported having 3 or more part-

TABLE 1—STD Behaviors, Perceptions, Symptoms, Testing, and Diagnosis Among US Teenaged and Young Adult Men

	Men Aged 18–19 Years				Men Aged 22–26 Years				
	All	Black	Non-Black	Sexually Active	All	Black	Non-Black	Married or Cohabiting	Single, Sexually Active
Unweighted sample size[a]	470	139	331	357	995	371	624	356	543
Sexual behaviors									
Had sex with female during past year	73.4%	85.6%	71.4%*	100.0%	88.6%	91.2%	88.2%	99.2%	100.0%
Had any unprotected sex during past year	54.9	60.7	54.0	73.3	77.3	78.5	77.1	98.2	76.4***
Had 3 or more female partners during past year	14.9	27.8	12.8***	20.3	18.0	29.3	16.0***	3.7	33.6***
Had sex with high-risk partner during past year[b]	2.9	4.6	2.7	3.9	6.8	13.6	5.6***	6.4	7.3
STD perceptions									
Believes he is at some risk for STDs	40.7	37.1	41.3	42.2	26.4	33.6	25.2*	12.7	39.9***
Believes last sexual partner is at some risk for STDs[c]	24.7	27.1	24.3	24.7	16.1	14.6	16.4	4.8	25.2***
STD symptoms									
Ever had STD symptoms[d]	20.3	25.4	19.5	21.9	20.9	31.5	19.1***	17.0	22.1
Had STD symptoms during past year	7.4	14.9	6.2*	8.4	8.2	11.4	7.7	6.1	8.5
Health access									
Has health insurance	78.1	83.2	77.3	77.9	74.1	74.3	74.1	76.1	71.8
Had physical exam during past year	67.6	68.0	67.6	67.2	55.3	63.5	53.8**	56.9	59.0*
STD testing									
Had STD test during past year	15.8	36.2	12.6***	19.4	17.3	32.0	14.8***	15.6	21.1**
Had STD test during past year (among those symptomatic)	25.6	56.4	13.8**	30.1	33.5	53.9	28.7*	29.8	46.5
STD diagnosis									
Medical professional said he had an STD during past year	2.7	13.7	1.0***	3.7	1.6	5.7	0.9***	0.4	2.8**
Told he had an STD during past year[e]	11.4	35.7	0.3***	12.4	8.1	15.9	5.4**	1.8	12.1

Note. Significance tests compare levels for Black vs. non-Black men among the 18–19 year olds and for Black vs. non-Black and for married/cohabiting vs. single sexually active men among 22–26 year olds.
[a]Sample sizes are unweighted; all other estimates are weighted.
[b]Defined as having had sexual intercourse with a prostitute or an injection drug user or having had anal or oral sex with a male.
[c]Among those sexually active.
[d]Symptoms include burning on urination or abnormal genital discharge.
[e]Among those with STD test.
*P < .1; **P < .05; ***P < .01.

ners during the past year. Only 3% of the teenagers reported having sexual intercourse with a high-risk partner during the past year; the percentage was twice as high for young adults.

Although most of the sexually active teenagers and young adults had unprotected intercourse during the past year, only two fifths felt themselves to be at some risk for contracting an STD. Fewer believed that their female partners were at risk; one fourth of both the sexually active teenagers and the single young adults believed that their last female partner was at some risk for contracting an STD.

Symptoms and Medical Care

About one fifth of both the teenagers and the young adults reported ever having had symptoms related to chlamydial infection, whereas 8% or fewer had had symptoms during the past year. Whether the respondents were sexually active was not significantly related to STD symptoms (data not shown). Some of the men who were not sexually active might have had urethritis of nonsexual etiology or might have misreported their sexual behaviors.

Although the great majority of the respondents had health insurance and had received a physical exam during the past year, only one sixth had been tested for STDs. (We do not know for which STDs they were tested.) Only one fourth of the symptomatic teenagers and one third of the symptomatic young adults had been tested. Very few of the respondents, even among those who had been tested, had been diagnosed with an STD.

Black men were more likely to have been tested for STDs or to have been diagnosed with an STD than were non-Black men. Black men also were more likely to have a history of STD symptoms. But even among the men with symptoms, Black men were more likely than non-Black men to have been tested for STDs.

Prevalence of Infection, Symptoms, and Medical Care

According to PCR test results, 3.1% of the teenagers and 4.5% of the young adults had chlamydial infections.[5] Most of those with symptoms did not test positive, however, and many of those who did test posi-

TABLE 2—STD Symptoms and Medical Care Received Among US Teenaged and Young Adult Men With and Without Chlamydial Infection

	Men Aged 18–19 Years		Men Aged 22–26 Years	
	Tested Positive for Chlamydial Infection	Tested Negative for Chlamydial Infection	Tested Positive for Chlamydial Infection	Tested Negative for Chlamydial Infection
Unweighted sample size	26	444	66	929
Had STD symptoms during past year	23.4%	6.9%*	7.5%	8.2%
Had a physical exam during past year	75.0	67.4	59.6	55.1
Had an STD test during past year	45.7	14.8**	16.9	17.3
Ever told he had an STD	26.5	2.0**	9.6	6.5
Told he had an STD during past year	26.5	2.0**	3.3	1.5
Had any unprotected sex during past year	87.3	53.9***	88.4	76.8*
Had 3 or more female partners during past year	37.7	14.1	31.0	17.4
Had sex with high-risk partner during past year	10.8	2.7	4.1	6.9

*$P < .1$; **$P < .05$; ***$P < .01$.

tive were asymptomatic: 10.1% of the symptomatic teenagers and 4.1% of the symptomatic young adults were infected, compared with 2.7% of the asymptomatic teenagers and 4.5% of the asymptomatic young adults.

As shown in Table 2, the great majority of infected teenagers and young adults were asymptomatic; only 23% of the infected teenagers and 8% of the infected young adults had experienced symptoms during the past year. Most of the infected men from both groups had access to routine health care, but only a minority had been tested for STDs. Teenagers who had been tested for STDs during the past year were more likely to be infected than those who had not, but no such relationship existed among the young adults. A small minority of the infected men had been diagnosed with an STD during the past year (27% of the teenagers and 3% of the young adults).

Men who had unprotected sexual intercourse were more likely to be infected than those who had not ($P < .01$ for teenagers; $P < .1$ for young adults). Some of those who were infected had reported no unprotected sexual intercourse in the past year. This apparent discrepancy could be attributable to (1) misreporting about condom use,[9] (2) condom failure,[10] (3) transmission of infection despite condom use (e.g., by epidermal contact rather than by fluid transfer), (4) false-positive laboratory results, or (5) becoming infected more than 1 year ago.

DISCUSSION

Most of the young men with chlamydial infection were asymptomatic and undiagnosed. A minority of those infected had been tested for STDs, and probably only a fraction of the STD tests were for chlamydial infection. Thus, most of the chlamydial infections were not detected. Chlamydial infection screening of women has increased in the past decade, and the CDC recommends that all sexually active adolescent females be screened for chlamydial infection at each pelvic exam.[11] However, the agency has no equivalent recommendations for young men, although in 1998 a CDC advisory committee recommended routine screening of young men to help prevent HIV transmission.[12]

At any point in time, 3% to 5% of teenaged and young adult men in the United States have a chlamydial infection. Given that some receive treatment or otherwise clear their infection, the percentage who become infected must be even higher. These estimates are conservative, because surveys probably undercount certain high-risk groups, such as homeless or incarcerated men. Men who engage in unprotected intercourse are more likely to have chlamydial infection than those who do not. Such risky sexual behavior has double consequences: when a man has unprotected sexual intercourse, he not only may become infected himself but also may transmit the infection to future partners. A study of couples found that about 70% of those with chlamydial infection also had an infected current partner.[13]

Primary prevention of chlamydial infection should begin with reducing the percentage of men who engage in unprotected intercourse; data indicate that this percentage fell during the late 1990s, primarily owing to increased condom use.[1] Efforts to reduce the incidence of chlamydial infection should continue with measures designed to increase public awareness of STD risks, especially among single young adults. Most single young adult men in the United States have occasional unprotected intercourse, but few view themselves as susceptible to STDs, and even fewer believe that their partners are susceptible. This lack of awareness means that men fail to seek testing or treatment for STDs, even when they have symptoms.

Because most men with chlamydial infection are asymptomatic, screening efforts broader than those currently in place should also be considered. Most young women obtain routine gynecological care, during which STD-related risks may be evaluated; however, a comparable system of routine reproductive health care for young men does not exist.[14] Those who provide primary care for men need to promote an increased awareness of STDs among their young patients.

Future research can help identify whether widespread screening of young men for chlamydial infection is appropriate. The cost-effectiveness of such screening depends in part on whether it will help prevent sequelae (e.g., pelvic inflammatory disease) among these men's female sexual partners. If STD screening for young men is to become more widespread, it could be incorporated into existing components of primary care (e.g., physical exams for sports, school, or employment). Although this strategy was infeasible in the past, the availability of new DNA-based methods now makes STD testing possible in the primary care setting.[15] ■

About the Authors

At the time of the study, Leighton Ku was with the Urban Institute, Washington, DC. Michael St. Louis is with the Centers for Disease Control and Prevention, Harare, Zimbabwe. Carol Farshy and Sevgi Aral are with the Centers for Disease Control and Prevention, Atlanta, Ga. Charles Turner is with Research Triangle Institute, Washington, DC, and City University of New York. At the time of the study, Laura Lindberg was with the Urban Institute, Washington, DC. Freya Sonenstein is with the Urban Institute, Washington, DC.

Requests for reprints should be sent to Leighton Ku, Center on Budget and Policy Priorities, 820 First St, NE, Suite 510, Washington, DC 20002 (e-mail: ku@cbpp.org).

This article was accepted May 1, 2001.

Contributors

Leighton Ku designed the study and wrote the article. Michael St. Louis analyzed and interpreted data and contributed to the writing of the article. Carol Farshy conducted laboratory analyses and interpreted laboratory findings. Sevgi Aral interpreted the relationship between behaviors and infection. Charles Turner provided statistical advice. Laura Lindberg interpreted risk behaviors and coordinated data editing. Freya Sonenstein conceptualized and led the National Surveys of Adolescent Males. All authors participated in conceptual design, reviewed analysis and reports, and reviewed drafts.

Acknowledgments

This research was supported under National Institutes of Health grant RO1 HD30861, with funding from the National Institute of Child Health and Human Development (NICHD), the Centers for Disease Control and Prevention (CDC), the Office of Population Affairs, and the National Institute of Mental Health.

Harvey Zelon and his colleagues at the Research Triangle Institute (RTI) administered data collection in 1995. Carolyn Black and Debra Candal of the CDC were instrumental in conducting laboratory analyses. Sean Williams, formerly of the Urban Institute, provided research assistance. Vicki Freedman of RAND, Ralph Folsom of RTI, and Roderick Little of the University of Michigan provided statistical consulting. We thank Christine Bachrach of NICHD and Heather Miller, formerly of RTI, for their inspiring ideas and sound advice.

References

1. Sonenstein FL, Ku L, Lindberg LD, Turner CF, Pleck JH. Changes in sexual behavior and condom use among teenaged males: 1988 to 1995. *Am J Public Health.* 1998;88:956–959.

2. Sonenstein FL, Pleck JH, Ku LC. Sexual activity, condom use, and AIDS awareness among adolescent males. *Fam Plann Perspect.* 1989;21:152–158.

3. Ku L, Sonenstein FL, Turner CF, Aral SO, Black CM. The promise of integrated representative surveys about sexually transmitted diseases and behavior. *Sex Transm Dis.* 1997;24:299–309.

4. Bauwens JE, Clark AM, Loeffelholz MJ, et al. Diagnosis of *Chlamydia trachomatis* urethritis in men by polymerase chain assay of first-catch urine. *J Clin Microbiol.* 1993;31:3013–3016.

5. St. Louis ME, Ku L, Aral S, Black C, Williams S, Sonenstein F. Prevalence of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections among young men in the United States: results of a representative national survey. Paper presented at: International Society for Sexually Transmitted Disease Research, July 27, 1999, Denver, Colo.

6. Little RJA, Rubin DB. *Statistical Analysis With Missing Data.* New York: John Wiley & Sons; 1987.

7. Heitjan D, Little R. Multiple imputation for the Fatal Accident Reporting System. *Applied Stat.* 1991; 40:13–29.

8. Porter LE, Ku L. Use of reproductive health services among young men, 1995. *J Adolesc Health.* 2000; 27:186–194.

9. Zenilman JM, Weisman CS, Rompalo AM, et al. Condom use to prevent incident STDs: the validity of self-reported condom use. *Sex Transm Dis.* 1995;22: 15–21.

10. Warner L, Clay-Warner J, Boles J, Williamson J. Assessing condom use practices: implications for evaluating method and user effectiveness. *Sex Transm Dis.* 1998;25:273–277.

11. Centers for Disease Control and Prevention. Recommendations for the prevention and management of *Chlamydia trachomatis* infections, 1993. *MMWR Morb Mortal Wkly Rep.* 1993;42:1–39.

12. Centers for Disease Control and Prevention. HIV prevention through early detection and treatment of other sexually transmitted diseases—United States. Recommendations of the Advisory Committee for HIV and STD Prevention. *MMWR Morb Mortal Wkly Rep.* 1998;47:1–24.

13. Quinn TC, Gaydos C, Shepherd M, et al. Epidemiologic and microbiologic correlates of *Chlamydia trachomatis* infection in sexual partnerships. *JAMA.* 1996; 276:1737–1742.

14. Sonenstein F, ed. *Young Men's Sexual and Reproductive Health: Toward a National Strategy.* Washington, DC: Urban Institute Press; 2000.

15. Turner CF, Rogers SM, Miller HG, et al. Untreated gonococcal and chlamydial infection in a probability sample of adults. *JAMA.* 2002;287:726–733.

Mammography Screening and Differences in Stage of Disease by Race/Ethnicity

| Jillian Jacobellis, PhD, MS, and Gary Cutter, PhD

Objectives. We examined the effect of routine screening on breast cancer staging by race/ethnicity.

Methods. We used a 1990 to 1998 mammography database (N=5182) of metropolitan Denver, Colo, women to examine each racial/ethnic cohort's incident cancer cases (n=1902) and tumor stage distribution given similar patterns of routine screening use.

Results. Regardless of race/ethnicity, women participating in routine screenings had earlier-stage disease by 5 to 13 percentage points. After control for possible confounding factors, White women were more likely to have early-stage disease compared with Black and Hispanic women.

Conclusions. Lack of screening coverage in certain racial/ethnic populations has often been cited as a reason for tumor stage differences at detection. In this study, correcting for screening did not completely reduce stage differentials among Black and Hispanic women. (*Am J Public Health.* 2002;92:1144–1150)

Breast cancer incidence rates have risen in the United States for the past 2 decades. Although the lifelong chance of developing breast cancer is higher for White women than for Black and Hispanic women, Black women and subgroups of Hispanic women have a lower breast cancer survival rate.[1–6] In the United States, Black and Hispanic women disproportionately have poor breast cancer outcomes. Black women diagnosed with breast cancer are twice as likely to die from the disease within 5 years after diagnosis and Hispanic women are 1.5 times as likely to die as White women.[1,3] Black and Hispanic women undergo fewer baseline and routine mammography screenings and have more advanced stage of disease at diagnosis, which in part explains the observed decreases in survival rates. Several investigators contend that race/ethnicity is in part a determinant of resource access and that it is a contributing factor in the racial/ethnic disparity.[7–13]

Racial/ethnic differentials in breast cancer incidence rates, staging, and survival seen nationwide are similar to those observed in the Colorado and the 6-county Denver metropolitan area study data. Non-Hispanic White (hereafter labeled White), Hispanic, and non-Hispanic Black (Black) women in the Denver metropolitan area demonstrated an 11- to 21-percentage-point increase in early-stage disease between the period 1985 to 1987 and the period 1996 to 1997. Although increases in early detection were seen in all groups, presumably as a result of screening, tumor stage differentials were still apparent among Black, Hispanic, and White women. In 1996 to 1997, more Black (48%) and Hispanic (46%) women than White women (40%) received a diagnosis of advanced-stage breast cancer.[1]

Several randomized controlled trials of screening found significant decreases in mortality rates. Notably, the Stockholm and Malmo trials reported that 85% to 100% of breast cancer deaths occurred among women diagnosed with stage II, III, or IV breast cancer disease.[14–16] Screening mammography for early breast cancer detection has been investigated quite thoroughly, but studies examining the association of screening with race/ethnicity have been limited. The current investigation was undertaken in light of the observational evidence regarding the association of race/ethnicity with breast cancer staging and survival and the public health importance attributed to showing that access to routine screening can reduce racial/ethnic differentials in tumor staging.

Conducting a randomized controlled trial of sufficient size to answer these questions was not feasible in our environment. Thus, we elected to explore data from a 9-year observational study designed to examine mammography performance in a community setting. From these data, we assessed the effect of routine screening on the identification of primary breast cancer and examined whether routine screening would have eliminated the excess of late-stage breast cancer found in Black and Hispanic women.

METHODS

Sources of Data

The Colorado Mammography Project longitudinal database (from 1990 to 1998) was used to examine the comparative experience of Black, Hispanic, and White women with respect to tumor stage and histological grade, controlling for education (as a surrogate for socioeconomic status), age, and screening practices. The Colorado Department of Public Health began the Colorado Mammography Project in 1989. As part of the Breast Cancer Screening Consortium, the National Cancer Institute currently funds the Colorado Mammography Project, which collects data from mammography facilities serving women in the 6-county Denver metropolitan area.

Facilities were recruited to the study by the use of outreach project coordinators who explained the project and offered evaluation data for the facility and the radiologists involved. In Denver, single radiologist groups serve multiple mammography facilities. This provides patient population diversity with some consistency in the radiological review. Colorado Mammography Project participation is voluntary for each facility and woman, but participating-facility compliance is nearly 100%. The Colorado Mammography Project

population consists of women who attend these clinics throughout the 6-county metropolitan area, complete personal history forms, and receive a mammogram and radiological report. Because these facilities serve all metropolitan areas and facility participation is driven more by the participating radiologist groups than by the individual patient, no specific biases in patient selection were obvious.

Race/ethnicity information is obtained by self-description. Only women identifying themselves as exclusively Black, Hispanic, or White were selected as study participants. The study area includes approximately 60% of the state of Colorado's female population. The study population is representative of the Denver area in terms of racial/ethnic subgroups and income and educational levels.

In 1999, the Colorado Mammography Project collects as data more than 50% of all mammograms conducted in the study area (T. Byers; 1999 CMAP Coverage Analysis; Department of Preventive Medicine and Biometrics, University of Colorado Health Science Center, Denver; 1999). Through linkages with the statewide population-based Colorado Central Cancer Registry, a total of 5182 breast cancers were identified during the observation period and determined eligible for the study. The Colorado Central Cancer Registry follows Surveillance, Epidemiology, and End Results (SEER) guidelines for collection and reporting of cancer data. Non–Colorado Mammography Project cases were defined as the breast cancer cases of racial/ethnic eligible women residing in the Denver metropolitan area and in the Colorado Central Cancer Registry that were not linked to a Colorado Mammography Project record during the observation period (n=5163).

Case Selection

A major goal of this investigation was to ascertain whether screening can eliminate stage differences found between Black and Hispanic women compared with White women. Few populations have sufficient numbers of routinely screened minorities with detected cancers. To address the 2 types of screening behavior—*episodic screenings,* which are more like prevalent screens, and *routine screenings,* which yield incident cases—we defined 2 types of cases: prevalent and incident. Our assumption for defining incident cases was that a single negative screen sufficiently conveys incident status.

Prevalent cases were defined as women with breast cancer detected on their first screen or women who had an interval of longer than 25 months between the detection mammographic sequence and their prior screen. *Incident cases* were defined as women with at least 1 documented negative mammogram result 10 to 25 months before their primary breast cancer diagnosis (mammogram results were defined as negative based on American College of Radiology BI-RADS codes 1, 2, and 3), excluding the detection mammographic sequence, and no positive mammogram results (American College of Radiology BI-RADS codes 0, 4, or 5).

Breast cancer cases were considered *unconfirmed incident* if a woman had stated on her completed questionnaire or if the radiology report indicated that a previous negative mammogram result existed. No record of a negative result, however, was available in the database for status verification.

Stage of disease was defined with the American Joint Committee on Cancer TNM system. Early-stage disease was defined as TNM stage 0 or I, and late-stage disease was defined as stage II, III, or IV cancer. Histological grade of tumor (grade 1–4) was determined with the *International Classification of Diseases for Oncology.* A tumor not identified as belonging to the original tissues is considered undifferentiated (grade 4) and is the most aggressive grade. The tumor data were obtained from the Colorado Central Cancer Registry records.

Analysis

To characterize the racial/ethnic cohorts, frequencies and cross-tabulations for the variables of interest were generated in both the study case cohort of all incident and prevalent cancers combined and the incident case cohort (Figure 1). Bivariate relationships between racial/ethnic groups and the independent variables were evaluated with a χ^2 test for independence using SAS.[17] A logistic regression model was developed via a backward elimination approach to evaluate the effects of race/ethnicity, education, and other related explanatory variables on the dependent variable—stage of disease at diagnosis. In all comparisons, the cohort of White women was used as the reference cohort. Odds ratios were calculated for each dichotomous and continuous independent variable to test the statistical significance. Because variables were found not to be associated with the results, simpler models were used to assess the differences among racial/ethnic groups, adjusting for the remaining covariates.

The models reported show the effect of various covariates on the resulting dependent variable. The dependent variable, stage of disease, was classified into 2 groups: early and late stage of disease.

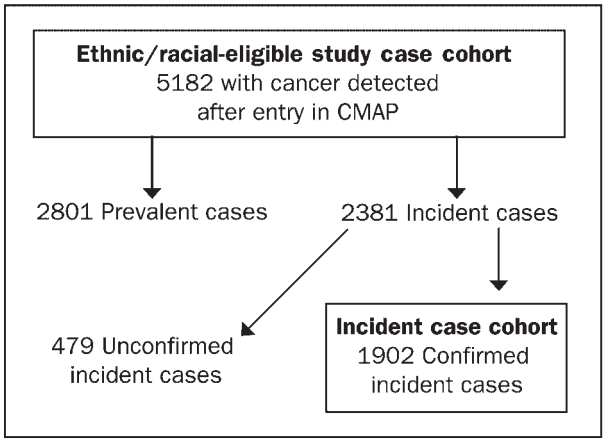


FIGURE 1—Study case selection process: selection of incident cases.

RESULTS

The Colorado Mammography Project database contained approximately 290 500 women who had received at least 1 mammogram through a participating screening facility. In the Colorado Mammography Project, 45% of the Black women, 40% of the Hispanic women, and 48% of the White women underwent at least 2 screens in the system.

The distribution of racial/ethnic groups shows a predominantly White female population. Of the women who reported their race/ethnicity, 83.5% were White, 9.6% were Hispanic, and 3.6% were Black. In Denver, approximately 66.4% of the population who report Hispanic ethnicity are of Mexican American descent.

Figure 1 shows the number of cases in the total study cohort (N=5182) and the distribution into incident (n=2381, or 45.9%) and prevalent cases (n=2801, or 54.1%). Note that the percentage of incident cases in our cohort—45.9%—was only slightly lower than the percentage of women with more than 1 mammogram in the total Colorado Mammography Project registry (55%). Of the 5182 cancers found, 1902 incident cases were confirmed (36.7%). The racial/ethnic breakdown of confirmed incident cases of the total (confirmed incident, unconfirmed incident, and prevalent cases) was as follows: Hispanic women, 89 of 324 (27.5%); Black women, 49 of 149 (32.9%); and White women, 1764 of 4709 (37.5%). Among the total incident cases (2381), 479 (20.1%) were unconfirmed cases, and 79.9% were confirmed.

Results From All Women (Prevalent and Incident Cases)

Stage information was available for 4933 (95.2%) of the 5182 women diagnosed with breast cancer. Among the White women, 63.0% of the cancer cases were detected in early stages (TNM stage 0 or I), compared with 51.4% of the Black women and 54.0% of the Hispanic women (Figure 2). Among the three racial/ ethnic cohorts, there was very little difference in the percentage of stage of disease that was classified as "unknown." White women had a more than 10–percentage point excess of early- over late-stage disease. Conversely, Black (P=.01) and Hispanic women (P=.03) had a statistically significant greater percentage of advanced-stage disease compared with White women and had less difference between late- and earlier-stage disease.

The breast cancer risk factor variables—personal cancer history, higher educational attainment, and age 50 years and older—were correlated with lower staging. These variables, when introduced into the model, had some level of explanatory power.

Black women had a 1.6-fold increased odds (P=.01) of late-stage cancer compared with White women (95% confidence interval [CI]=1.1, 2.2) when the covariates age and cancer history were controlled (Figure 3). In contrast, Hispanic women had a 1.4-fold increased odds (P<.01) of late-stage disease compared with White women (95% CI=1.1, 1.7) when age and cancer history were controlled for in the model (Figure 3).

When education was included in the equation, Hispanic women had a 1.2-fold increased odds (with a confidence interval that includes 1) of late-stage disease; the difference was not statistically significant (Figure 3). This was not the case for Black women, however; when the covariates age, cancer history, and education were controlled, Black women continued to have a l.6-fold increased odds (P=.02) of late-stage disease at diagnosis (95% CI=1.1, 2.3).

After age, cancer history, and education were controlled, there was an approximately 60% increase in late stage of disease among Black women compared with White women when all cancers (prevalent or incident cases) were examined.

Results From the Incidence Cohort

Among the White women, 69% of the confirmed incident cases were early stage (compared with 58% of the prevalent cases); the rates were 63% (vs 50%) for Hispanic women and 54% (vs 49%) for Black women (Table 1). The frequency distribution of early-detection breast cancer, by race/ethnicity, among non–Colorado Mammography Project, Colorado Mammography Project–prevalent, and Colorado Mammography Project–incident cases is shown in Figure 4.

Within each racial/ethnic group's incident cases, the percentage of early-stage cancers was higher in the incident cancer cases than

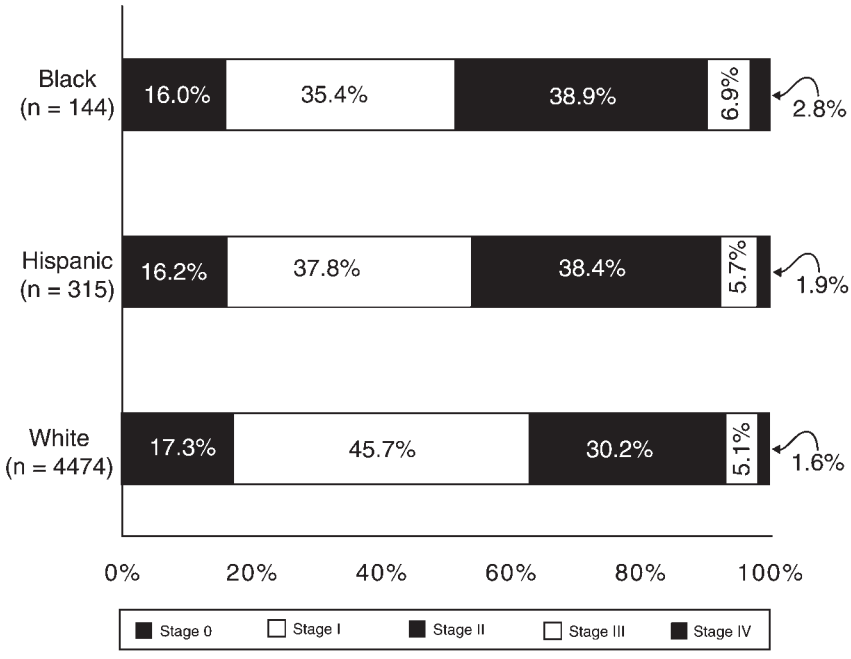
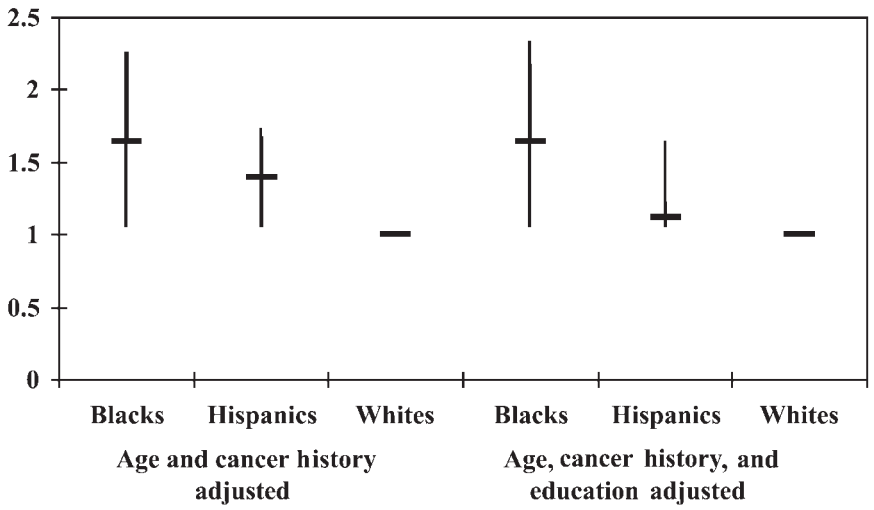


FIGURE 2—Stage of disease at diagnosis, by race/ethnicity.



Note. Total study case cohort: Black (n = 144), Hispanic (n = 315), White (n = 4474).

FIGURE 3—Association between race/ethnicity and late stage of disease (≥ stage II).

TABLE 1—Variable Composition (Percentage), by Race/Ethnicity, for Incident Cases (n = 1902)[a]

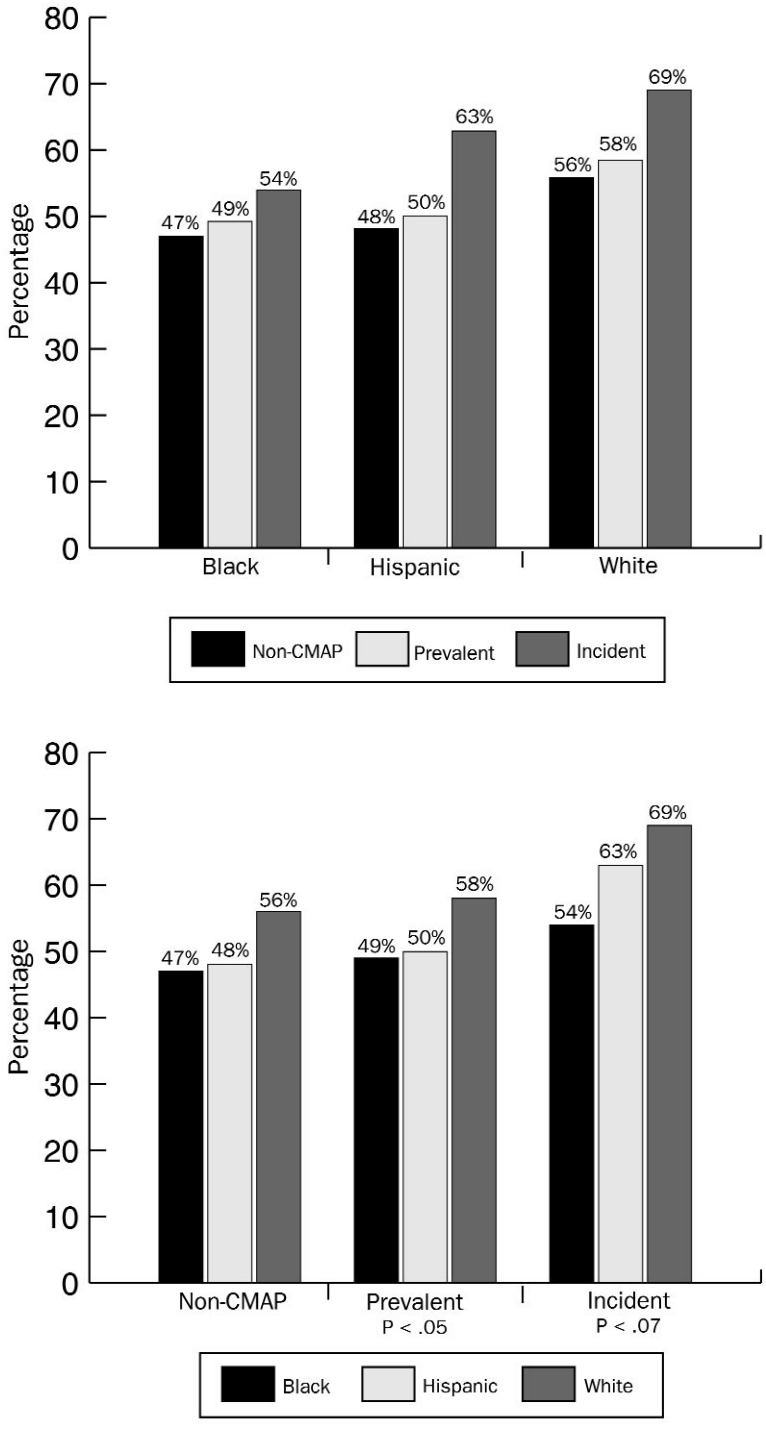
	Hispanic	White	Black
Age group, y			
<50	32.58	20.98	32.65
≥50	67.42	79.03	67.35
Education level			
≤Eighth grade or some high school	17.97	4.04	4.76
High school graduate or some college	49.25	54.99	59.52
College graduate or postgraduate degree	32.83	40.98	35.72
American Joint Committee on Cancer stage			
0	24.14	19.63	13.04
I	39.08	49.13	41.30
II	31.03	26.85	36.96
III	2.30	3.37	6.52
IV	3.45	1.02	2.17
Histological grade			
1 (well)	21.35	17.52	10.20
2 (moderate)	28.09	29.25	26.53
3 (poor)	15.73	18.42	28.57
4 (undifferentiated)	4.49	5.05	10.20
9 (unknown)	30.34	29.76	24.49

Note. CMAP = Colorado Mammography Project.
[a]Not all variables are present for each and every woman.

in the prevalent cases, suggesting that cancer detected under an incident screening protocol, regardless of race/ethnicity, leads to earlier-stage detection. The early-stage cancer differential between incident and prevalent cancers was smallest for Black women. Figure 4 shows for incident cases that the 3 racial/ethnic cohorts differed in regard to early detection (P=.07). There was a 15–percentage point difference between White and Black women in early-stage detection (69% vs 54%) and a 6–percentage point difference between White and Hispanic women (69% vs 63%). These results suggest that after correcting for screening behavior by selecting only the confirmed incident cancer cases, racial/ethnic stage differentials were not eliminated. Whites had higher rates of early-stage cancers in both of the other categories (non–Colorado Mammography Project and Colorado Mammography Project–prevalent cases), whereas Black and Hispanic women were more similar, differing by only 1 percentage point in these other groupings.

With the expectation that the residual differences in early-stage distribution among Black and Hispanic incidence cohorts would be eliminated when covariates were included, a formal logistic model for the adjusted odds ratio of late-stage disease was constructed. When the covariates age and cancer history were introduced into the model, the 95% confidence interval for both Black and Hispanic women overlapped unity (Figure 5). Black women, however, still had a 1.7-fold increased odds of detection of advanced-stage disease compared with White women (P=.08). Therefore, adjusting for age and cancer history did not completely eliminate stage-of-disease differences. These incident cases of Black women yielded an estimated odds ratio almost identical to the unadjusted values and the values based on the entire cohort of women. No statistically significant difference in the adjusted odds ratio was seen for Hispanic women compared with White women (Figure 5).

To further assess these apparent stage differences, we examined the distribution of another dependent variable separate from tumor stage—the histological grade—for early- and late-stage cancer by racial/ethnic group. A slide review was not performed. Specific multivariate modeling with histological grade as the outcome variable was conducted. Because Hispanic and Black women have a higher percentage of late-stage disease, we included stage as a confounding variable in the model.



Note. Non–CMAP = Non–Colorado Mammography Project; Prevalent = Colorado Mammography Project–prevalent; Incident = Colorado Mammography Project–incident.

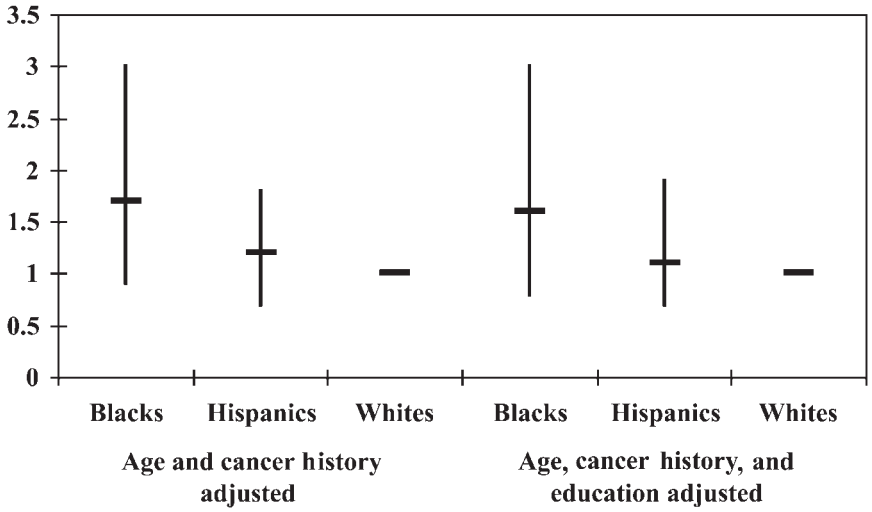
FIGURE 4—Case percentage with early detection (< stage II), by race/ethnicity and by case group.

When the variables of stage of disease, age, and cancer history were controlled in the model, Black women had a 2.2-fold increased odds (P=.03) of higher histological grade (95% CI=1.2, 4.3) compared with White women. No statistically significant difference was found in the adjusted odds ratio for Hispanic women compared with White women. The results confirmed that despite the similar screening histories in these incidence cohorts, and similarity in education level and other variables, Black women in the incident cohort had an excess of higher histological grades (P=.03).

Finally, we reported only on the combined cases (i.e., symptomatic and asymptomatic women). As expected, the percentage of women with later-stage disease was higher among the symptomatic or diagnostically screened women. The patterns, however, were similar. For diagnostic mammograms in the series, the rates of late-stage disease were 44.0% (11 of 25) for the Hispanic women, 42.3% (148 of 350) for the White women, and 50.0% (6 of 12) for the Black women. For asymptomatic or screen-detected women, the rates were 36.2% (17 of 47) for the Hispanic women, 28.0% (330 of 1180) for the White women, and 39.3% (11 of 28) for the Black women.

DISCUSSION

This article examines whether routine screening can be expected to eliminate racial/ethnic disparities in breast cancer staging. We specifically studied proper assignment of a cancer case as incident or prevalent. This study confirmed the results of numerous studies that reported a more advanced stage of disease in Black and Hispanic women at the time of diagnosis.[18–24] This finding is important because researchers consider advanced stage at diagnosis to be a strong contributor to survival differences seen between Black women and Hispanic women.[3,25] It also has been determined in the literature that to a large extent, racial/ethnic differentials seen in staging are the result of lower screening rates.[26–29] This study also has shown that although screening lowers stage at detection, even among women who participate in routine screening and who began the study



Note. Incident case cohort: Black (n = 49), Hispanic (n = 89), White (n = 1764).

FIGURE 5—Association between race/ethnicity and late stage of disease (≥ stage II).

observation period with a documented negative screening result, racial/ethnic stage differentials exist.

Adjusting for measurable risk and socioeconomic factors did not completely eliminate differences in stage of diagnosis by racial/ethnic group. White women were still more likely to have early-stage diagnosis compared with Black and Hispanic women. When routine screening was part of the equation, the odds of advanced stage were reduced, but residual effects still were seen, especially among Black women.

Hispanic women were younger and less educated than White women (2 factors correlated with late-stage disease), yet the difference in advanced-stage disease between Hispanic and White women was smaller than that between Black and White women. Overall, adjustment of individual variables produced modest modifications of the racial/ethnic group–stage association. This was clearly seen when education was added to models. Modifications to the estimates and confidence intervals were minimal to nonexistent. Routine screening may operate differently on stage of disease in Black women compared with Hispanic women. Screening does not appear sufficient to remove the staging differential seen in Black women.

By definition, incident cases reflected a fairly compliant screening population of women. In comparison to White women, no statistically significant difference (P=.07) in the distribution of advanced-stage disease among incident cases of Black and Hispanic women existed. The lack of significance, however, may be due to the decreasing sample size of Black and Hispanic women rather than to a lack of real differences between the groups, because the results parallel those of the total group in magnitude of effect.

Of course, it is difficult to determine whether these findings result from small sample size. The 70% increased odds of late-stage disease in Black women compared with White women, however, seems large enough not to ignore. To further investigate the plausibility of these increased odds of late-stage disease in Black women, we approached the race/ethnicity–screening behavior relationship by examining histological grade. Histological grade is considered a marker for the biological behavior of tumors.[28]

More-advanced histological grade tumors, which are considered more proliferative, were found among Black women who participated in routine screening, even after controlling for stage of disease. We believe that this intriguing result is consistent with the late-stage excess also shown.

This study offers no evidence that the observed remaining differences in tumor histology were due to genetic makeup or imply a biological uniqueness in any one group. The differences seen may be caused by other factors including incomplete control for screening behavior, postmenopausal estrogen exposure or obesity, diet, or subtle measures of socioeconomic status. Investigation of any and all of these factors may provide insight into why we may continue to see racial/ethnic staging differentials in cohorts with similar screening histories. The relation between staging and histological grade among Black women needs to be further explored within a larger cohort, because small numbers limit the interpretive power of this finding.

Routine screening alone may not, however, be sufficient to remove the staging differentials seen among racial/ethnic subgroups of women. Even though the literature suggests that the higher proportions of late-stage disease found in Black and Hispanic women relative to White women is related to these subgroups' greater difficulty in accessing mammography screening and early intervention, our data indicate that this explanation is insufficient—or, alternatively, that perhaps a longer observation period is needed to see the positive effects of screening. Nevertheless, the finding that routine screening does not completely eliminate the staging differentials finding is quite important in refocusing expectations by public health programs regarding the markers of success expectations of the amount of morbidity and mortality that it is possible to prevent.

Limitations

This study had several potential limitations. Education and cancer history depended on study participants' responses; therefore, misclassification was possible because of potential differential recall bias among the groups. We are reasonably confident regarding the generalizability of this cohort sample, given that the database captures more than 50% of the women undergoing mammography in the Denver metropolitan area. Also, the racial/ethnic composition and racial/ethnic educational attainment of Colorado Mammography Project study women were similar to those of the Denver metropolitan population. Colorado Mammography Project women reflect the overall Denver metropolitan population except for their tendency to have been screened.

Although the screening interval for all racial/ethnic cohorts was 10 to 25 months, variation in the average number of months before cancer diagnosis among the racial/ethnic cohorts could contribute to the differences seen. These differences, however, did not explain the staging differences. Interestingly, among all the incident cohorts, additional negative screening results (i.e., over 1) were not associated with lower stages of disease (women who had *more than* 1 negative screening result before the screening in which cancer was detected did not evidence lower stages of disease than women who had *only* 1 negative screen).

Summary

Despite mandatory mammography insurance coverage; free screening programs for low-income women; screening guidelines; national, state, and local public health outreach campaigns; and a public health agenda to eliminate disparities, Black and Hispanic women undergo fewer periodic screenings and have more-advanced disease at diagnosis. Although the research shows that screening leads to improvement across all racial/ethnic groups, a result evidenced in our data as well, the present study—admittedly based on a single study site and a relatively small population—suggests that screening may not completely reduce staging differentials.

Our results may need to be confirmed in a larger cohort, but they illustrate the complex relationship between stage and race/ethnicity. Great strides have been made in the screening of Black and Hispanic women. However, a better understanding of modifiable risk factors and the racial/ethnic beliefs and practices that affect screening is essential. ■

About the Authors

Jillian Jacobellis is with the Colorado Department of Public Health and Environment, Denver, and the Department of Preventive Medicine and Biometrics, School of Medicine, University of Colorado Health Science Center. Gary Cutter is with the Center for Research Methodology and Biometrics, AMC Cancer Research Center, and the University of Nevada, Reno.

Requests for reprints should be sent to Jillian Jacobellis, PhD, MS, Colorado Department of Public Health and Environment, 4300 Cherry Creek Dr S, Denver, CO 80246-1530 (e-mail: jillian.jacobellis@state.co.us).

This article was accepted January 9, 2002.

Contributors

J. Jacobellis planned and designed the study, analyzed the data, and wrote the paper. G. Cutter was a major contributor to the study design and data analysis and contributed to the writing of the paper.

Acknowledgments

We would like acknowledge Monika Baier and Keri Holthaus for their assistance with statistical analysis, Jack Finch for his help in linking the data sets used in the analysis, and Dr Holly Hedegaard for her helpful comments in reviewing the final paper. We also wish to thank the Colorado Central Cancer Registry staff and the Colorado Mammography Project staff for their diligent attention to data quality and Louann Dicus for her technical and editing assistance.

References

1. Colorado Central Cancer Registry. *Cancer in Colorado 1992-1997 Incidence Mortality and Survival.* Denver: Colorado Department of Public Health and Environment; 1999.

2. Delgado DJ, Lin WY, Coffey M. The role of Hispanic race/ethnicity and poverty in breast cancer survival. *P R Health Sci J.* 1995;14:103–116.

3. Eley JW, Hill HA, Chen VW, et al. Racial differences in survival from breast cancer—results of the National Cancer Institute Black/White Survival Study. *JAMA.* 1994;272:947–954.

4. Fox SA, Stein JA, Gonzalez RE, Farrenkopf M, Dellinger A. A trial to increase mammography utilization among Los Angeles Hispanic women. *J Health Care Poor Underserved.* 1998;9:309–321.

5. Elledge RM, Clark GM, Chamness GC, Osborne CK. Tumor biologic factors and breast cancer prognosis among white, Hispanic, and black women in the United States. *J Natl Cancer Inst.* 1994;86:705–712.

6. Weiss SE, Tartter PI, Ahmed S, et al. Ethnic differences in risk and prognostic factors for breast cancer. *Cancer.* 1995;76:268–274.

7. Breen N, Kessler LG, Brown ML. Breast cancer control among the underserved—an overview. *Breast Cancer Res Treat.* 1996;40:105–115.

8. Kerner JF. Breast cancer prevention and control among the medically underserved. *Breast Cancer Res Treat.* 1996;40:1–9.

9. Hahn RA, Stroup DF. Race and ethnicity in public health surveillance: criteria for scientific use of social categories. *Public Health Rep.* 1994;109:7–15.

10. Liu ET. The uncoupling of race and cancer genetics. *Cancer.* 1998;83(suppl):1765–1769.

11. Neale AV. Racial and marital status influences on 10 year survival from breast cancer. *J Clin Epidemiol.* 1994;47:475–483.

12. Wells BL, Horm JW. Stage at diagnosis in breast cancer: race and socioeconomic factors. *Am J Public Health.* 1992;82:1383–1385.

13. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health.* 1992;82:703–710.

14. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. *J Natl Cancer Inst.* 1993;85:1644–1656.

15. Frisell J, Eklund G, Hellstrom L, et al. Randomized study of mammography screening—preliminary report on mortality in the Stockholm trial. *Breast Cancer Res Treat.* 1991;18:49–56.

16. Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *BMJ.* 1988; 297:943–948.

17. *SAS System for Regression.* Cary, NC: SAS Institute Inc; 1991.

18. Ansell D, Whitman S, Lipton R, Cooper R. Race, income, and survival from breast cancer at two public hospitals. *Cancer.* 1993;72:2974–2978.

19. Villar HV, Menck HR. The National Cancer Data Base report on cancer in Hispanics: relationships between ethnicity, poverty, and the diagnosis of some cancers. *Cancer.* 1994;74:2386–2395.

20. Bentley JR, Delfino RJ, Taylor TH, Howe S, Anton-Culver H. Differences in breast cancer stage at diagnosis between non-Hispanic white and Hispanic populations, San Diego County 1988–1993. *Breast Cancer Res Treat.* 1998;50:1–9.

21. Chen F, Trapido EJ, Davis K. Differences in stage at presentation of breast and gynecologic cancers among whites, blacks, and Hispanics. *Cancer.* 1994;73: 2838–2842.

22. Horm JW, Devisa SS, Berhansstipanov L. Cancer incidence, mortality, and survival among racial and ethnic minority groups. In: Schottenfeld D, Fraumeni JF, eds. *Cancer Epidemiology and Prevention.* New York, NY: Oxford University Press; 1996:192–235.

23. Perkins P, Cooksley CD, Cox JD. Breast cancer: is ethnicity an independent prognostic factor for survival? *Cancer.* 1996;78:1241–1247.

24. Mandeblatt J, Andrews H, Lerner J, Zauber A, Burnett W. Determinants of late stage diagnosis of breast cancer and cervical cancer: the impact of age, race, social class, and hospital type. *Am J Public Health.* 1991;81:646–649.

25. Bassett MT, Krieger N. Social class and black–white differences in breast cancer survival. *Am J Public Health.* 1986;76:1400–1403.

26. Zaloznik AJ. Breast cancer stage at diagnosis: Caucasians versus Afro-Americans. *Breast Cancer Res Treat.* 1995;34:195–198.

27. Zaloznik AJ. Breast cancer stage at diagnosis: Caucasians versus Hispanics. *Breast Cancer Res Treat.* 1997;42:121–124.

28. Hunter CP, Redmond CK, Chen VW, et al. Breast cancer: factors associated with stage at diagnosis in black and white women. Black/White Cancer Survival Study Group. *J Natl Cancer Inst.* 1993;85:1129–1137.

29. Richardson JL, Langholz L, Bersnstein C, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. *Br J Cancer.* 1992;65:922–926.

Optimal Indicators of Socioeconomic Status for Health Research

Greg J. Duncan, PhD, Mary C. Daly, PhD, Peggy McDonough, PhD, David R. Williams, PhD

Objectives. In this study we examined the relationship between indicators of socioeconomic status (SES) and mortality for a representative sample of individuals.

Methods. The sample included 3734 individuals aged 45 and older interviewed in 1984 in the Panel Study of Income Dynamics. In the current study, mortality was tracked between 1984 and 1994 and is related to SES indicators of education, occupation, income, and wealth.

Results. Wealth and recent family income were the indicators that were most strongly associated with subsequent mortality. These associations persisted after we controlled for the other SES indicators and were stronger for women than for men and for nonelderly than for elderly individuals.

Conclusions. We found that the economic indicators of SES were usually as strongly associated with mortality as, if not more strongly associated with mortality than, the more conventional indicators of completed schooling and occupation. (*Am J Public Health.* 2002;92:1151–1157)

Although numerous studies have documented the associations between socioeconomic status (SES) indicators and a variety of health outcomes,[1–6] comprehensive indicators of SES are not routinely collected in the United States. In addition, most SES data that are obtained are not reported.[7–9] This data deficiency was highlighted at a 1996 federally sponsored health conference on SES[1,10] and has been noted by the National Committee on Vital and Health Statistics.[11] In both cases, the recommendation was for regular collection of SES data and for the use of SES variables in studies of differential health outcomes.

Despite growing awareness of the need for regular collection of SES indicators, however, there is little agreement on which indicators should be gathered.[12] One problem is that numerous indicators of SES, including occupation,[13] education,[14,3] and household income,[4,5,15,16] have been shown to affect health outcomes, but these indicators are not interchangeable.[12,17–19] Moreover, the impact on health of any particular SES indicator—such as one based on sex and age—varies across different population subgroups.[3,20–22] The fact that various SES indicators may capture different aspects of overall health risk suggests that a systematic examination of the explanatory power of a variety of SES indicators is required before an optimal set of indicators can be recommended.

We contribute to this examination by analyzing the empirical relationship between a set of SES indicators (available from both administrative and survey data sources) and mortality for a nationally representative sample of individuals. We used a unique data set, the Panel Study of Income Dynamics (PSID), to evaluate the predictive power of a variety of SES indicators. Although it includes the traditional SES indicators of education and occupation, our analysis focuses on the relatively neglected economic indicators of SES.

In general, indicators of SES are meant to provide information about an individual's access to social and economic resources. As such, they are markers of social relationships and command over resources and skills that vary over time.[23–24] Among the most frequently used socioeconomic indicators are education and occupation. Economic indicators such as household income and wealth are used less frequently but are potentially as important as or more important than education and occupation. We describe the benefits and drawbacks of each indicator below.

Education is an important determinant of individuals' work and economic circumstances,[25] which are themselves linked to health through specific work conditions and levels of consumption. Education may also be associated with health through its connection to health behaviors. The higher one's level of education, the more likely one is to engage in a range of health-enhancing self-maintenance activities.[26,27] Years of completed schooling are reported with reasonable ease and reliability and are a meaningful indicator of SES for virtually all adults. Because education is typically completed early in adulthood, it serves as a marker of early life circumstances,[28] and no reverse-causation problems result from linking education with health outcomes at older ages. It is for these reasons that the National Center for Health Statistics (NCHS) selected education for inclusion in death certificates in 1989[7] and that the National Committee on Vital and Health Statistics has offered the preliminary assessment that education may be the most useful SES indicator for administrative databases.[11] However, education captures neither the differential on-the-job training and other career investments made by individuals with similar levels of formal schooling nor the volatility in economic status during adulthood that has recently been shown to adversely affect health.[16]

Usual or most recent occupation has long been used as an SES indicator for persons in the workforce, and it can have direct and indirect effects on health. For example, occupation represents exposure to the psychosocial and physical dimensions of work arrangements[29,30] as well as a range of expected earnings and social capital in the form of relative standing or prestige. Indicators of occupational class are widely used in other industrialized countries and have been found to be robust in predicting variations in health status.[17] The National Institutes of Health (NIH) conference on Measuring Social Inequalities in Health called for including occupation as a core SES variable in the US health status re-

porting system.[1] Nevertheless, using occupation as an SES indicator is problematic for subgroups such as teenaged mothers and others with little labor market experience. Moreover, later-career occupations, unlike education, are subject to reverse-causation problems if poor health leads to declines in occupational status.

Household income has been more widely used as an indicator of SES in US studies than in studies undertaken elsewhere. Whereas education and occupation capture individually based dimensions of SES, household income is more indicative of a standard of living and of life chances household members experience through sharing goods and services. The most typical income-based indicator is a household's total cash income, measured over the month, calendar year, or 12-month period before the point of health measurement.[5,15,22] Our examination of "optimal" SES indicators was decidedly empirical and was based on the indicators' sensitivity to mortality risk. We find that economic indicators are considerably more sensitive than traditional ones and suggest that the former should be a standard feature of the US measurement system for monitoring links between SES and health. Although they require difficult-to-obtain tax data, measures of disposable household income—obtained by subtracting from total cash income the taxes households pay—better approximate a household's flow of resources than do measures of total cash income.

One problem with using household income to examine relationships between SES and health is that household members may have unequal access to household income. Specifically, research points to a female disadvantage in resource sharing in households.[31,32] A second problem is that current household income may be an inadequate representation of the standard of living of retired individuals because it may not reflect available financial resources, and it disregards the cumulative effects of a lifetime of deprivation or privilege.[33] Moreover, because current income may be a product of recent health, associations between income and health are subject to reverse-causation problems.

In contrast to income, which consists of a flow of resources over a defined time period, wealth captures the accumulated stock of assets or economic reserves at a given point in time. Income and wealth are positively correlated. For example, wealth is higher for families with histories of higher earnings, lower consumption, more savings and, in some cases, fewer expenditures on health care. But wealth and income are also distinct. For example, elderly individuals frequently have little income but substantial wealth. For most of the US population, wealth is tied up in cars and homes, items for which survey nonresponse bias can be minimized. Several studies in both the United States and the United Kingdom have found that indicators of wealth are related to health, independent of the more traditional indicators of SES.[33–36] Concurrent associations between wealth and health are subject to problems of reverse causation, although perhaps less so than are concurrent associations between income and health, given that accumulating wealth typically takes a long time.

Most health inequalities research undertaken in the United States relies on SES ascertained at a single point in time. Although this measure provides some indication of the relative pattern of health differentials, the cumulative and dynamic nature of socioeconomic structures and experiences is rarely considered. Persistent low income and income volatility may be especially problematic for health,[16] and degree of vulnerability to socioeconomic conditions may vary across the life course.[37] Thus, in assessing the relevance of cross-sectional and longitudinal measures of SES for public policy regarding data collection and reporting, it is important to evaluate the relative utility of these measures.

Finally, although cumulative research points to a robust association between SES and health, the magnitude of the effect of SES on health may vary across social groups. For example, a weaker socioeconomic gradient in mortality has been observed for retired individuals[21,38] and women.[24,39,40] The survival of those with lower levels of health risk, the postponement of morbidity among the socioeconomically advantaged, the universality of certain social programs (e.g., Medicare), and the inadequacy of commonly used indicators of SES to capture the experiences of diverse groups may account for the differential effects of SES by age,[20,41] race/ethnicity,[42,43] and sex.[44]

With a view to providing concrete information that could lead to routinizing the gathering of socioeconomic information in various data-collection modes, we examined the relationship between SES and mortality using data from the PSID. We considered both individual and household indicators of SES as well as the relative merit of short-term vs long-term appraisal of selected indicators. All analyses were stratified by age and sex. Insufficient case counts precluded an additional level of stratification by race/ethnicity.

METHODS

Data

The PSID is an ongoing longitudinal study of a representative sample of individuals living in the United States and of the family units in which they reside. The survey began in 1968 with the most recent mortality follow-up through 1994. The emphasis of the survey is on dynamic aspects of household economic and demographic characteristics, and study staff have been careful to edit and code occupation, earnings, and family income data consistently across waves. Beginning with a representative national sample of households and individuals in 1968, the PSID has collected data on individuals from those households on an annual basis. The initial-wave response rate among sampled dwellings in 1968 was 76%. Attrition was 11% between 1968 and 1969 and has remained between 2% and 3% for each year since 1969. Approximately 55% of the still-living original sample were still participating in the study in the interviewing year 1995. Studies evaluating the national representativeness of the surviving PSID sample at various points (including the 1984 point used to define our sample) have found no significant problems.[45] Probability-of-selection weights are available to adjust for differential nonresponse not related to mortality, as well as for the design-driven unequal selection probabilities of the original sample. This makes it possible to generate estimates that are representative of the US population (omitting some immigrants who arrived after 1968).

Death is recorded in the PSID as a reason for attrition from the sample. In the majority of instances, deaths are reported in the next annual interview by a surviving household member. For individuals who were living alone when last interviewed, information about death comes from a variety of sources, including a surviving contact person, the administrator of the deceased individual's estate, or the post office (via returned mail). Comparisons of PSID data with vital statistics mortality data from the NCHS generally show close agreement.

Sample

The analysis of PSID data is based on 3734 individuals aged 45 and older who participated in the 1984 interview. Deaths among these individuals were tracked between 1984 and 1994. Over this period 298 deaths (11.8%) were recorded for the nonelderly cohort (aged 45 through 64; 67.8% of the sample) and 535 deaths (44.3%) were recorded for the elderly cohort (aged 65 and older; 32.2% of the sample). Mortality was linked with SES indicators by means of Poisson regression models that included additive controls for age in 1984, race/ethnicity (Black vs all other), and sex. Although the relatively small sample available in the PSID precluded estimation of separate models for most demographic subgroups, we did estimate a complete set of models separately for nonelderly men and women (aged 64 years or younger in 1984; n=1091 and 1435, respectively). In all cases, we calculated Huber-White robust standard errors using Stata (version 6.0) to account for the geographically clustered nature of the sample.[46]

We distinguished three kinds of SES indicators: (1) administrative data indicators, which can be collected in most health data, including death and birth certificates; (2) survey indicators that can be collected in a household survey; and (3) exogenous indicators measured a decade or more before the measurement of the health outcome, which are likely free from the serious bias caused by health status affecting SES.

Administrative Data Indicators

Included in our set of readily collected SES indicators were years of completed schooling, most recent occupation, and total family income. Several of the PSID's interviewing waves included a direct question about completed schooling; we took the most recent report before the 1984 interview. Descriptive information about this and all other indicators is presented in Table 1.

The PSID asked for information about occupation whenever a respondent reported being employed at the time of the interview

TABLE 1—Descriptive Statistics for Analysis Variables[a]

	Age 45-64			Age ≥65
	Total	Men	Women	Total
Administrative Data Indicators				
Education				
N[b]	2526	1091	1435	1208
≤8 years	0.21 (0.411)	0.24 (0.425)	0.20 (0.399)	0.41 (0.493)
9-11 years	0.21 (0.409)	0.19 (0.395)	0.23 (0.418)	0.16 (0.371)
12 years	0.35 (0.478)	0.28 (0.451)	0.41 (0.491)	0.26 (0.439)
≥13 years	0.21 (0.410)	0.28 (0.449)	0.16 (0.370)	0.15 (0.358)
Last Occupation				
N[b]	2369	1075	1294	919
Farmer	0.02 (0.129)	0.03 (0.179)	na	0.02 (0.147)
Service	0.15 (0.359)	0.07 (0.258)	0.21 (0.410)	0.05 (0.209)
Laborer	0.03 (0.176)	0.06 (0.244)	0.01 (0.091)	0.02 (0.151)
Operative	0.17 (0.378)	0.16 (0.370)	0.18 (0.384)	0.23 (0.422)
Crafts	0.09 (0.292)	0.21 (0.405)	0.01 (0.087)	0.05 (0.227)
Sales	0.04 (0.193)	0.04 (0.196)	0.04 (0.19)	0.02 (0.139)
Clerical	0.13 (0.338)	0.05 (0.217)	0.19 (0.395)	0.09 (0.289)
Manager	0.10 (0.303)	0.16 (0.368)	0.06 (0.232)	0.04 (0.185)
Professional	0.09 (0.293)	0.12 (0.328)	0.07 (0.261)	0.06 (0.243)
1983 Pre-Tax Family Income				
N[b]	2526	1091	1435	1208
Overall Mean	34 233 (32 467)	39 715 (34 337)	30 065 (30 327)	19 252 (18 242)
Survey-Based Indicators				
Total Wealth				
N[b]	2526	1091	1435	1208
Overall Mean	118 356 (345 277)	135 801 (369 394)	105 093 (325 254)	114 855 (424 458)
1979-1983 Post-Tax Family Income				
N[b]	2526	1091	1435	1208
Overall Mean	29 040 (21 338)	32 644 (21 825)	26 301 (20 549)	19 289 (29 755)
1979-1983 Family Income-to-Needs				
N[b]	2526	1091	1435	1208
Overall Mean	3.7 (3.4)	4.2 (3.5)	3.4 (3.3)	3.0 (4.9)
"Exogenous" Indicators				
1969-1975 Post-Tax Family Income				
N[b]	2526	1091	1435	1208
Overall Mean	26 433 (14 480)	28 064 (14 177)	25 193 (14 590)	20 110 (14 181)
1969-1975 Family Income-to-Needs				
N[b]	2526	1091	1435	1208
Overall Mean	2.7 (1.9)	2.9 (1.9)	2.6 (2.0)	3.0 (2.5)

[a]Unweighted mean (standard deviation).
[b]Number of nonmissing observations on analysis variable.
Source: Authors' calculations based on the Panel Study of Income Dynamics.

or in the calendar year preceding it; again, we took the most recent report before the 1984 interview. Questions used to determine occupation were identical to those asked in Census Bureau surveys, and responses were coded to the 1970 US Census occupational classifications. The ordinal scale used in the inequality index method (discussed below) was based on the following ranking of occupations: professional, managerial, clerical, sales, crafts, operatives, laborers, service, and farmer. This ordering follows the pattern of mortality risk across occupations reported by Moore and Hayward.[47]

Total household income comes from a series of questions asked in the 1984 interview about income received by all family members during the calendar year 1983. The detailed nature of the questions is likely to yield more reliable measurement of income than would be obtained from a single question. In contrast to the case for other income-based indicators described in the next paragraph, we do not subtract taxes from household income, given that the data required for such an adjustment are not likely to be available in administrative data sources. Similarly, it may not be feasible to collect sufficiently high-quality income information from death certificates.

Survey-Based Indicators

Our list of survey-based indicators consists of SES indicators that can be collected in a cross-sectional or short-run longitudinal household survey. Our household-income indicator averages reports of household income over the 5 calendar years between 1979 and 1983. We inflated all dollar-based indicators in our analysis to 1984 price levels using the CPI-UX1 component of the Consumer Price Index. To approximate disposable household income, we subtracted federal income taxes and social security taxes from the household's total cash income.

We obtained a household size–adjusted indicator of household income by dividing an individual's household income by a Census Bureau–based poverty threshold that accounts for family size. For example, in 1999, the poverty threshold for a family of 4—2 adults and 2 children—was $16 895. An individual with that level of household income would have an "income-to-needs" ratio of 1.0; an income of $33 790 would produce a ratio of 2.0. We constructed a measure of household wealth at the time of the 1984 interview from a sequence of questions designed to gather comprehensive information about the assets and liabilities of the household.

Exogenous Indicators From Long-Term Prospective Studies

That SES may reflect rather than cause health status is a persistent problem for studies of SES–health linkages. The PSID data span a long period dating back to 1968 and thus provide researchers with SES measurements from a decade or more before the period over which mortality is measured. Our strategy for compiling a set of exogenous SES indicators was to measure everything before the 1976 interview and to adjust our regression estimates to reflect whether individuals reported health limitations during that interview. Our 2 exogenous indicators were household income and family size–adjusted household income, both averaged over the years between 1967 and 1975. To minimize the possible effects of health selection in the analysis, we also controlled for disability, as defined by a 1976 self-reported response to the question, "Do you have any physical or nervous condition that limits the type of work or the amount of work that you can do?"

We related SES indicators to mortality by creating indices of inequality based on each of our socioeconomic indicators[28,48] and estimating the relationship between these indices and mortality using Poisson regression. In analyses not included here, we also used Cox regression models to estimate the relative mortality risk of individuals in the bottom vs the top deciles of the income and wealth distributions. Results were similar to those reported here.

Following Pamuk,[49] Kunst and Mackenback,[48] and Smith et al.,[28] we created indices for each of our SES indicators by assuming that the SES of a group (e.g., those who did not complete high school) is determined by the group's relative position in the hierarchy for that indicator (e.g., education). Thus, the socioeconomic position of each group is assigned a value between 0 and 1 based on the proportion of the population with a higher position on the SES indicator than the midpoint of the given group. For example, if 10% of the population were in the highest educational group, the relative position of its members would be between 0 and 0.10, the midpoint being 0.05. If the next group contained an additional 16% of the population, this group would be assigned an index value of 0.18 (=0.1 + [0.16/2]). We calculated an index of this type for each of our SES indicators.

We used Poisson (log-linear) regression to examine the relation of numerical indicators of SES to mortality. Coefficients transformed by (expβ) were used to show the relative risk of mortality for those at the bottom of the social hierarchy compared with those at the top. Following Smith et al.,[28] we referred to the relative inequality index as RII.

RESULTS

Table 2 presents Poisson regression–based risk estimates and 95% confidence intervals for our various SES indicators. We obtained the estimates in each cell in the table from regressions containing age, race/ethnicity, and the indicated SES indicator only. In cases where the male and female samples are combined, the regressions also control additively for sex.

The first three rows of Table 2 reveal that the inequality index method produces a significant mortality association for occupation and family income, but not education, in the nonelderly cohort. (Because all SES contrasts are reverse scaled, a risk estimate that exceeds 1.0 indicates higher mortality associated with lower SES levels.) Breaking the nonelderly cohort down by sex produces somewhat higher SES-related mortality risks for women than for men (the exception being occupation), but only in the case of women's 1983 pretax family income is the risk estimate statistically significant at conventional levels. Only the family income administrative data indicator had significant mortality effects among individuals in the elderly cohort.

Results for the survey-based SES indicators (all of which measure economic resources) are presented in the fourth through sixth rows of Table 2. For the nonelderly cohort, all SES indicators were significantly associated with mortality, with relative rates of mortality standing at about 3.0. In all cases, these rates

TABLE 2—Age-, Sex-, and Race/Ethnicity-Adjusted Relative Rates of Mortality According to Administrative, Survey, and Exogenous Data Indicators[a]

	Aged 45–64			Aged ≥65
	Total	Men	Women	Total
Administrative data indicators				
Education RII	1.59 (0.96, 2.64)	1.67 (0.89, 3.10)	2.04 (0.83, 5.02)	1.48 (0.95, 2.31)
Occupation RII	2.34 (1.19, 4.57)*	2.37 (1.00, 5.63)*	2.01 (0.71, 5.70)	0.73 (0.38, 1.39)
1983 pre-tax family income RII	3.46 (2.07, 5.78)*	1.64 (0.77, 3.50)	3.87 (1.89, 7.93)*	1.58 (1.17, 2.14)*
Survey data indicators				
Wealth RII	2.86 (1.50, 5.45)*	2.51 (1.15, 5.44)*	4.51 (1.84, 11.0)*	2.05 (1.48, 2.85)*
1979–1983 post-tax family income RII	2.95 (1.67, 5.20)*	1.68 (0.82, 3.44)	4.60 (2.20, 9.64)*	1.50 (1.12, 2.00)*
1979–1983 family income-to-needs RII	3.04 (1.63, 5.68)*	1.74 (0.73, 4.16)	3.68 (1.54, 8.78)*	2.06 (1.34, 3.19)*
Exogenous indicators				
1969–1975 post-tax family income RII	1.94 (0.93, 4.08)	1.26 (0.62, 2.54)	3.97 (1.61, 9.79)*	1.38 (1.04, 1.85)*
1969–1975 family income-to-needs RII	1.95 (1.01,3.75)*	1.49 (0.72,3.06)	2.31 (0.96, 5.53)	1.47 (1.06, 2.03)*

Note. Calculations are the authors' and are based on the Panel Study of Income Dynamics. CI = confidence interval.
[a]CIs are in parentheses.
*Significant at the .05 level.

were higher for women than for men. As a final generalization, the rates were universally smaller for the elderly than for the nonelderly cohort. Comparing income indicators in the administrative and survey categories reveals that the relative mortality rates are not substantially affected either by lengthening the accounting period from 1 year to 5 years or by adjusting family income for family size.

Results for the final, exogenous set of indicators were drawn a decade or more before the beginning of the interval over which mortality is assessed (final two rows of Table 2). To enhance our efforts to assess the exogenous effects of these SES components, we introduced an additive control for 1976 self-reported work limitations into all the regressions. By and large, measuring the economic indicators of SES in the late 1960s and early 1970s reproduced the patterns found for the economic indicators measured in the late 1970s and early 1980s. Household income, in this case averaged over the years between 1967 and 1975, continued to be more strongly associated with subsequent mortality for women than for men and for the nonelderly than for the elderly cohort.

Risk ratios for the elderly cohort were evident for both the more distant and the more recent SES indicators, suggesting that economic status before and during retirement is an important determinant of postretirement health. The fact that the risk ratios are never as strong for the elderly cohort as they are for the nonelderly cohort suggests that some of the health effects of SES may take the form of survival until age 65.

To assess which SES indicators retain their explanatory power in the presence of controls for other SES indicators, we undertook the series of regressions summarized in Table 3. In all cases we used the inequality index method and controlled for sex, age, and race/ethnicity. The first row shows the relative risk (1.59) associated with education for the nonelderly cohort and includes our demographic controls but no other SES indicators. By design, this estimate is identical to the one presented in the first row and column of Table 2. In contrast, the relative risk shown in the second row (.56) includes controls for occupation and for posttax family income in the years between 1979 and 1983. In neither case are these estimates significantly different from 1.0 at conventional levels. The occupation effect is also reduced in the presence of economic controls, but the associations of income and wealth with mortality are not diminished after controlling for occupation and education. A qualitatively similar result applies to the elderly cohort, although the absolute levels of the risk ratios are considerably lower than are those for the nonelderly cohort.

DISCUSSION

Our examination of "optimal" SES indicators was decidedly empirical and was based on the indicators' sensitivity to mortality risk. We find that economic indicators are considerably more sensitive than traditional ones and suggest that the former should be a standard feature of the US measurement system for monitoring links between SES and health.

Our objective here was to enumerate alternative SES indicators and to assess their associations with mortality using prospective data from a nationally representative survey. Although we found some SES–mortality gradients for education and occupation, the most powerful associations were seen for the economic indicators—wealth and family income. These associations were generally stronger for women than for men and for the nonelderly cohort than for the elderly cohort. They diminished little when they were measured 9 to 15 years before the mortality observation window or in the presence of adjustments for education and occupation.

One of our most striking findings was the high mortality risk for women with low fam-

TABLE 3—Multivariate Analysis of Age-, Sex-, and Race/Ethnicity -Adjusted Relative Rates of Mortality According to Administrative and Survey Data Indicators (95% CI)[a]

Variables Used in Regression[b]	Age, Sex, and Race/ethnicity	Education	Occupation	1979–1983 Posttax Family Income	Relative Risk Index RII (95% CI)
Age 45–64					
Education	x				1.59 (0.96, 2.64)
Education	x		x	x	0.56 (0.25, 1.23)
Occupation	x				2.34 (1.19, 4.57)*
Occupation	x	x		x	1.67 (0.61, 4.55)
Wealth	x				2.86 (1.50, 5.45)*
Wealth	x	x	x		2.75 (1.40, 5.41)*
1979–1983 posttax family income	x				2.95 (1.67, 5.20)*
1979–1983 posttax family income	x	x	x		3.58 (1.78, 7.18)*
Age ≥65					
Education	x				1.48 (0.95, 2.31)
Education	x		x	x	1.05 (0.66, 1.69)
Occupation	x				0.73 (0.38, 1.39)
Occupation	x	x		x	0.41 (0.15, 1.15)
Wealth	x				2.05 (1.48, 2.85)*
Wealth	x	x	x		1.92 (1.38, 2.66)*
1979–1983 post-tax family income	x				1.50 (1.12, 2.00)*
1979–1983 post-tax family income	x	x	x		1.27 (0.89, 1.79)

Note. Calculations are the authors' and are based on the Panel Study of Income Dynamics. CI = confidence interval.
[a]CIs are in parentheses.
[b]x indicates variables included in regression.
*Significant at the .05 level

ily incomes during their preretirement years. This finding is in stark contrast to those of other studies that have reported a weaker socioeconomic gradient in mortality for women, particularly when education and occupational class are considered,[24] and it supports the argument that those indicators may not adequately capture women's SES. For example, women receive lower income returns from education than do men, and occupational classification systems based on the characteristics of male-dominated occupations do not capture the hierarchy of women's occupations.[39] Further research is needed to elucidate the ways in which the material and symbolic dimensions of SES differentially affect the health of men and women.

Although our analysis distinguished SES–mortality associations between elderly and nonelderly cohorts and was able to measure some components of SES more than a decade before the mortality observation window, we only began to exploit the potential of a life-course-analysis perspective on links between SES and mortality. A more complete analysis would better examine the ways in which socioeconomic resources are acquired through training and lost through failing health.

Our results suggest that economic components of SES should be a standard feature of the measurement system for monitoring links between SES and health. It is feasible to gather reasonably valid information about income and wealth in surveys without compromising response rates.[50] Efforts to gather this information as part of collecting administrative data may be more difficult. However, the far superior explanatory power of income- and wealth-based indicators of SES as compared with the more conventional indicators of education and occupation suggests the value of methodological efforts to support the collection of economic indicators as part of administrative data systems. ■

About the Authors

Greg Duncan is with Northwestern University. Mary Daly is with the Federal Reserve Bank of San Francisco. At the time of the study, Peggy McDonough was with York University. David Williams is with the University of Michigan.

Requests for reprints should be sent to Greg J. Duncan, Institute for Policy Research, Northwestern University, 2040 Sheridan Road, Evanston, IL 60208 (e-mail: greg-duncan@northwestern.edu).

This article was accepted April 18, 2001.

Contributors

All the authors contributed to the design of the study and the writing of the article. Mary Daly analyzed the data.

Acknowledgments

We are grateful for financial support from the Robert Wood Johnson Foundation, the Investigator Awards in Health Policy Research (David Williams, PI, grant 026422), the National Institute on Aging (Benjamin Amick, PI, grant R01AG13036–01A1; George Kaplan, PI, grant P50-HD38986 Michigan Interdisciplinary Center on Social Inequalities), and the John D. and Catherine T. MacArthur Foundation Research Network on Socioeconomic Status and Health.

We also appreciate the comments of Benjamin Amick and two referees. Louise Tench provided research assistance, and Carol D'Souza and Anita Todd provided editorial assistance.

Note. The views expressed here are those of the authors and should not be attributed to the Federal Reserve Bank of San Francisco or the Federal Reserve System.

References

1. Moss N, Krieger N. Measuring social inequalities in health. *Public Health Rep.* 1995;110:302–305.

2. Adler NE, Boyce T, Chesney MA, Folkman S, Syme SL. Socioeconomic inequalities in health: no easy solution. *JAMA.* 1993;269:3140–3145.

3. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989;129:919–933.

4. Duleep HO. Measuring socioeconomic mortality differentials over time. *Demography.* 1989;26: 345–351.

5. Pappas G, Queen S, Hadden W, Fisher G. The Increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. *N Engl J Med.* 1993;329:103–109.

6. Hahn RA, Eaker ED, Barker ND, Teutsch SM, Sosniak WA, Krieger N. Poverty and death in the United States–1973 and 1991. *Epidemiology.* 1995;6: 490–497.

7. Krieger N, Chen JT, Ebel G. Can we monitor socioeconomic inequalities in health? A survey of US Health Departments' data collection and reporting practices. *Public Health Rep.* 1997;112:481–491.

8. Williams DR. Missed opportunities in monitoring socioeconomic status. *Public Health Rep.* 1997;112: 492–494.

9. Navarro V. Race or class versus race and class: mortality differentials in the United States. *Lancet.* 1990;336:1238–1240.

10. Recommendations of the conference on Measuring Social Inequalities in Health. *Int J Health Serv.* 1996;26:521–527.

11. National Committee on Vital and Health Statistics. *Diversity and Health Care Data Base: A Preamble for Policymakers.* Washington, DC: US Dept of Health and Human Services; 1993. DHHS publication PHS 93-1205:3–19.

12. Winkleby M, Jatulis D, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health.* 1992;82:816–820.

13. Kotler P, Wingard D. The effect of occupational, marital, and parental roles on mortality: the Alameda County study. *Am J Public Health.* 1989;79:607–612.

14. Lahelma E, Valkonen T. Health and social inequities in Finland and elsewhere. *Soc Sci Med.* 1990; 31:257–265.

15. Kitagawa EM, Hauser PM. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology.* Cambridge, Mass: Harvard University Press; 1973.

16. McDonough P, Duncan GJ, Williams DR, House J. Income dynamics and adult mortality in the US, 1972–1989. *Am J Public Health.* 1997;87:1476–1483.

17. Krieger N, Williams DR, Moss N. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health.* 1997;18: 341–378.

18. Williams DR, Collins C. US socioeconomic and racial differences in health. *Annual Review of Sociology.* 1995;21:349–386.

19. Liberatos P, Link B, Kelsey J. The measurement of social class in epidemiology. *Epidemiol Rev.* 1988;10: 87–121.

20. House JS, Kessler RC, Herzog AR, Mero R, Kinney A, Breslow M. Age, socioeconomic status, and health. *Milbank Q.* 1990;68:383–411.

21. House JS, Lepowski JM, Kinney AM, Mero RP, Kessler RC, Herzog AR. The social stratification of aging and health. *J Health Soc Behav.* 1994;35: 213–234.

22. Elo IT, Preston SH. Educational differences in mortality: United States 1979–1985. *Soc Sci Med.* 1996;42:47–57.

23. Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav.* 1995(extra issue):80–94.

24. Macintyre S, Hunt K. Socio-economic position, gender and health: how do they interact? *Journal of Health Psychology.* 1997;2:315–334.

25. Psacharoupoulos G. Returns to education: a further international update and implications. *Journal of Human Resources.* 1985;20:583–604.

26. Lynch JW, Kaplan GA, Salonen JT. Why do poor people behave poorly? Variations in adult health behaviours and psychosocial characteristics by stages of the socioeconomic life course. *Soc Sci Med.* 1997;44: 809–819.

27. Ross C, Wu C. The links between education and health. *American Sociological Review.* 1995;60: 719–745.

28. Davey Smith G, Hart C, Hold D, et al. Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Community Health.* 1998;52:153–160.

29. Johnson JV, Stewart W, Hall EM, Fredlund P, Theorell T. Long-term psychosocial work environment and cardiovascular mortality among Swedish men. *Am J Public Health.* 1996;86:324–331.

30. Karasek RA, Theorell T. *Healthy Work.* New York, NY: Basic Books; 1990.

31. Volger C, Pahl J. Money, power and inequality within marriage. *The Sociological Review.* 1994;42: 263–288.

32. Pahl J. Household spending, personal spending, and the control of money in marriage. *Sociology.* 1990; 24:119–138.

33. Robert S, House JS. SES differentials in health by age and alternative indicators of SES. *J Aging Health.* 1996;8(3):359–388.

34. Mare RD. Socio-economic careers and differential mortality among older men in the United States. In: Vallin J, D'Souza S, eds. *Measurement and Analysis of Mortality—New Approaches.* Oxford, England: Clarendon Press; 1990:362–387.

35. Smith JP, Kington R. *Demographic and Economic Correlates of Health in Old Age.* Santa Monica, Calif: RAND; 1996. Laboratory Population Progress Working Paper Series, No. 96–04.

36. Filakti H, Fox J. Differences in mortality by housing tenure and by car access from the OPCS Longitudinal Study. *Popul Trends.* 1995;(81):27–30.

37. Wadsworth ME. Health inequalities in the life course perspective. *Soc Sci Med.* 1997;44:859–869.

38. Beckett M. Converging health inequalities in later life—an artifact of mortality selection. *J Health Soc Behav.* 2000;41:106–119.

39. McDonough P, Williams DR, House J, Duncan GJ. Gender and the socioeconomic gradient in mortality. *J Health Soc Behav.* 1999; 40:17–31.

40. Backlund ESPD, Johnson NJ. A comparison of the relationships of education and income with mortality: the National Longitudinal Mortality Study. *Soc Sci Med.* 1999;49:1173–1384.

41. Kaplan G, Seeman T, Cohen R, Knudsen L, Guralnik J. Mortality among the elderly in the Alameda County Study: behavioral and demographic risk factors. *Am J Public Health.* 1987;77:307–312.

42. Williams DR, Takeuchi D, Adair R. Socioeconomic status and psychiatric disorder among blacks and whites. *Social Forces.* 1992;71:179–194.

43. Kessler RC, Neighbors HW. A new perspective on the relationships among race, social class, and psychological distress. *J Health Soc Behav.* 1986;27:107–115.

44. Krieger N, Rowley DL, Herman AA, Avery B, Phillips MT. Racism, sexism, and social class: implications for studies of health, disease, and well-being. *Am J Prev Med.* 1993;9(suppl 6):82–122.

45. Fitzgerald J, Gottschalk P, Moffitt R. An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics. *Journal of Human Resources.* 1998;23:251–299.

46. *Stata. Version 6.0* [computer program]. College Station, Tex: Stata Corp; 1999.

47. Moore DE, Hayward MD. Occupational careers and mortality of older men. *Demography.* 1990;27: 31–53.

48. Kunst AE, Mackenback JP. The size of mortality differences associated with educational level in nine industrialized countries. *Am J Public Health.* 1994;84: 932–937.

49. Pamuk ER. Social class and inequality in mortality from 1921 to 1972 in England and Wales. *Popul Stud (Camb).* 1985;39:17–31.

50. Duncan GJ, Petersen E. The long and the short of asking questions about income, wealth, and labor supply. *Soc Sci Res.* 2001;30:248–263.

Checking the Pulse: Midwestern Reporters' Opinions on Their Ability to Report Health Care News

Melinda Voss, MPH

Objectives. Newspapers play a key role in disseminating information and shaping perceptions about health, research, and policies. Inadequate or misleading reporting constitutes a public health threat that can jeopardize individual health and lead to harmful health policies.

Methods. Surveys were mailed to 165 reporters at 122 newspapers in 5 Midwest states. The association of training, newspaper size, and experience with reporter's self-perceived reporting ability was assessed.

Results. The response rate was 69.6% (115/165). Between 66% and 85% of the reporters assessed 4 tasks vital to sound health reporting as "sometimes difficult" to "nearly always difficult." No significant differences in perceived ability were found by training or newspaper size. Respondents with less experience reported higher perceived ability.

Conclusions. These findings show that reporters may have difficulty understanding complex health issues and interpreting statistics because they are inadequately trained. (*Am J Public Health.* 2002;92:1158–1160)

The mass media provide important information to the public about research, policies, and the health business,[1–3] yet a study of science reporters found that 77% acknowledged that they do not understand the complexities of scientific subjects.[2] Reporters have been criticized for careless, inadequate, or unfair coverage,[4–13] and even reporters themselves have criticized the quality of health news.[14–16] Inadequate, misleading, or incomplete news reporting constitutes a public health threat. Such reporting can lead people to make misguided choices that may put their health at risk or influence policymakers to adopt inadequate or harmful laws, regulations, or policies.

Traditionally, reporters are not trained in the subjects they cover, although evidence indicates that they want such training.[16–19] For health reporters, vital skills may include understanding complex health issues, finding reliable sources, placing research findings in context, and producing balanced, thorough stories on tight deadlines, as well as interpreting statistics.

This study examines health reporters' perceptions about their ability to report health news; additionally, it may help public health professionals, who need to become more aware of reporters' constraints, to aid and work with the media to ensure quality reporting.

METHODS

Participants were health reporters at daily newspapers in 5 Midwestern states (Minnesota, Iowa, South Dakota, North Dakota, and Wisconsin). Reporters at 122 newspapers identified by Burrelle's Media Directory (Livingston, NJ; July 1998) received surveys. Identification of individuals surveyed occurred through telephone conversations with reporters or editors. Surveys were anonymous. A return of 115 completed surveys and 1 uncompleted survey from 165 surveys mailed yielded a response rate of 115/165, or 69.6%.

Researchers asked participants how often they report health news, from full-time to rarely, as well as how much they report on health policy, medical research, consumer health, health business, public health, or other related topics. Participants rated their ability in 5 skill areas (e.g., understanding key health issues) on a 5-point Likert scale ranging from 1 (nearly always easy to do) to 5 (nearly always difficult to do). The Cronbach α coefficient for the 5-item measure was .67. This measure has not been independently validated for reliability.

Participants rated their interest in covering health issues on a 4-point Likert scale (1 = very strong: I want to cover it indefinitely; 4 = weak: I don't like it and would prefer another beat). Questions about training for reporting health news employed yes-or-no answers. Researchers did not define training or ask respondents to do so. The validity of this measure has not been independently assessed.

For questions about confidence in reporting health news and interpreting statistics, respondents' ratings could range from 1 (not at all confident) to 4 (very confident). Finally, participants' agreement with 9 statements assessing their perceptions of health news (e.g., news media who cover health concentrate too much on spot news) was measured on a 5-point Likert scale (1 = strongly agree; 5 = disagree strongly). Demographic variables were self-reported.

Frequencies were examined to assess need for training, attitudes about covering health news, and perceptions of the quality of health news. Differences in perceived ability by training (training vs no training), newspaper size (>25 000 vs <25 000 weekday circulation), and health reporting experience (>5 years vs <5 years) were assessed. *t* tests were conducted with perceived ability as a continuous variable. *P* values of <.05 were considered significant.

RESULTS

About three quarters (77.4%) of respondents (see Table 1 for their demographic characteristics) stated that they reported on health half or less than half of the time, whereas about 14% covered health full-time. Nearly 83% (n = 94) reported having received no training for covering health news. Of those, about 73% (n = 62) said that training would be helpful. Nearly 84% (n = 96) reported hav-

TABLE 1—Characteristics of Survey Respondents (N = 115)

	% (n)
Gender	
Female	60.5 (n = 69)
Male	39.5 (n = 45)
Age	
20–30	21.1 (n = 24)
31–40	25.4 (n = 29)
41–50	35.1 (n = 40)
> 50	18.4 (n = 21)
Experience in journalism	
< 2 years	4.4 (n = 5)
> 2 years but < 5 years	18.6 (n = 21)
> 5 years but < 10 years	13.3 (n = 15)
> 10 years but < 20 years	31.9 (n = 36)
> 20 years	31.9 (n = 36)
Experience covering health	
< 2 years	22.8 (n = 26)
> 2 years but < 5 years	22.8 (n = 26)
> 5 years but < 10 years	21.9 (n = 25)
> 10 years	32.5 (n = 37)
Race/ethnicity	
White	95.6 (n = 109)
African American	0.9 (n = 1)
American Indian	0.0 (n = 0)
Asian or Pacific Islander	2.6 (n = 3)
Hispanic	0.0 (n = 0)
Other	0.9 (n = 1)
Highest level of education completed	
High school or equivalent	1.8 (n = 2)
Some college or technical training	10.5 (n = 12)
Two-year degree or certificate	4.4 (n = 5)
Bachelor's degree	66.7 (n = 76)
Master's degree	15.8 (n = 18)
PhD	0.9 (n = 1)
Weekday circulation of newspaper	
< 10 000	25.0 (n = 28)
> 10 001–25 000	25.9 (n = 29)
> 25 001–50 000	17.9 (n = 20)
> 50 001–100 000	12.5 (n = 14)
> 100 001–250 000	8.0 (n = 9)
> 250 000	10.7 (n = 12)

TABLE 2—Self-Perceived Reporting Ability by Category (N = 115)

		Perceived Ability, %				
		Easy				Difficult
Reporting Task	Mean Score (SD)	1	2	3	4	5
Finding reliable sources	2.10 (0.87)	**26.1**	**43.5**	25.2	4.3	0.9
Understanding key health issues	3.16 (0.78)	1.8	15.9	**49.7**	**31.0**	2.7
Putting health news in context	3.30 (1.00)	5.3	**26.3**	**47.4**	18.4	2.6
Producing balanced stories on deadlines	3.38 (1.11)	1.7	**31.3**	**43.5**	18.3	5.2
Interpreting statistical data	3.64 (1.37)	1.8	13.3	**51.3**	**27.4**	6.2
Overall perceived ability (n = 115)	3.05 (0.60)					

Note. 1 = Nearly always easy to do, 2 = Usually easy to do, 3 = Sometimes easy to do, sometimes difficult, 4 = Often difficult to do, 5 = Nearly always difficult to do. Bold percentages show where the majority of responses fall.

ing received no training in interpreting health statistics. Of those, nearly 68% (n=80) said that training would be helpful.

Fifty-one percent of respondents said that they had a strong interest in covering health. Of those, roughly half said that they wanted to cover health indefinitely. Another 44.1% said that they liked covering health but could easily shift to another subject. Reporters at newspapers with weekday circulation under 25 000 had a mean score of 2.34 (SD=0.91) in rating their interest in covering health, and reporters at larger newspapers had a mean score of 2.22 (SD=0.94). A *t* test showed no significant difference (P=.51, t=.65, df=108).

Table 2 shows percentages of perceived ability in 5 categories. About one-third of respondents said that understanding key health issues and interpreting health statistics were often or nearly always difficult to do. Putting health news in context was nearly always easy or usually easy to do for about 32%.

Nearly three quarters of respondents said that they were moderately or very confident in their ability to report h0ealth news. Another 25% said that they were not at all or were somewhat confident. Less than half said that they were moderately or very confident about reporting statistics. In rating confidence in health-reporting ability, small-newspaper reporters (n=63) had a mean score of 3.12 (SD=0.77), and respondents from larger papers (n=49) had a mean score of 2.00 (SD=0.82). A *t* test showed no significant difference (P=.21, t=−1.24, df=110).

The average score for perceived ability in reporting health news was 2.98 (SD=0.59) among respondents with less than 5 years' experience (n=52) and 2.76 (SD=0.55) among those with more than 5 years' experience (n=62). In this instance, a significant difference *was* found (P=.04, t=2.0, df=112): respondents with less experience reported higher perceived ability.

Average perceived ability was 2.87 (SD=0.61) among respondents at smaller newspapers (weekday circulation <25 000, n=49) and 2.87 (SD=0.54) among those at larger newspapers (n=64). No significant difference was found (P=.94, t=.06, df=111).

Fifty-five percent agreed that the news media often do not provide context for health stories. Nearly 23% disagreed. About 40% agreed that most health reporters lack adequate training to cover health; less than one third (29.2%) disagreed.

DISCUSSION

This study examined reporters' perceptions about their ability to report health news and whether health reporting experience, training, or newspaper size affects perceived ability. Results strongly suggest that health reporters are aware that they lack proficiency and want help. Only 31% and 9.7%, respectively, felt very confident in reporting health news and interpreting health statistics. In contrast, between 66% and 85% of respondents assessed 4 of 5 critical skills required for sound health reporting as sometimes difficult to nearly always difficult. Four troublesome skills for respondents are understanding key health issues, putting health news in context, producing balanced stories on deadline, and interpreting statistics.

These difficulties may stem from inadequate training. Nearly 83% of respondents re-

ported that they had no training for covering health. Over half indicated a strong interest in covering health, suggesting that many reporters want to do a good job. Reporters at smaller papers were just as likely as those at larger papers to express a strong interest, suggesting that training should include small newspapers.

Respondents gave the news media low marks for health news coverage. About half of the respondents agreed that the media often do not provide context. Respondents' views mirror science reporters' perceptions reported in a study based on a national survey.[2] This result suggests that certain criticisms of news coverage by those outside journalism are valid and that newspapers should address those criticisms.

Study limitations included a measure for perceived ability that has not been validated outside this study. Also, the use of a Midwestern sample may limit generalizability. Finally, because the study is cross-sectional, no data are available to indicate whether perceptions about training and the survey questions changed over time. This study contributes to research on training needs of health reporters by showing relatively high degrees of perceived difficulty with essential tasks. This result may mean that the quality of health reporting is low, contradicting a widespread belief held by the journalism community that reporters are trained to ask the right questions, analyze complexity, and write understandable stories.[2] This assumption does not consider that health matters often rely on statistics, science, economics, and related disciplines. Most journalism schools have not required such training. Most news executives have not perceived statistics courses as important,[20] yet without proficiency in these subjects, many health professionals and scientists believe,[2,8–13] reporters may shortchange or harm readers with poor reporting.

Thus, the need for better training of health reporters seems clear. Public health professionals could help educate reporters who contact them for information and could also place pressure on newspapers to train health reporters. Further research is needed to examine how much and what training would be most effective. ■

About the Author

Melinda Voss is with the Association of Health Care Journalists.

Requests for reprints should be sent to Room 204 Murphy Hall, School of Journalism and Mass Communication, University of Minnesota, 206 Church St SE, Minneapolis, MN 55455 (e-mail: vossx017@umn.edu).

This article was accepted June 5, 2001.

Acknowledgment

The author would like to thank her advisors, John R. Finnegan Jr, Phyllis L. Pirie, and Albert R. Tims, for their assistance in this study.

References

1. Finer D, Tomson G, Bjorkman N. Ally, advocate, analyst, agenda-setter? Positions and perceptions of Swedish medical journalists. *Patient Education and Counseling.* 1997;30:71–81.

2. Hartz J, Chappell R. *Worlds Apart—How the Distance Between Science and Journalism Threatens America's Future.* Arlington, VA: First Amendment Center. Freedom Forum 1997.

3. Roper Start Worldwide, Inc. News poll. *America's Watching.* 1995:7–18.

4. Van Trigt A, De Jong–Van Den Berg L, Haaijer-Ruskamp F, Willems J, Tromp T. Journalists and their sources of ideas and information on medicines. *Soc Sci Med.* 1994;38:637–643.

5. Lebow M, Arkin EB. Women's health and the mass media: the reporting of risk. *Women's Health Issues.* Winter 1993;3(4):181–190.

6. Davis R. Health education on the six-o'clock news: motivating television coverage of news in medicine. *JAMA.* 1988;259:1036–1038.

7. Huebner J, Fan DP, Finnegan J. Death of a thousand cuts—the impact of media coverage on public opinion about Clinton's Health Security Act. *J Health Communication.* 1997;2(4):253–270.

8. Maclean G, Wessely S. Professional and popular views of chronic fatigue syndrome. *BMJ.* 1994;308: 776–777.

9. Meacham J. Myth information (media errors in federal news). *Washington Monthly.* July–August 1993; 25(7–8):20.

10. Oxman AD, Guyatt GH, Cook DJ, Jaeschke R, Heddle N, Keller J. An index of scientific quality for health reports in the lay press. *J Clin Epidemiol.* 1993; 46:987–1001.

11. Schwartz J, Murray D. AIDS and the media. *The Public Interest.* Fall 1996;125:57(15).

12. Wilkes S. Medical scientists and health news reporting: a case of miscommunications. *Ann Intern Med.* 1997;126(12):976–982.

13. Nelkin D. An uneasy relationship: the tensions between medicine and the media. *Lancet.* 1996;347: 1600–1603.

14. Ward B. Environmental reporters: their own worst critics. *Safety & Health.* August 1993;148(2):33.

15. Lieberman T. You can't take it with you. *Columbia J Rev.* 1997;36(2)50.

16. Saari, MA, Gibson C, Osler A. Endangered species—science writers in the Canadian daily press. *Public Understanding of Science.* 1998;7:1:61–81.

17. Franklin J. How to keep training when budgets are tight. *The American Editor.* July–August 1998;763:22.

18. Freedom Forum Survey supplement. No train, no gain. *American Journalism Review. Am J Rev.* July–August 1993;15:6:F4.

19. McCleneghan S. Training and service: the killing fields of small newspapers. *Editor & Publisher.* November 8, 1997:130:45:56(2).

20. Lindenmann W, Lyon L, Nickelsberg J. Views of print and broadcast media executives toward journalism education. Ketchum Public Relations for the Virginia Commonwealth University School of Mass Communications and the Associated Press Managing Editors. 1997.

Increasing Rural–Urban Gradients in US Suicide Mortality, 1970–1997

Gopal K. Singh, PhD, MS, MSc, and Mohammad Siahpush, PhD, MS

Objectives. This study examined rural–urban gradients in US suicide mortality and the extent to which such gradients varied across time, sex, and age.

Methods. Using a 10-category rural–urban continuum measure and 1970–1997 county mortality data, we estimated rural–urban differentials in suicide mortality over time by multiple regression and Poisson regression models.

Results. Significant rural–urban gradients in age-adjusted male suicide mortality were found in each time period, indicating rising suicide rates with increasing levels of rurality. The gradient increased consistently, suggesting widening rural–urban differentials in male suicides over time. When controlled for geographic variation in divorce rate and ethnic composition, rural men, in each age cohort, had about twice the suicide rate of their most urban counterparts. Observed rural–urban differentials for women diminished over time. In 1995 to 1997, the adjusted suicide rates for young and working-age women were 85% and 22% higher, respectively, in rural than in the most urban areas.

Conclusions. The slope of the relationship between rural–urban continuum and suicide mortality varied substantially by time, sex, and age. Widening rural–urban disparities in suicide may reflect differential changes over time in key social integration indicators. (*Am J Public Health.* 2002;92:1161–1167)

Suicide is the eighth leading cause of death in the United States.[1] Although the US age-adjusted suicide mortality rate for men has changed very little over the past 3 decades, national mortality data indicate that the rate has declined fairly consistently among women, at an average rate of 2.25% per year between 1970 and 1997.[2] Social isolation (or lack of social integration) has long been recognized as one of the major determinants of suicide.[3–12] Although social isolation can be measured at the individual level by such measures as living alone and absence of social support, it can also be considered as a measure of the social environment.[3–5,7,12] Levels of rurality and of urbanization can be viewed as one such macrosocietal measure.[12,13]

A number of studies have shown that rural areas tend to have higher suicide mortality rates than urban areas.[12,14–16] Besides physical and social isolation and limited opportunities for social interaction and networks in rural areas, a number of other societal factors, such as unfavorable changes in the demographic structure and socioeconomic and industrial activity that causes social instability and disruption, have been cited as possible reasons for comparatively higher suicide rates in rural areas.[12,14–16] The extent to which suicide rates vary in response to the degree of rurality or urbanization has not been well studied. Furthermore, the degree to which rural–urban disparities in US suicide mortality among men and women have changed over time has not been examined. This report empirically examines rural–urban patterns in US suicide mortality and the extent to which rural–urban gradients in male and female suicide mortality have changed, both overall and for those aged 15 to 24, 25 to 64, and 65 years and older, during the period 1970 to 1997.

METHODS

To analyze time trends in rural–urban suicide differentials, we used a 10-category rural–urban continuum variable developed by the US Department of Agriculture.[17] This variable classifies 3103 US counties into 10 distinct groups on the basis of the counties' population size and their proximity to metropolitan areas. The 10 categories, in declining order of urbanization, are defined as follows: (1) central counties in metropolitan areas with 1 million people or more; (2) fringe counties in metropolitan areas with 1 million people or more; (3) counties in metropolitan areas with 250 000 to 1 000 000 people; (4) counties in metropolitan areas with fewer than 250 000 people; (5) urban counties with a population of 20 000 or more, adjacent to a metropolitan area; (6) urban counties with a population of 20 000 or more, not adjacent to a metropolitan area; (7) urban counties with a population of 2500 to 19 999, adjacent to a metropolitan area; (8) urban counties with a population of 2500 to 19 999, not adjacent to a metropolitan area; (9) rural counties with a population of less than 2500, adjacent to a metropolitan area; and (10) rural counties with a population of less than 2500, not adjacent to a metropolitan area. The number of counties in each of these 10 county groups was 169, 132, 323, 203, 137, 110, 616, 643, 247, and 523, respectively. The 10 county groups respectively accounted for the following percentages of the total US population in 1990: 45.5, 3.6, 22.3, 7.9, 4.0, 2.5, 6.5, 5.1, 1.0, and 1.4.

Using national mortality data files maintained by the National Center for Health Statistics, we obtained age-, sex-, and county-specific annual suicide deaths from 1970 through 1997.[1,2] Age-, sex-, and county-specific population estimates for 1970 to 1997 prepared by the US Bureau of the Census were used as denominators.[18,19] Each of the 3103 counties on the mortality data set was assigned 1 of the 10 rural–urban continuum codes. Alaska was not assigned a code and was therefore excluded from our analysis. Age-, sex-, and county-specific deaths and populations were summed within the 10 categories of the rural–urban variable and the following 6 time periods: 1970 to 1974, 1975 to 1979, 1980 to 1984, 1985 to 1989, 1990 to 1994, and 1995 to 1997. The adjustment by age of suicide mortality rates was performed by the direct method, with the age composition of the 1990 US population used as the standard. We calculated age-adjusted death rates and stan-

dard errors across each rural–urban category and time period by using 5-year age-specific death rates for the following 18 age groups: 0 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, 25 to 29, 30 to 34, 35 to 39, 40 to 44, 45 to 49, 50 to 54, 55 to 59, 60 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 years and older.[1,20]

The rural–urban differentials in suicide mortality were estimated separately for men and women. To examine the extent to which the rural–urban gradient varied across time, we modeled, using multiple linear regression, county-specific age-adjusted suicide rates (the dependent variable) as a function of rural–urban continuum (treated as a continuous independent variable) separately for each of the 6 time periods. The yearly suicide mortality trends for the most rural and most urban county groups were summarized by annual exponential rates of change.[21] To estimate relative risks of suicide for each rural–urban continuum category, we fitted Poisson regression models to the age-, sex-, and county-specific suicide death data with a log link function and the corresponding stratum-specific log population as an offset variable for the time periods 1980 to 1984, 1985 to 1989, 1990 to 1994, and 1995 to 1997.[22] The age- and sex-specific Poisson models were also fitted for 1970 to 1974 and for 1975 to 1979, but these models are not presented or discussed here for the sake of brevity. In all Poisson models, the most urban county group was selected as the reference category. All models, fitted by the SAS GENMOD procedure,[23] showed reasonable fit as determined by the likelihood ratio statistic or deviance.

We also estimated the effect of the rural–urban continuum on suicide mortality after adjusting for county-level variations in ethnic composition and for divorce rate, an important indicator of social disintegration.[3,4] High divorce rates are associated with rising suicide rates,[3,4,11,12] and Whites and American Indians have substantially higher suicide rates than other racial/ethnic groups in the United States.[24] Thus, we included county-level divorce rates (number of divorces per 1000 people aged 15 years and older), the percentage of Whites, and the percentage of American Indians in 1970, 1980, and 1990 as control variables along with rural–urban continuum (coded as a continuous variable) in the sex- and time-specific regression models of county suicide rates. Other important social integration indicators, such as unemployment rate, percentage of population living alone, percentage employed in agricultural or farming occupations, and income inequality, were considered but were excluded from the county-level regression models because of their high collinearity with divorce rate.

RESULTS

Table 1 shows age-adjusted male and female suicide rates for the 10 rural–urban county groups over each of the 6 study time periods. For men, suicide rates were generally higher in more rural areas than in urban areas, especially for the more recent time periods. For the periods 1980 to 1984, 1985 to 1989, 1990 to 1994, and 1995 to 1997, the male suicide rate for the most rural county group was respectively 21%, 26%, 37%, and 54% higher than the rate for the most urban county group. The rural–urban gradient for males, measured by the regression slope of county-level age-adjusted suicide rates on the rural–urban continuum, was positive and statistically significant for each time period except 1970 to 1974, suggesting rising suicide rates with increasing levels of rurality. Moreover, the size of the rural–urban gradient increased consistently across time, indicating widening rural–urban differentials in male suicide mortality.

For women, different rural–urban patterns in suicide mortality can be noted, especially from 1970 to 1989. For the first 4 time periods, female suicide rates were generally higher in more urban areas than in more rural areas. For the periods 1970 to 1974, 1975 to 1979, 1980 to 1984, and 1985 to 1989, the female suicide rate for the most rural county group was respectively 52%, 30%, 24%, and 16% lower than the rate for the most urban county group. The negative unadjusted rural–urban gradient also indicates lower female suicide rates at higher levels of rurality, but the declining gradient over time implies diminishing rural–urban differentials in observed female suicide mortality. In the period 1995 to 1997, no significant association between rural–urban continuum and female suicide mortality was observed.

Figure 1 shows yearly trends in age-adjusted suicide rates for the most rural and most urban county groups. Consistent with the aforementioned results in Table 1, the widening of rural–urban differentials for men and the narrowing of differentials over time for women is clearly evident. During the period 1970 to 1997, while the suicide rate for men in the most urban areas decreased at an average annual rate of 0.46% (95% confidence interval [CI]=0.30%, 0.61%), the rate for men in the rural areas grew at an average annual rate of 1.08% (95% CI=0.85%, 1.31%). The observed suicide rate for women in the most urban areas decreased at an average annual rate of 3.19% (95% CI=3.03%, 3.36%), whereas the rate for women in the rural areas declined only slightly or remained fairly stable from 1970 to 1997.

As seen in Table 1, not only did rural–urban patterns in suicide mortality vary by sex, but the crude rural–urban gradients in suicide were generally much steeper for men than for women. Men had substantially higher suicide rates than women in each county group and time period. However, sex differentials in suicide mortality appear to have increased over time in almost all county groups. For example, the suicide rate was 5 times greater in 1970 to 1974 and 6.7 times greater in 1995 to 1997 for men than for women in the most rural county group. The suicide rate was 2.3 times greater in 1970 to 1974 and 4.3 times greater in 1995 to 1997 for men than for women in the most urban county group.

Table 1 also shows the effect of rural–urban continuum on suicide mortality after adjustment for ethnic composition and divorce rate. Adjustment for ethnic composition and divorce rates widened the rural–urban differentials in suicide for men across time, with the adjusted gradient varying from 0.34 in 1970 to 1974 to 1.30 in 1995 to 1997. However, adjustment for ethnic composition and divorce rates to a large extent accounted for rural–urban differences in suicide mortality among women between 1970 and 1989 and widened the differentials for 1990 to 1994 and 1995 to 1997 such that rural

TABLE 1—Age-Adjusted Suicide Mortality Rates for US Men and Women by Rural–Urban Continuum, 1970–1997

Rural–Urban Continnum	1970–1974		1975–1979		1980–1984		1985–1989		1990–1994		1995–1997	
	Rate	SE	Rate	SE	Rate	SE	Rate	SE	Rate	SE	Rate	SE
Men												
Most urban	19.84	0.11	20.36	0.10	19.17	0.09	19.58	0.09	18.74	0.08	17.45	0.10
2nd most urban	20.28	0.41	20.21	0.37	20.08	0.35	22.00	0.34	20.66	0.31	19.85	0.38
3rd most urban	20.13	0.16	20.93	0.15	20.80	0.14	22.09	0.14	21.56	0.13	20.74	0.16
4th most urban	19.93	0.25	20.47	0.24	20.70	0.23	22.21	0.23	22.23	0.22	21.22	0.27
5th most urban	20.83	0.36	20.38	0.33	20.16	0.32	22.88	0.33	22.27	0.31	22.00	0.39
6th most urban	22.07	0.46	22.11	0.44	22.10	0.42	24.67	0.43	25.84	0.43	25.64	0.54
7th most urban	20.67	0.27	21.26	0.26	20.93	0.24	23.23	0.25	23.87	0.25	23.18	0.31
8th most urban	20.77	0.29	21.73	0.28	21.68	0.27	23.79	0.28	24.98	0.28	25.59	0.37
2nd most rural	20.06	0.65	20.63	0.63	21.49	0.61	24.35	0.64	24.44	0.63	24.40	0.80
Most rural	20.71	0.54	21.31	0.52	23.24	0.53	24.57	0.54	25.73	0.56	26.88	0.73
Rate ratio (most rural/most urban)	1.044		1.047		1.212**		1.255**		1.373**		1.540**	
Rural–urban gradient (regression slope)	0.149		0.155*		0.307**		0.550**		0.658**		1.000**	
Adjusted rural–urban gradient[a]	0.342**		0.331**		0.662**		0.856**		0.990**		1.304**	
Women												
Most urban	8.70	0.06	7.61	0.06	6.00	0.05	5.25	0.04	4.57	0.04	4.05	0.05
2nd most urban	6.27	0.22	5.90	0.19	5.01	0.16	4.46	0.14	4.21	0.13	3.88	0.16
3rd most urban	6.79	0.08	6.79	0.08	5.84	0.07	5.28	0.06	4.80	0.06	4.62	0.07
4th most urban	6.27	0.13	6.17	0.12	5.26	0.11	5.27	0.10	4.62	0.09	4.47	0.12
5th most urban	5.94	0.18	5.94	0.17	4.92	0.15	4.87	0.14	4.42	0.13	4.48	0.17
6th most urban	6.36	0.24	6.10	0.22	5.20	0.19	5.50	0.19	4.95	0.18	4.73	0.22
7th most urban	5.43	0.13	5.37	0.13	4.50	0.11	4.44	0.11	4.33	0.10	4.42	0.13
8th most urban	5.25	0.14	5.26	0.13	4.74	0.12	4.80	0.12	4.48	0.12	4.60	0.15
2nd most rural	5.36	0.34	5.33	0.31	4.53	0.28	3.94	0.25	4.06	0.25	4.29	0.33
Most rural	4.13	0.23	5.31	0.26	4.58	0.23	4.43	0.23	4.14	0.22	4.01	0.29
Rate ratio (most rural/most urban)	0.475**		0.698**		0.763**		0.844**		0.906		0.990	
Rural–urban gradient (regression slope)	–0.253**		–0.124**		–0.140**		–0.078**		–0.064**		0.002	
Adjusted rural–urban gradient[a]	–0.132**		–0.005		–0.002		0.057		0.094*		0.144**	

Note. Death rates per 100 000 population are age-adjusted by using the 1990 age composition of the US population as standard. Most urban = central metropolitan county with a population of over 1 million; 2nd most urban = fringe metropolitan county with a population of over 1 million; 3rd most urban = metropolitan county with a population of 250 000 to 1 000 000; 4th most urban = metropolitan county with a population of less than 250 000; 5th most urban = urban population of 20 000 or more, adjacent to a metropolitan area; 6th most urban = urban population of 20 000 or more, not adjacent to a metropolitan area; 7th most urban = urban population of 2500 to 19 999, adjacent to a metropolitan area; 8th most urban = urban population of 2500 to 19 999, not adjacent to a metropolitan area; 2nd most rural = rural population of less than 2500, adjacent to a metropolitan area; most rural = rural population of less than 2500, not adjacent to a metropolitan area.

[a]Adjusted rural–urban gradients are derived by regressing age-adjusted county suicide rates on rurality, ethnic composition, and divorce rate.

*P < .05; **P < .01.

areas had significantly higher suicide rates than urban areas—a pattern consistent with that for men.

Analysis of time trends in age-adjusted rates may mask important differences in trends in age-specific suicide rates. Therefore, in Table 2 we present relative risks of suicide for each county group across time, estimated separately for the young (15–24 years), those at working age (25–64 years), and the elderly (≥65 years). As seen in Table 2, the higher the degree of rurality, the higher the suicide rate among young men, especially in the 1990s. Compared with young men in the most urban areas, young men living in the most rural areas had 17% higher suicide rates in 1980 to 1984 and 60% higher suicide rates in 1995 to 1997. Rural–urban differentials in suicide mortality were somewhat less pronounced for men aged 25 to 64 years, with the relative rate ratios between the most rural and the most urban areas varying from 1.16 in 1980 to 1984 to 1.48 in 1995 to 1997. For men aged 65 years and older, rural–urban differentials in suicide were substantial in each of the time periods, although they did not appear to increase across time as sharply as those for young men.

We also estimated adjusted relative risks of male suicide for each county group by fitting age-specific Poisson models that included rural–urban continuum, ethnic composition, and divorce rate as covariates. To conserve

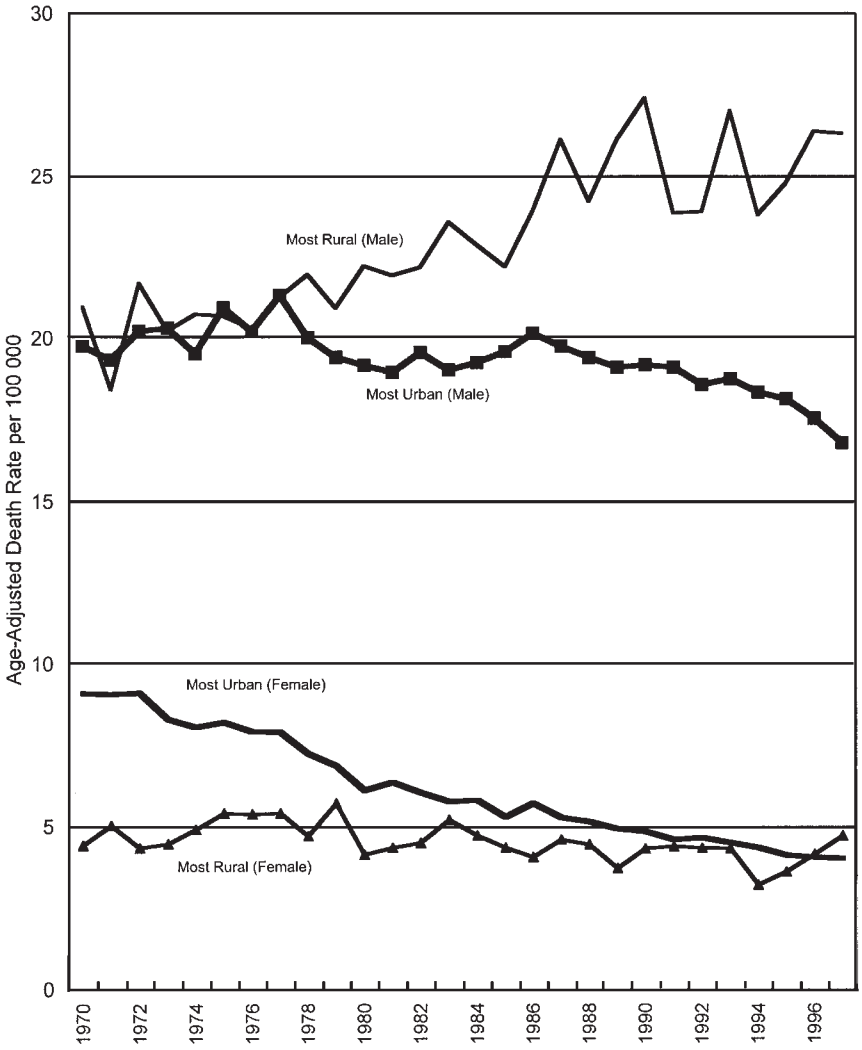


FIGURE 1—Age-adjusted US suicide mortality rates for the most urban counties (metropolitan, 1 million people or more) and the most rural counties (fewer than 2500 people): 1970 to 1997.

space, only the adjusted relative suicide risks between the most rural and most urban county groups are presented for each of the 3 age cohorts in Table 2. Controlling for ethnic composition and divorce rate substantially increased the male suicide differential between the most rural and most urban county groups for each age cohort, and the differential grew across time for young and working-age men. For the period 1995 to 1997, men in each age cohort in the most rural areas had almost twice the suicide rate of their counterparts in the most urban areas.

The data in Table 2 indicate that rural–urban patterns in female suicide mortality vary by age. While there are no consistent rural–urban patterns in female youth suicide from 1980 to 1994, young women living in the most rural areas had a 55% higher suicide rate in 1995 to 1997 than their most urban counterparts. Among the working-age women, the suicide rate tended to be 20% to 25% lower in the most rural areas during 1980 to 1984 and 1985 to 1989; however, by 1995 to 1997, the suicide rate appeared to have increased with decreasing levels of urbanization. For the elderly women, the decrease in suicide rate was associated with increasing levels of rurality in all time periods, and the magnitude of the rural–urban differentials did not appear to change appreciably across the 4 time periods. Compared with women aged 65 and older living in the most urban areas, women of the same age group living in the most rural areas had a 42% lower suicide rate in 1980 to 1984 and a 31% lower suicide rate in 1995 to 1997. Adjustment for ethnic composition and divorce rate explained most of the observed rural–urban suicide differentials among elderly women, but the suicide differential between the most rural and most urban county groups increased over time for young women aged 15 to 24. The 1995–1997 suicide rate for young women in the most rural areas was 85% higher than the rate for those in the most urban areas.

DISCUSSION

In this study, we used a rural–urban continuum measure to stratify all US counties and examined the effect of rurality on suicide mortality over time. To our knowledge, this is the first national study that has systematically examined temporal trends in the extent of rural–urban differentials in suicide mortality. The results of the study showed that the slope of the relationship between rural–urban continuum and suicide mortality varies substantially by time, sex, and age. Specifically, the study revealed an increasing rural–urban gradient in suicide mortality among men during the period 1970 to 1997. The gradients were generally steeper for men than for women, for the young and the elderly than for the working-age population, and for the more recent time periods. The relationship (positive gradient) between the degree of rurality and male suicides became even stronger when county-level variations in divorce rates and ethnic composition were controlled for. Observed female suicide rates in the most urban areas in 1970 to 1974 were substantially higher than those in the most rural areas, but by 1995 the gap had closed. In fact, when divorce rates and ethnic composition were controlled for, the pattern was reversed such that suicide rates for women, especially for young

TABLE 2—Relative Risk Estimates of Suicide Derived From Poisson Regression Models by Rural–Urban Continuum, Age, Sex, and Time Period, United States, 1980–1997

Rural–Urban Continuum	Age Group 15–24 y								Age Group 25–64 y								Age Group ≥65 y							
	1980–1984		1985–1989		1990–1994		1995–1997		1980–1984		1985–1989		1990–1994		1995–1997		1980–1984		1985–1989		1990–1994		1995–1997	
	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI
Men																								
Most urban	Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference	
2nd most urban	1.07	0.99, 1.15	1.21**	1.12, 1.30	1.17**	1.09, 1.26	1.10	1.00, 1.22	0.98	0.94, 1.03	1.05*	1.01, 1.10	1.01	0.97, 1.05	1.09**	1.04, 1.14	1.18**	1.09, 1.28	1.23**	1.14, 1.31	1.23**	1.15, 1.31	1.21**	1.11, 1.32
3rd most urban	1.00	0.97, 1.04	1.07**	1.03, 1.11	1.14**	1.10, 1.18	1.13**	1.07, 1.18	1.06**	1.04, 1.08	1.10**	1.08, 1.12	1.12**	1.10, 1.14	1.18**	1.15, 1.21	1.11**	1.07, 1.15	1.13**	1.09, 1.17	1.11**	1.08, 1.15	1.12**	1.08, 1.17
4th most urban	0.99	0.94, 1.04	1.04	0.99, 1.09	1.18**	1.12, 1.24	1.18**	1.10, 1.26	1.08**	1.05, 1.11	1.14**	1.11, 1.17	1.18**	1.15, 1.21	1.23**	1.19, 1.27	1.16**	1.10, 1.23	1.19**	1.14, 1.25	1.20**	1.14, 1.26	1.18**	1.11, 1.26
5th most urban	0.96	0.90, 1.04	1.16**	1.08, 1.24	1.17**	1.09, 1.26	1.16**	1.05, 1.27	1.04	0.99, 1.08	1.13**	1.09, 1.18	1.15**	1.11, 1.19	1.26**	1.20, 1.32	1.17**	1.09, 1.25	1.26**	1.18, 1.33	1.28**	1.21, 1.36	1.29**	1.19, 1.39
6th most urban	0.98	0.91, 1.05	1.04	0.97, 1.11	1.23**	1.15, 1.31	1.20**	1.10, 1.31	1.01	0.97, 1.05	1.03	0.99, 1.07	1.18**	1.13, 1.22	1.35**	1.29, 1.41	1.09*	1.01, 1.18	1.15**	1.07, 1.22	1.12**	1.05, 1.20	1.04	0.95, 1.14
7th most urban	1.00	0.95, 1.06	1.18**	1.11, 1.25	1.33**	1.26, 1.41	1.39**	1.29, 1.49	1.08**	1.05, 1.12	1.15**	1.11, 1.18	1.22**	1.19, 1.26	1.25**	1.20, 1.29	1.22**	1.16, 1.28	1.32**	1.26, 1.38	1.35**	1.29, 1.41	1.43**	1.35, 1.52
8th most urban	1.09**	1.02, 1.15	1.19**	1.12, 1.27	1.42**	1.34, 1.50	1.49**	1.38, 1.61	1.14**	1.10, 1.18	1.18**	1.14, 1.22	1.28**	1.24, 1.32	1.41**	1.35, 1.46	1.16**	1.10, 1.23	1.32**	1.26, 1.39	1.36**	1.30, 1.43	1.42**	1.33, 1.52
2nd most rural	1.01	0.88, 1.17	1.29**	1.13, 1.47	1.53**	1.35, 1.73	1.39**	1.17, 1.65	1.14**	1.06, 1.23	1.20**	1.11, 1.28	1.22**	1.14, 1.31	1.34**	1.24, 1.46	1.21**	1.08, 1.36	1.37**	1.24, 1.51	1.37**	1.24, 1.51	1.41**	1.23, 1.61
Most rural	1.17**	1.05, 1.30	1.51**	1.36, 1.67	1.49**	1.34, 1.66	1.60**	1.39, 1.84	1.16**	1.09, 1.23	1.18**	1.11, 1.25	1.36**	1.28, 1.44	1.48**	1.38, 1.59	1.38**	1.27, 1.50	1.32**	1.22, 1.43	1.31**	1.20, 1.42	1.53**	1.38, 1.70
Adjusted most rural[a]	1.33**	1.19, 1.49	1.63**	1.46, 1.82	1.69**	1.50, 1.89	1.76**	1.52, 2.04	1.54**	1.44, 1.64	1.48**	1.39, 1.57	1.66**	1.57, 1.76	1.77**	1.65, 1.91	1.85**	1.69, 2.03	1.79**	1.65, 1.96	1.72**	1.57, 1.88	2.06**	1.84, 2.30
Women																								
Most urban	Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference		Reference	
2nd most urban	1.00	0.85, 1.18	1.01	0.85, 1.20	1.21*	1.02, 1.44	1.05	0.82, 1.34	0.82**	0.76, 0.89	0.85	0.79, 0.92	0.89**	0.82, 0.96	0.95	0.87, 1.05	0.75**	0.64, 0.90	0.73**	0.62, 0.86	0.83*	0.71, 0.97	0.91	0.75, 1.12
3rd most urban	0.95	0.88, 1.03	0.89**	0.82, 0.96	1.07	0.98, 1.16	1.02	0.91, 1.15	0.98	0.95, 1.02	1.03	0.99, 1.06	1.06**	1.03, 1.10	1.18**	1.13, 1.23	0.80**	0.74, 0.86	0.89**	0.83, 0.95	0.88**	0.82, 0.94	0.93	0.85, 1.01
4th most urban	0.85**	0.76, 0.96	0.97	0.86, 1.08	1.00	0.88, 1.14	1.05	0.89, 1.23	0.91**	0.87, 0.96	1.06*	1.00, 1.11	1.07**	1.02, 1.13	1.16**	1.09, 1.24	0.70**	0.62, 0.78	0.84**	0.76, 0.93	0.77**	0.70, 0.86	0.91	0.79, 1.04
5th most urban	0.76**	0.64, 0.90	0.88	0.74, 1.04	0.91	0.76, 1.10	0.91	0.70, 1.17	0.84**	0.78, 0.91	0.95	0.88, 1.02	1.02	0.94, 1.09	1.19**	1.09, 1.30	0.75**	0.64, 0.87	0.87*	0.76, 0.99	0.78**	0.68, 0.90	0.86	0.71, 1.04
6th most urban	0.86	0.74, 1.01	0.80*	0.67, 0.95	1.06	0.89, 1.25	1.24*	1.00, 1.53	0.83**	0.77, 0.89	1.01	0.94, 1.08	0.99	0.92, 1.06	1.21**	1.11, 1.33	0.74**	0.63, 0.88	0.93	0.81, 1.07	0.97	0.85, 1.11	0.75**	0.61, 0.93
7th most urban	0.76**	0.67, 0.88	0.96	0.84, 1.10	1.10	0.95, 1.27	1.18	0.98, 1.42	0.78**	0.74, 0.83	0.87**	0.82, 0.92	0.97	0.91, 1.03	1.13**	1.05, 1.22	0.61**	0.54, 0.69	0.66**	0.59, 0.74	0.74**	0.66, 0.83	0.81**	0.70, 0.94
8th most urban	0.94	0.82, 1.08	1.06	0.92, 1.22	1.14	0.98, 1.33	1.34**	1.11, 1.62	0.81**	0.76, 0.87	0.93*	0.87, 0.99	0.99	0.92, 1.05	1.14**	1.05, 1.24	0.55**	0.48, 0.63	0.71**	0.63, 0.80	0.70**	0.61, 0.79	0.70**	0.58, 0.83
2nd most rural	0.87	0.63, 1.20	0.86	0.61, 1.22	0.91	0.62, 1.34	0.97	0.59, 1.58	0.77**	0.66, 0.89	0.75**	0.64, 0.88	0.94	0.81, 1.09	1.17	0.99, 1.39	0.59**	0.44, 0.79	0.76*	0.59, 0.98	0.69**	0.52, 0.91	0.55**	0.36, 0.85
Most rural	0.84	0.64, 1.10	1.01	0.77, 1.33	1.15	0.85, 1.56	1.55*	1.10, 2.19	0.80**	0.71, 0.90	0.86*	0.76, 0.97	0.89	0.79, 1.02	0.93	0.79, 1.10	0.58**	0.46, 0.74	0.69**	0.56, 0.86	0.65**	0.52, 0.82	0.69**	0.50, 0.94
Adjusted most rural[a]	1.31	0.99, 1.73	1.29	0.97, 1.71	1.43*	1.04, 1.94	1.85**	1.29, 2.64	1.16*	1.03, 1.31	1.17*	1.03, 1.33	1.19**	1.04, 1.36	1.22*	1.03, 1.45	0.96	0.75, 1.22	1.13	0.90, 1.41	1.01	0.80, 1.28	1.04	0.75, 1.44

Note. RR = relative risk; CI = confidence interval. For explanation of rural–urban continuum, see note to Table 1.

[a]The adjusted relative suicide risk for the most rural vs the most urban county group was estimated by a Poisson regression model that included rural–urban continuum, ethnic composition, and divorce rate as covariates.

*P < .05; **P < .01.

women, appeared to be significantly higher in the most rural than the most urban areas in 1995 to 1997.

The results of the study are highly consistent with the Durkheimian theory of social integration and suicide.[3–5,8,9,11,12] According to Durkheim, low levels of social integration are associated with high suicide rates. Durkheim defined social integration as being attached to social groups, maintaining interpersonal ties, and feeling allegiance to social groups. Low levels or an absence of social integration, as measured by living alone or experiencing marital disruption, represent social isolation and the atomization of individuals in a community.[3] The increasing rural–urban gradients in suicide observed in this study might to some degree reflect differential changes over time in the social integration indicators for rural and urban areas. Changes in the socioeconomic and demographic structure of a community can have both short- and long-term implications for the health of that population, including suicide rates. Declines in traditional farm activity, the change from an agriculture-based economy to a more service- and manufacturing-oriented economy, and the substantial population loss due to birth deficits and out-migration in most rural communities can lead to a deemphasizing of traditional institutions such as the family and religion, a progressive weakening of social and community ties, and a loss of people's sense of community.

Both rural and urban areas have experienced profound social and demographic changes during the past 3 decades. However, these changes have affected rural areas much more adversely than urban areas. Our own analysis of the 1970 to 1990 census and vital statistics data (not presented here for the sake of brevity) indicates that a variety of social integration and demographic factors are substantially correlated with county-level suicide rates in each of the 6 time periods.[25] For example, between 1970 and 1990 the population increase was almost 5 times greater in the most urban areas than in the most rural areas. While the most rural areas of the country as a whole experienced a substantial drop (20%) in the crude birth rate, the most urban areas saw their birth rate increase by 4%. The birth deficit (excess of deaths over births) was particularly dramatic in the rural areas, which experienced a 61% increase in deaths over births. In the most urban areas, there was a 25% increase in births over deaths. Although the rate of agriculture-sector employment was considerably greater in the rural than the most urban areas, the rate declined by 32% for the rural areas and increased by 95% for the most urban areas. Divorce rates increased twice as rapidly, and the percentage of those living alone rose 3 times more rapidly in the most rural areas than in the most urban areas. Household crowding, as indexed by the percentage of housing units with more than 1 person per room, declined by 67% in the most rural areas and by 12% in the most urban areas. The rate of female participation in the labor force and income inequality grew at a much higher rate in the most rural areas than in the most urban areas.

Changing societal attitudes toward suicide in terms of its recognition and labeling may also have contributed to rural–urban disparities over time.[26] Furthermore, to the extent that those migrating to rural areas or those returning from urban areas had higher suicide rates than native rural residents, they might have had an important effect on trends.[12,15] Despite the substantial birth deficit, the rural population during the period 1970 to 1990 grew owing to net migration. Those migrating to rural and nonmetropolitan areas during the 1970s and 1990s tended to be elderly retirees, blue-collar workers, and disenchanted city residents.[27] The higher suicide rates observed among rural migrants have been attributed to social disruption, lack of social support and networks, change in lifestyle, and the increased alienation that accompanies migration.[12,15]

Geographic variations in access to firearms may also have contributed to the observed rural–urban disparities in suicide. Access to firearms is more common in rural than in urban areas.[15,16,28] Studies have shown a strong association between increased firearm availability and a steadily upward trend in firearm suicide rates from 1946 to 1982.[29–31] Our own analysis of the mortality data (not shown) indicates consistently growing rural–urban disparities in firearm suicide rates from 1979 to 1997, with rural areas having rates more than twice those of the most urban areas in 1997. Firearm suicides accounted for over 75% of all rural suicides during the period 1979 to 1997, whereas firearm suicides represented half of all suicides in the most urban areas.[2]

To help reduce rural–urban disparities in suicide, public health researchers and policymakers need to closely monitor geographic and temporal trends in social integration measures. Social and public health policies that emphasize investment in social integration or social capital through job creation, provision of gainful employment and social services, and improved social support and networks through community organization and involvement, especially for the rural young and elderly, may lower suicide rates.[3] Restricting access to firearms may also reduce suicide rates, particularly in rural areas.

Although suicide is one of the most individualistic acts, we examined variations in county-level suicide rates rather than individual suicide risks. Our study design was therefore ecological, and the present analysis is not likely to be characterized by ecological fallacy or bias. We modeled ecological variations in suicide rates primarily as a function of an ecological variable, degree of rurality or urbanization, which is not quite reducible to the level of the individual. This study has certain limitations, however. The rural–urban continuum variable we used to stratify all of the nation's counties for the period 1970 to 1997 was developed in 1993 by using population and commuting criteria from the 1990 census.[17,25] It is possible that these criteria may not apply as well to the earlier time periods, especially the 1970s. In other words, the classification of counties into specific rural–urban county groups, particularly into the small and medium-sized urban county groups, may have been somewhat different during the 1970s and the early 1980s. This may have introduced some inconsistency in the estimation of rural–urban gradients in suicide across time. Second, the accuracy with which suicide is recorded as a cause of death on the death certificate may differ between rural and urban areas. Because of social stigma, suicide deaths may be more likely to be underreported in rural areas than in urban areas. As a result, the observed rural–urban sui-

cide differentials reported here may be underestimated for men and overestimated for women. Trends in suicide rates could also be affected if the registration of suicide and errors in classifying suicide deaths varied systematically between rural and urban areas and over time. ■

About the Authors

Gopal K. Singh is with the Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Md. Mohammad Siahpush is with the VicHealth Centre for Tobacco Control, Cancer Control Research Institute, Anti-Cancer Council of Victoria, Carlton, Australia.

Requests for reprints should be sent to Gopal K. Singh, PhD, MS, MSc, National Cancer Institute, Division of Cancer Control and Population Sciences, 6116 Executive Blvd, Suite 504, MSC 8316, Bethesda, MD 20892-8316 (e-mail: gopal_singh@nih.gov).

This article was accepted March 23, 2001.

Note. *The views expressed in this report are the authors' and not necessarily those of their institutions.*

Contributors

Both authors planned and designed the study, analyzed the data, and wrote the paper.

References

1. Hoyert DL, Kochanek KD, Murphy SL. Deaths: final data for 1997. *Natl Vital Stat Rep.* 1999;47(19): 1–108.

2. National Center for Health Statistics. *Vital Statistics of the United States, 1970–1997, Volume II: Mortality, Parts A and B.* Washington, DC: Public Health Service; 1999.

3. Siahpush S, Singh GK. Social integration and mortality in Australia. *Aust N Z J Public Health.* 1999;23: 571–577.

4. Singh GK, Wilkinson AV, Song FF, et al. *Health and Social Factors in Kansas: A Data and Chartbook, 1997–98.* Lawrence, Kan: Allen Press; 1998.

5. Kposowa AJ, Breault KD, Singh GK. White male suicide in the United States: a multivariate individual-level analysis. *Soc Forces.* 1995;74:315–323.

6. Trout DL. The role of social isolation in suicide. *Suicide Life Threat Behav.* 1980;10:10–23.

7. Seeman TE. Social ties and health: the benefits of social integration. *Ann Epidemiol.* 1996;6:442–451.

8. Trovato F, Jarvis GK. Immigrant suicide in Canada: 1971 and 1981. *Soc Forces.* 1986;65:433–457.

9. Travis R. Halbwachs and Durkheim: a test of two theories of suicide. *Br J Sociol.* 1990;41:225–243.

10. Fernquist RM, Cutright P. Societal integration and age-standardized suicide rates in 21 developed countries, 1955–1989. *Soc Sci Res.* 1998;27:109–127.

11. Leenaars AA, Lester D. Domestic integration and suicide in the provinces of Canada. *Crisis.* 1999;20: 59–63.

12. Wilkinson KP, Israel GD. Suicide and rurality in urban society. *Suicide Life Threat Behav.* 1984;14: 187–200.

13. Dewey R. The rural–urban continuum: real but relatively unimportant. *Am J Sociol.* 1963;66:60–66.

14. Saunderson T, Haynes R, Langford IH. Urban–rural variations in suicides and undetermined deaths in England and Wales. *J Public Health Med.* 1998;20: 261–267.

15. Morrell S, Taylor R, Slaytor E, Ford P. Urban and rural suicide differentials in migrants and the Australian-born, New South Wales, Australia 1985–1994. *Soc Sci Med.* 1999;49:81–91.

16. Dudley MJ, Kelk NJ, Florio TM, Howard JP, Waters BG. Suicide among young Australians: an interstate comparison of metropolitan and rural trends. *Med J Aust.* 1998;169:77–80.

17. Butler MA, Beale CL. *Rural–Urban Continuum Codes for Metro and Nonmetro Counties, 1993.* Washington, DC: Economic Research Service, US Dept of Agriculture; 1994. Staff report 9425.

18. Sink L. *Estimates of the Population of Counties by Age, Sex, Race and Hispanic Origin: 1990 to 1998.* Washington, DC: US Bureau of the Census; 1999.

19. Hollmann FW. *United States Population Estimates, by Age, Sex, Race and Hispanic Origin: 1980 to 1988.* Washington, DC: US Bureau of the Census; 1990. Current Population Reports, Series P-25, No. 1045.

20. Singh GK, Kochanek KD, MacDorman MF. Advance report of final mortality statistics, 1994. *Month Vital Stat Rep.* 1996;45(3)(suppl):1–80.

21. Singh GK, Yu SM. Infant mortality in the United States: trends, differentials, and projections, 1950 through 2010. *Am J Public Health.* 1995;85:957–964.

22. Agresti A. *An Introduction to Categorical Data Analysis.* New York, NY: John Wiley & Sons Inc; 1996.

23. *SAS/STAT Software: Changes and Enhancements Through Release 6.12: The GENMOD Procedure.* Cary, NC: SAS Institute Inc; 1997.

24. *Health, United States, 1999.* Hyattsville, Md: National Center for Health Statistics; 1999.

25. *The Area Resource File (ARF): Public Use File Technical Documentation.* Rockville, Md: Health Resources and Services Administration; 1996.

26. Morrell S, Taylor R, Quine S, Kerr C. Suicide and unemployment in Australia 1907–1990. *Soc Sci Med.* 1993;36:749–756.

27. Johnson KM. The rural rebound. *Reports on America.* 1999;1(3):1–19.

28. Wiktor SZ, Gallaher MM, Baron RC, Watson ME, Sewell CM. Firearms in New Mexico. *West J Med.* 1994;161:137–139.

29. Zwerling C, McMillan D, Cook PJ, et al. Firearm injuries: public health recommendations. *Am J Prev Med.* 1993;9(suppl 1):52–55.

30. Baker SP, O'Neill B, Ginsburg MJ, Li G. *The Injury Fact Book.* New York, NY: Oxford University Press; 1987.

31. Lester D, Clarke RV. Note on "suicide and increased availability of handguns in the United States": the influence of firearm ownership on accidental deaths. *Soc Sci Med.* 1991;32:1311–1313.

Assessment of a New Approach to Family Planning Services in Rural Pakistan

Mehboob Sultan, MSc, John G. Cleland, MA, and Mohamed M. Ali, PhD

Objectives. In 1993, the government of Pakistan started a new approach to the delivery of contraceptive services by training literate married women to provide doorstep advice and supplies in their own and neighboring communities. This report assesses whether this community-based approach is starting to have an impact on contraceptive use in rural areas.

Methods. A clustered nationally representative survey was used to collect data on contraceptive use and access to services in each cluster. Two-level logistic regression was applied to assess the effects of service access, after potential confounders were taken into account.

Results. Married women living within 5 km of 2 community-based workers were significantly more likely to be using a modern, reversible method of contraception than those with no access (odds ratio=1.74; 95% confidence interval=1.11, 2.71).

Conclusions. After decades of failure, the managers of the family planning program have designed a way of presenting modern contraceptives that is appropriate to the conditions of rural Pakistan. The new community-based approach should be steadily expanded. (*Am J Public Health.* 2002;92:1168–1172)

With 137 million inhabitants, Pakistan is the world's seventh most populous nation. According to United Nations projections, the population will grow to 285 million by 2050, at which time Pakistan will rank as the world's fourth largest country.[1] The main reason for this huge projected increase and the rise in relative ranking is the slow pace of fertility decline. The level of childbearing started to fall in the 1980s, from about 7 births per woman; for the period 1995 to 2000 it was estimated to be 5.0 births per woman, a value considerably higher than for other countries in the region (3.1 for India, 4.4 for Nepal, 3.1 for Bangladesh).[1,2]

Socioeconomic factors offer no obvious explanation for this divergence between Pakistan and its neighbors. For instance, the country is considerably richer and more urbanized than Nepal and Bangladesh, and levels of literacy and life expectancy are similar.[3] Furthermore, there is no evidence that Pakistani couples want particularly large families. The 1975 Pakistan Fertility Survey showed that, on average, married women wanted 4 children, a number typically found in Asian surveys at that time.[4] The most recent national survey (1996–1997) found that the average desired family size was 3.5 children.[5] It also confirmed the existence of a large unmet need for contraception: 38% of married women wanted no more children but were practicing no method of birth control. What distinguishes Pakistani couples most clearly from their neighbors is a reluctance or inability to translate reproductive preferences into appropriate behavior. The wide gap between preferences and practice stems partly from a prevailing (but erroneous) belief that Islam is opposed to contraception, and concerns about side effects and health hazards of modern methods.[6,7] It also reflects the low political priority given to family planning for much of the last 30 years.

In the 1960s, President Ayub Khan initiated a vigorous family planning program that was widely applauded as a model for other Islamic countries. However, it had serious design defects. It relied heavily on one method, the intrauterine device, and on financial incentives to clients and providers. It achieved little and was discredited when Ayub Khan fell from power.[8] For the next 20 years, first under the regime of Prime Minister Zulfiqar Ali Bhutto (1971–1977) and then under that of President Zia-ul-Haq (1977–1988), family planning was a low priority. During this period, several different approaches were tried, but implementation was poor and a sense of urgency was lacking. The reasons stemmed partly from domestic political considerations. Family planning was a low priority for Bhutto probably because it was so closely identified with his bitter political rival, Ayub Khan. Zia was reluctant to promote family planning vigorously because he drew much of his political support from conservative religious elements.[9] During his tenure, for instance, expenditure on family planning was sharply reduced and television advertising of family planning messages was banned.

The political climate improved in the 1990s. Family planning has received steady support from successive regimes, and there have been several encouraging developments. Social marketing of contraceptives has expanded, and efforts are being made to involve private medical practitioners more closely in service provision. Perhaps the most promising initiative is the deployment of specially trained literate women to provide contraceptive information and basic services in their own and surrounding villages. A large body of international evidence shows that community-based initiatives are often effective at raising contraceptive use,[10,11] and the results of small-scale projects in Pakistan have been positive.[12] This approach has worked particularly well in Bangladesh and is thought to be largely responsible for the unexpectedly large decline in fertility there.[13] In April 1992, a team of senior officials from the Ministry of Population Welfare, Pakistan's lead agency for family planning, visited Bangladesh. Later that year, a plan was announced to recruit and train 12 000 female village-based family planning workers by 1998.

Recruitment criteria were specified. Family planning workers are married women aged

18 to 50, with at least 10 years of schooling, who currently reside in a rural area. Recruits are trained for 7 months; the training comprises 4 months of classroom instruction interspersed with field practice. Duties include the compilation of a register of local married women of reproductive age who should be visited at home at regular intervals. Workers are supplied with oral and injectable contraceptives and condoms for distribution, together with a range of medications for the treatment of sick children. For this work, staff receive a monthly salary of 1500 rupees, equivalent to about US $25. Supervision takes the form of monthly visits from female managers, who are provided with transport. Workers also visit district centers to collect supplies and salary.

In 1994, the Ministry of Health launched a very similar but larger program involving female community-based health workers, called "lady health workers." Recruitment criteria, training, remuneration, supervision, and method of service delivery are almost identical to those adopted by the Ministry of Population Welfare. While the main emphasis of the Ministry of Health program is on maternal and child health, the provision of family planning services is part of the package and lady health workers are supplied with oral contraceptives and condoms. The 2 ministries collaborated to minimize overlap in the placement of workers. By the end of 1996, about 5500 village-based family planning workers and 30 000 lady health workers had been trained and were operational. While the Ministry of Population Welfare program has been carefully monitored by several small-scale studies, which have documented practical problems of implementation[14,15] but also encouraging signs of success in 4 Punjabi communities,[16] the Ministry of Health program remains unevaluated.

This report assesses the impact of this new approach to the provision of family planning services in Pakistan. Specifically, we sought to determine whether or not use of modern reversible methods of contraception is higher in rural localities served by these community-based workers than in other localities. Attention was restricted to rural areas because the family planning workers, unlike the lady health workers, do not operate in towns and cities. The focus on reversible methods (thereby excluding sterilization) stemmed from the fact that many sterilizations preceded the start of the outreach programs. Moreover, the primary emphasis of the programs is on reversible rather than permanent methods.

METHODS

The data for this assessment come from the rural portion of the Pakistan Fertility and Family Planning Survey, a nationally representative survey conducted in late 1996 and early 1997 by the National Institute of Population Studies, Islamabad, in collaboration with the London School of Hygiene and Tropical Medicine. The detailed methodology of the survey has been previously published.[5] In brief, the sample frame was prepared by the Federal Bureau of Statistics. In the rural domain, 175 geographical clusters were selected and a household census was carried out in each. From the household lists, 31 households were selected randomly in each cluster. Because of its small population, Baluchistan province was oversampled, and all results reported below have been weighted to adjust for this overrepresentation. Selected households were visited by specially trained female staff, who first identified and then interviewed ever-married women aged 15 to 49. The contact–response rate was 89%. Interviews with women covered a wide range of fertility, family planning, and health topics, including current contraceptive status and the number of household visits, if any, received from health or family planning workers in the previous 12 months.

Concurrently with the survey of women, an inventory was made of all health and family planning facilities and staff within a radius of 5 km of the center of each rural cluster. All facilities and medical practitioners were visited to ascertain the precise nature of family planning services offered. All data on service access, including access to community-based workers, are derived from this facility survey. Because Pakistani villages are typically compact, a single cluster-level measure of access is a good approximation of individual women's access. In order to obtain a complete profile of each selected cluster, the presence of other types of modern facilities or services (e.g., post office, bank, schools, and electricity) was also noted.

For the analysis, the 2 data files were linked and 2-level statistical modeling was performed with the software package Stata[17] to estimate the effect of access to services, including those from community-based workers, on the use of reversible methods of contraception. As the outcome is binary (i.e., use or nonuse of a modern reversible method), logistic regression with a random effect term for the intercept is applied.[18] The purpose of the random effect term is to take account of unobserved heterogeneity at the cluster level. Results are shown as unadjusted and adjusted odds ratios, with 95% confidence intervals. Widowed, divorced, and separated women (n=171), along with sterilized couples (n=224), were excluded. In addition, 372 women living in 12 clusters where the facilities were either closed or there was no staff to be interviewed were excluded from the analysis because key access variables could not be measured. These exclusions reduced the sample size to 4676 women residing in 163 rural clusters.

RESULTS

The unweighted results of the 1996–1997 survey show that 7% of married women in rural Pakistan were using a reversible modern method. After data are weighted to adjust for the overrepresentation of Baluchistan (a low-use province), this estimate rises to 9%, with a fairly even spread between intrauterine devices, condoms, oral pills, and injectables. In addition, 4% of couples had been sterilized and 6% were using coitus interruptus or periodic abstinence. Thus, the overall prevalence of contraceptive use in rural areas was 19%, compared with 11% recorded by the previous 1994–1995 national survey.[19] Nearly half (43%) of women lived within 5 km of a static facility with staff trained in family planning, and 24% had similar access to a private practitioner who provided contraceptive services. In terms of access to the new outreach programs, 37% of rural women lived within 5 km of a lady health worker or a village-based family planning worker and an additional 10% lived within 5 km of both types of

TABLE 1—Unadjusted and Adjusted Effect of Service Access and Other Factors on Use of Reversible Modern Methods of Contraception in Rural Pakistan: Odds Ratios (ORs) and 95% Confidence Intervals (CIs) From Logistic Regression

	n[a]	Use, %	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Access				
Community-based worker, n				
0	1820	5.3	1.00	1.00
1	2128	9.4	1.83 (1.30, 2.56)	1.14 (0.80, 1.61)
2 or more	728	16.8	4.05 (2.71, 6.07)	1.74 (1.11, 2.71)
Any static facility				
No	2644	8.3	1.00	
Yes	2033	9.8	1.12 (0.81, 1.56)	0.84 (0.62, 1.14)
Private practitioner				
No	3561	7.3	1.00	
Yes	1116	14.4	2.41 (1.74, 3.34)	1.47 (1.00, 2.15)
Community Characteristics				
Modernization				
Low	954	4.7	1.00	
Medium	1793	6.7	1.23 (0.78, 1.94)	0.75 (0.44, 1.26)
High	1929	12.7	2.63 (1.71, 4.07)	0.72 (0.39, 1.31)
School availability				
Low	826	3.9	1.00	
Medium	1129	5.6	1.45 (0.84, 2.52)	1.42 (0.80, 2.54)
High	1113	10.5	3.07 (1.82, 5.19)	2.57 (1.36, 4.88)
Very high	1609	12.8	3.90 (2.39, 6.36)	2.67 (1.35, 5.27)
Distance to town, km				
<10	1958	11.2	1.00	
10–19	1205	6.9	0.55 (0.37, 0.82)	0.68 (0.46, 1.02)
≥20	1513	7.8	0.61 (0.42, 0.88)	0.87 (0.60, 1.26)
Individual and Household Factors				
Household wealth score				
Poor	1614	5.2	1.00	
Moderate	1890	8.7	1.65 (1.23, 2.19)	1.30 (0.95, 1.76)
High	1173	9.0	2.95 (2.19, 3.98)	1.69 (1.17, 2.45)
Education level				
No education	3979	7.6	1.00	
Up to primary	470	14.6	1.93 (1.43, 2.61)	2.08 (1.49, 2.90)
Above primary	229	21.1	2.79 (1.94, 4.02)	2.62 (1.70, 4.04)
Watch TV				
Never	2388	6.0	1.00	
Not regularly	1141	9.1	1.42 (1.06, 1.91)	1.23 (0.91, 1.68)
Daily	1148	15.0	2.80 (2.15, 3.64)	1.75 (1.26, 2.42)
Age group, y				
<25	1296	3.8	1.00	
25–34	1955	11.3	3.13 (2.26, 4.33)	1.13 (0.77, 1.66)
≥35	1425	10.5	2.90 (2.06, 4.07)	0.79 (0.51, 1.22)
Living children				
0–1	1250	1.4	1.00	
2–3	1336	9.3	7.21 (4.30, 12.08)	6.91 (3.97, 12.05)
4–5	1033	12.6	10.83 (6.46, 18.17)	13.66 (7.59, 24.59)
≥6	1058	13.8	12.01 (7.18, 20.09)	19.20 (10.45, 35.27)
Boy–girl difference				
1 or more girls	1606	7.8	1.00	
Same number	1395	6.3	0.84 (0.63, 1.13)	1.43 (1.04, 1.95)
1 or more boys	1676	12.3	1.86 (1.45, 2.37)	1.88 (1.45, 2.43)

[a]Weighted frequency.

worker. In rural Pakistan as a whole, 16% of women said that they had been visited at home by a health or family planning worker within the previous 12 months. This proportion rises to 20% in clusters with access to one type of worker and to 32% in clusters with access to both types of worker ($P<.001$). Partly because workers display boards outside their dwelling that indicate their institutional affiliation, most women were able to distinguish between Ministry of Health workers and those from the Ministry of Population Welfare, despite the considerable overlap in the nature of their services. When the most recent visit had been from a family planning worker, 92% of women recalled that contraception had been discussed and 73% that health matters had been raised. The corresponding figures for a health worker's visit were 79% and 86%.

The middle column of Table 1 shows the crude or unadjusted effects of access to services, other cluster characteristics, and household and individual characteristics on the use of modern reversible methods. Three types of access are examined: the availability of community-based workers, the presence of a static health or family planning center having at least 1 staff member with special training in contraception, and the presence of at least 1 private practitioner who offers contraceptive services. There is no relationship between access to a static center and contraceptive use, but the other 2 types of service access are significantly related.

Other features of clusters are represented by 3 variables. The indicator of modernization represents the number of modern institutions, such as banks and post offices, in the locality. Similarly, the indicator of school availability represents the number and types of schools within 5 km of the cluster. Along with the third variable, distance from the nearest town, these variables are included in the analysis because of theoretical or commonsense expectations that they may influence the uptake of contraception. Indeed, Table 1 shows that all have statistically significant associations.

The final block of factors to be considered is derived from the survey of individuals; it includes characteristics that are known from previous studies to act as powerful determinants of contraceptive uptake. As expected,

contraceptive use rises in line with household wealth, extent of schooling of the wife, and exposure to television. It is also higher among older couples, those with larger families, and those who have more sons than daughters than among other types of couples.

These effects on contraceptive use were reassessed with a multivariate model; the results in terms of adjusted odds ratios are shown in the right column of Table 1. The effect of access to community-based workers is severely attenuated, suggesting that conditions for contraceptive uptake are favorable in localities served by such workers. Nevertheless, the net effect remains significant. In localities where both types of worker are present, the probability of using a modern reversible method is increased by 74%. The difference in use between localities served by one worker and those with no worker access is in the expected direction but is not significant. The effect of access to private practitioners offering contraceptive services is also significant. Among the other cluster-level factors, only the presence of schools retains a net effect on contraceptive use. All the household and individual factors (except women's age) remain significantly associated with use.

DISCUSSION

In a population as large and rapidly growing as Pakistan's, the design of effective family planning services is an issue of both national and international importance. The key to success in many countries has been to expand information and services beyond the restrictions of a clinic-based approach. In this regard, Pakistan appears to be no exception. For many years, family planning services have been available at over 1200 family welfare centers run by the Ministry of Population Welfare. Contraceptive services are also provided at many of the Ministry of Health's rural health centers and basic health units. These facilities are severely underutilized. For instance, one evaluation showed that, on average, a family welfare center received only 2 clients per day.[20] One reason for this underutilization is the limited mobility of Pakistani women, which stems from deep-rooted traditions of female seclusion, or *purdah*. Specific questions were asked in the 1996–1997 survey about mobility. Only 15% of rural women had been outside their village for any purpose in the previous month without being accompanied by another adult, and only 20% said that they would be able to visit a hospital by themselves. Thus, accessing a health or family planning center in a nearby village or town is logistically complex because, typically, the husband, mother-in-law, or another adult family member must be persuaded to act as an escort, thereby preserving the family's *izzat*, or honor.

There are many ways to expand access beyond static clinics and, over the past 30 years, many of them have been tried in Pakistan. Mobile units have proved to be expensive and ineffective, partly because of the practical difficulties of vehicle maintenance.[21] The training of traditional birth attendants, or *dais*, has been tried on many occasions in South Asia, but their low social status prevents them from being plausible agents of social change, and their impact on family planning has been negligible.[22–24] The most interesting of Pakistan's attempts to improve access occurred in the early 1970s with the deployment of 2-person teams (1 man and 1 woman) to provide contraceptive advice and supplies at the doorstep.[22,25] Unfortunately, recruitment of workers became excessively politicized and the program was dismantled. It remains uncertain whether proper and sustained implementation would have had an impact. Unlike the current community-based programs, there was no insistence that staff reside in their work areas. Moreover, the deployment of mixed-sex teams in conservative rural Pakistan is problematic.

The current approach may be a sociologically more appropriate way of presenting contraception in rural Pakistan than that begun in the early 1970s. The workers are members of the communities they serve. Being literate in a society where nearly 90% of married rural women cannot read or write means that workers typically belong to respected and influential families and can thus act as bridges between village life and the outside world, which may be regarded as alien and threatening. By allowing wives or daughters-in-law to work for either ministry, male members of workers' families implicitly endorse family planning. Indeed, it is not an uncommon sight to see husbands actively helping their wives by transporting them to nearby villages on their motorbikes.

The results reported here suggest that a steady expansion of current community-based programs may be the most effective way in the short term of meeting the huge potential demand for contraceptive services in rural Pakistan and in facilitating fertility decline. The deployment of thousands of salaried workers is not a cheap option, nor should the difficulties of maintaining adequate logistical and supervisory systems be understated. The Ministry of Population Welfare currently spends about US $37 million per year on family planning; much of this expenditure is on types of service that clearly do not meet the needs of women. The village-based family planning program accounts for only 18% of expenditure. A strong case exists for increasing this fraction. In the long term, of course, integration of the health and family planning outreach programs is both desirable and inevitable. Since it lost direct responsibility for family planning in the mid-1960s, the Ministry of Health has made rather little effort to provide contraceptive advice and supplies as part of its overall health service. One of the most encouraging results to emerge from the 1996–1997 survey is that the Ministry of Health workers are actively promoting contraception. It is to be hoped that an era of closer collaboration between the 2 ministries has begun.

The results also suggest that expansion of schools in rural Pakistan is likely to have a profound effect on reproductive behavior. While the effects of parental education on fertility have been intensively studied, the influence of children's schooling opportunities has attracted comparatively little empirical attention. The positive link between access to schools and contraceptive use found in this analysis is consistent with theoretical expectations that new opportunities to invest in children will encourage parents to choose to have fewer children.[26,27]

CONCLUSION

This analysis shows that use of reversible modern methods of contraception is significantly higher in localities having good access

to literate, female community-based workers than in localities with little or no access. Because this result is not based on an experimental study design, causal attribution has to be cautious. Nevertheless, the link between access and contraceptive use persisted despite the introduction of a wide range of controls into the analysis. It is thus highly probable that the new community-based programs of the ministries of health and population welfare are starting to have an impact. Access to private practitioners who offer contraceptive services also had a significant effect on use. In contrast, this analysis detected no impact from access to static family planning services.

The availability of schools also exerts a powerful influence on contraceptive uptake, but the presence of other modern institutions, or proximity to a town, had no effect. ■

About the Authors

Mehboob Sultan is with the National Institute of Population Studies, Islamabad, Pakistan. John G. Cleland and Mohamed M. Ali are with the Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, England.

Requests for reprints should be sent to John G. Cleland, MA, Centre for Population Studies, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, 49-51 Bedford Square, London WC1B 3DP, England (e-mail: john.cleland@lshtm.ac.uk).

This article was accepted July 24, 2001.

Contributors

M. Sultan designed the study, supervised data collection, and reviewed all drafts. J.G. Cleland was responsible for drafting and interpretation. M.M. Ali analyzed the data and reviewed all drafts. All authors designed the analysis.

Acknowledgments

Funding for the national survey on which this analysis is based was provided by the United Kingdom Department for International Development.

We are grateful for the support of Dr Abdul Hakim, the director of the survey and technical director of the National Institute of Population Studies.

References

1. *World Population Prospects: The 1998 Revision.* New York, NY: Population Division, United Nations; 1999.

2. Sathar Z, Casterline JB. The onset of fertility transition in Pakistan. *Popul Dev Rev.* 1998;24:773–796.

3. United Nations Development Program. *Human Development Report 1997.* New York, NY: Oxford University Press; 1997.

4. Lightbourne RE, MacDonald AL. *Family Size Preferences.* Voorburg, Netherlands: International Statistical Institute; 1982. World Fertility Survey Comparative Study No. 14.

5. Hakim A, Cleland J, Bhatti MH. *Pakistan Fertility and Family Planning Survey 1996–97: Main Report.* Islamabad, Pakistan: National Institute of Population Studies and London School of Hygiene and Tropical Medicine; 1998.

6. Hashmi S, Alam K, Sheraz A. *Non-Users and Unmet Need for Contraception in Punjab: First Report.* Islamabad, Pakistan: National Institute of Population Studies; 1993.

7. *The Gap Between Reproductive Intentions and Behaviour. A Study of Punjabi Men and Women.* Islamabad, Pakistan: Population Council; 1997.

8. Robinson WC, Shah MA, Shah NM. Family planning program in Pakistan: what went wrong? *Int Fam Plann Perspect.* 1981;7:85–92.

9. Khan A. Policy making in Pakistan's population programme. *Health Policy Plann.* 1996;11:30–51.

10. Ross JA, Frankenberg E. *Findings From Two Decades of Family Planning Research.* New York, NY: Population Council; 1993.

11. Population and Human Resources Department *Effective Family Planning Programs.* Washington, DC: World Bank; 1993.

12. Shelton JD, Bradshaw L, Hussein B, Zubair Z, Drexter T, McKenna MR. Putting unmet need to the test: community-based distribution of family planning in Pakistan. *Int Fam Plann Perspect.* 1999;25:191–195.

13. Phillips JF, Simmons R, Koenig MA, Chakraborty J. Determinants of reproductive change in a traditional society: evidence from Matlab, Bangladesh. *Stud Fam Plann.* 1988;19:313–334.

14. Ali SR. *Village Relations Study.* Islamabad: Ministry of Population Welfare and the Population Council, Pakistan; 1998. Research Report No. 9.

15. Ministry of Population Welfare and the Population Council. *Situation Analysis of Village Based Family Planning Workers in Pakistan.* Islamabad, Pakistan: Population Council; 1995. Research Report No. 3.

16. Ministry of Population Welfare and the Population Council. *Initial Performance and Impact of Village Based Family Planning Workers in Four Districts in Punjab.* Islamabad, Pakistan: Population Council; 1997. Research Report No. 5.

17. *Stata Statistical Software, Release 6.0* [computer program]. College Station, Tex: Stata Corp; 1999.

18. Longford NT. *Random Effect Models.* Oxford, England: Clarendon Press; 1993.

19. Population Council, Ministry of Population Welfare, and the UN Population Fund. *Pakistan Contraceptive Prevalence Survey 1994–95. Final Report.* Islamabad, Pakistan: Population Council; 1998.

20. Cernada GP, Rob AKU, Ameen SI, Ahmad MS. *A Situation Analysis of Family Welfare Centers in Pakistan.* Islamabad, Pakistan: Population Council; 1993.

21. Ahmad T. Evaluation of the Mobile Service Units of the population welfare program in Pakistan. *Pakistan Popul Rev.* 1993;4:121–127.

22. Ahmed W. Field structures in family planning. *Stud Fam Plann.* 1971;2:6–13.

23. Ali SM, Zahid GM. Population planning in Pakistan: how to meet the challenge. *Pakistan Dev Rev.* 1998;37:523–540.

24. Population Information Program. *Traditional Midwives and Family Planning.* Baltimore, Md: Johns Hopkins University School of Public Health. 1980. Population Reports J No. 22.

25. Osborn RW. The Sialkot experience. *Stud Fam Plann.* 1974;5:123–129.

26. Becker GS, Lewis HG. On the interaction between the quantity and quality of children. *J Polit Econ.* 1973; 81:S279–S288.

27. Caldwell JC. Mass education as a determinant of the timing of fertility decline. *Popul Dev Rev.* 1980;6: 225–255.

Use of Topical Sunscreens and the Risk of Malignant Melanoma: A Meta-Analysis of 9067 Patients From 11 Case–Control Studies

Michael Huncharek, MD, MPH, and Bruce Kupelnick, BA

Objectives. This study examined the methodology of epidemiological studies that suggest use of topical sunscreen preparations is associated with increased risk of malignant melanoma.

Methods. We pooled data from observational studies using a general variance–based meta-analytic method that employed confidence intervals (previously described). The outcome of interest was a summary relative risk (RR) reflecting the risk of melanoma associated with sunscreen use versus nonuse. Sensitivity analyses were performed when necessary to explain any observed statistical heterogeneity.

Results. Combining studies that used nonheterogeneous data yielded a summary RR of 1.01, indicating no association between sunscreen use and development of malignant melanoma.

Conclusions. The available epidemiological data do not support the existence of a relationship between topical sunscreen use and an increased risk of cutaneous malignant melanoma. (*Am J Public Health.* 2002;92:1173–1177)

Malignant melanoma is 1 of the most increasingly common solid tumors over the last three decades.[1] Although increased detection may possibly account for some of the increase in incidence, other behavioral and environmental factors are likely to contribute to the current "epidemic" of this disease. Sun exposure in the form of ultraviolet-B (UVB) radiation is considered a major risk factor for the development of melanoma. Unfortunately, uncertainties regarding the impact of various host characteristics, frequency and type of sun exposure, and behavioral factors on melanoma development complicate assessment of the exact relationship between sun exposure and cancer risk.[2]

Sunscreens are able to delay sunburns and to reduce some UV-induced skin lesions, such as nonmelanoma tumors in rodents, local immunological depression, mutations of the p53 gene in keratinocytes, and the incidence of actinic keratoses in humans. As a consequence, sunscreen use is often recommended as a sun protection method, although its true impact on melanoma prevention remains obscure.

Despite uncertainties in the available epidemiological data, experimental evidence using both animal models and humans suggests that sunscreen preparations capable of reducing exposure to UVB radiation from the sun can prevent melanoma.[2] Regrettably, this finding has not been universal. In fact, some investigators suggest that sunscreen use could be a risk rather than a protective factor for malignant melanoma.[1] Although it is considered unlikely that available sunscreen preparations contain compounds with carcinogenic effects, other factors may account for this observed relationship; they include uncontrolled confounding caused by host factors and behavioral factors, such as increased sun exposure among patients who use sunscreen preparations.

This article presents the results of a meta-analysis designed to examine the impact of sunscreen use on melanoma risk. In addition to calculating an overall summary estimate of effect, the analysis also explores characteristics of the included studies that may contribute to heterogeneity of observed outcome. The resulting data may provide a clearer understanding of the role of sunscreen in preventing malignant melanoma.

METHODS

The methods used in the design and execution of this study have been described previously.[3,4] The study protocol initially developed outlined a meta-analysis to examine the risk of developing malignant melanoma associated with topical sunscreen use. Eligibility criteria for study inclusion were determined prospectively, as were the specific data elements to be extracted from each published report. The study protocol also included details of the planned statistical analysis.

We used a data extraction form designed for recording relevant information from each selected report. Two researchers performed data extraction, with differences in extraction forms resolved by consensus. Other data collected but not included in the eligibility criteria were number of patients in each study, study odds ratios and 95% confidence intervals (CIs), and type of statistical adjustments made, if any, by individual study authors.

Literature Search

Information retrieval was performed with previously described methods.[3] Briefly, we conducted a MEDLARS search of literature published between January 1966 and December 1999, as well as a review of CancerLit and the CD-ROM version of Current Contents. The search criteria included all languages. If a series of articles was published, all data were retrieved from the most recent article. The literature search also included hand searches of bibliographies of published reports, review articles, and textbooks.

The initial citations (in the form of abstracts) from this literature search were screened by a physician–investigator to exclude those that did not meet protocol-specified inclusion criteria. Reasons for rejection included studies of designs other than case–control; cohort or randomized controlled trials; animal or in vitro studies; stud-

ies including nonmelanoma skin cancer patients not stratified by tumor type; abstracts; and review articles. Copies of full articles for the remaining citations were obtained and screened according to the following additional eligibility criteria: (1) published case–control or cohort studies, (2) studies enrolling adult patients only (i.e., ≥18 years of age), (3) availability of data on frequency of sunscreen use, (4) specified selection criteria for case and control subjects, and (5) availability of data on the outcome of interest (i.e., proportion of patients with a diagnosis of malignant melanoma).

Statistical Analysis

We performed data analysis according to meta-analytic procedures described by Greenland.[5] This method of meta-analysis is a general variance–based method employing confidence intervals. Because the variance estimates are based on the adjusted measures of effect and on the 95% confidence interval for the adjusted measures, the confidence interval methods do not ignore confounding factors and are the preferred methodology for nonrandomized data.

For each included study, we derived odds ratios reflecting the risk of developing malignant melanoma associated with sunscreen use and determined the natural logarithm of the estimated relative risk (RR) for each data set followed by an estimate of the variance. We used the estimate of the 95% confidence interval from each study to calculate the variance of each study's measure of effect.

We calculated a weight for each included study as 1/variance followed by a summation of the weights. We then determined the product of the study weight and the natural logarithm of the estimated relative risk and performed a summation of these products. Finally, we calculated a summary RR and 95% confidence interval.[5]

Before estimation of the summary RR, we performed a statistical test for heterogeneity (Q). This procedure tests the hypothesis that the effect sizes are equal in all studies.[3] If Q exceeds the upper-tail critical value of the χ^2 distribution at $k-1$ degrees of freedom (where k is the number of studies analyzed or the number of statistical comparisons), the observed variances in study effect sizes are significantly greater than would be expected by chance if all studies shared a common population effect size. If the hypothesis that the studies are homogenous is rejected, the studies are not measuring an effect of the same size, and calculation of a pooled estimate of effect must be done cautiously. Possible explanations for the observed heterogeneity must then be sought to provide the most rational interpretation of the summary RR. Therefore, we performed sensitivity and/or further stratified analyses as needed based on the magnitude of Q; these analyses are discussed below.

RESULTS

We obtained a total of 166 citations from the electronic and manual literature search. Initial screening of these citations yielded 13 that appeared to meet specified protocol criteria.[6–18] On further review of the full published manuscripts, we found that 2 articles did not meet inclusion criteria. Herzfeld et al.[17] did not clearly distinguish between the use of suntan lotion and sunscreen preparations. Because of its lack of stratification, this study was not included in the meta-analysis. The study by Autier and Dore[18] also did not meet inclusion criteria, because it examined only the influence of sun exposure during childhood and adulthood on melanoma risk using a "sun exposure index" created by the authors. Data on sunscreen use had been collected in a earlier case–control study by this group[15] that was included in the meta-analysis. The remaining 11 published articles composed the database for the present analysis.

Table 1 provides an overview of the 11 case–control studies in the meta-analysis. Overall, the 11 study reports encompassed a total of 9067 patients. Also shown in the table are the odds ratios calculated for each individual report included in the pooled analysis, along with its 95% confidence interval. An odds ratio greater than 1.0 indicates an increased risk of melanoma associated with sunscreen use. Frequency of sunscreen use is given as noted by the authors of each study. The most frequent reported use was compared with "never used" in the pooled analysis.

All but 3 studies[6,7,16] had odds ratios greater than 1.0, demonstrating that the vast majority of case–control studies indicate that sunscreen users have a greater risk of melanoma than do nonusers. Combining data from all 11 reports gave a summary RR of 1.11 (95% CI=0.37, 3.32), a statistically nonsignificant result. Calculation of Q for this meta-analysis resulted in a value of 42.0 (Table 2). With 10 degrees of freedom, this yielded a P value of <.001, a highly significant result. A Q of this magnitude indicates that the pooled studies are heterogeneous—that is, the studies are not measuring an effect of the same size. Therefore, the validity of the summary RR is questionable, and sources of heterogeneity needed to be sought.

We performed several sensitivity analyses to evaluate possible sources of the observed statistical heterogeneity. As indicated in Table 1, all but 2 studies adjusted for potential confounders. The analyses by Klepp et al.[10] and Graham et al.[11] found a positive association between sunscreen use and increased melanoma risk. A sensitivity analysis we performed excluded both of these reports from the meta-analysis. Recalculation of Q yielded a value of 31.2 (P<.001), which indicated that heterogeneity remained despite removal of these data from the analysis (i.e., other factors were accounting for the variation across studies).

Our examination of the data presented in Table 2 showed that the study by Rodenas et al.[16] had a variance substantially greater than that of any other study in the pooled analysis (0.674). An additional sensitivity analysis, omitting these data from the calculation of a summary RR, yielded a Q of 37.7; with 9 degrees of freedom, the corresponding P value was <.001. A Q of this magnitude indicates persistent heterogeneity.

Table 3 outlines selection criteria for case and control subjects and also indicates whether study data were derived from population-based or hospital-based sources. Seven studies used hospital-derived case and control patients, totaling 4231 subjects.[6,9–11,14–16] Because the source of study subjects may bias results through such factors as referral patterns, we stratified the available data to explore this possibility. We pooled the 4 reports that used population registry–derived subjects[7,8,12,13] and calculated a Q statistic (4836 study subjects total); Q equaled 4.9 (P=.18). With 3 degrees of freedom, this result was

TABLE 1—Overview of Included Studies

Reference Authors	No. Patients	No. Controls	Frequency of Sunscreen Use	Odds Ratio (95% CI)	Adjustments
Espinoza Arranz et al.[6]	116	235	Ever vs never	0.48 (0.34, 0.71)	Skin type, nevi count, age
Holly et al.[7]	452	930	Almost always vs never	0.48 (0.33, 0.67)	Sunburns up to 12 yrs of age, skin reaction to sun, host factors
Westerdahl et al.[8]	400	640	Almost always vs never	1.80 (1.10, 2.80)	History of sunburn, history of sunbathing, employment, host factors.
Wolf[9]	193	319	Often vs never	3.34 (1.81, 6.64)	Age, sex, sunbathing, host factors
Klepp et al.[10]	89	227	Often vs rarely or never	2.27 (1.26, 4.12)	None
Graham[11]	404	521	Use vs never used	2.20 (1.2, 4.1)	None
Holman et al.[12]	507	507	Ever vs never	1.15 (0.78, 1.68)	Host factors, age at arrival in Australia, ethnic origin
Osterlind et al.[13]	474	926	>10 yrs vs never	1.2 (0.9, 1.5)	Constitutional factors, sex, age
Beitner et al.[14]	523	505	Very often/often vs never	1.80 (1.2, 2.7)	Age, sex, hair color
Autier et al.[15]	418	438	Regular use vs never	1.50 (1.09, 2.06)	Age, sex, hair color, no. of holiday weeks spent in sunny climate
Rodenas et al.[16]	105	138	Always vs never	0.2 (0.01, 0.8)	Age, skin color/type, no. of nevi, no. of hrs sun exposure

Note. CI = confidence interval.

TABLE 2—Data for Analysis of Heterogeneity

Reference Authors	Weight	Variance	Odds Ratio (95% CI)
Espinoza Arranz et al.[6]	32.3	0.031	0.48 (0.34, 0.71)
Holly et al.[7]	27.0	0.037	0.48 (0.33, 0.67)
Westerdahl et al.[8]	15.9	0.063	1.80 (1.10, 2.80)
Wolf et al.[9]	10.2	0.098	3.34 (1.81, 6.64)
Klepp et al.[10]	11.1	0.090	2.27 (1.26, 4.12)
Graham et al.[11]	10.4	0.096	2.20 (1.20, 4.10)
Holman et al.[12]	25.6	0.039	1.15 (0.78, 1.68)
Osterlind et al.[13]	45.5	0.022	1.20 (0.9, 1.50)
Beitner et al.[14]	23.3	0.043	1.80 (1.20, 2.70)
Autier et al.[15]	37.0	0.027	1.50 (1.09, 2.06)
Rodenas et al.[16]	1.48	0.674	0.20 (0.04, 0.79)

Note. CI = confidence interval

not statistically significant—that is, the data were not heterogeneous and could therefore be pooled to calculate a summary RR. The resultant summary RR was 1.01 (95% CI= 0.46, 2.28), a statistically nonsignificant result. These data failed to show any relationship between sunscreen use and increased risk of melanoma.

Next, we combined the 7 reports that used hospital patient databases in a meta-analysis. Five of the 7 studies had odds ratios greater than 1.09,[9–11,14,15] suggesting an association between sunscreen use and melanoma risk. Our analysis for heterogeneity yielded a Q of 36.9. With 6 degrees of freedom, the corresponding *P* value for a Q of this size was <.001, a highly heterogeneous result. Substantial heterogeneity therefore exists across these 7 studies. We obtained a summary RR of 2.41 (95% CI=0.32, 18.1). This finding provides evidence that bias associated with hospital-derived data is probably accounting for the observed positive association between sunscreen use and melanoma risk in many of the available case–control studies.

DISCUSSION

The sustained increase in malignant melanoma incidence over the past few decades highlights the fact that this disease represents a major public health management issue worldwide. In the United States alone, more than 42 000 cases are diagnosed and more than 7 000 deaths result each year.[19] Sun exposure is recognized as the most important environmental risk factor for malignant melanoma.[20] Behaviors that increase sun exposure have been suggested to be major contributors to the rising incidence. This suggestion has led to the development of measures to protect individuals from the potentially harmful effects of solar ultraviolet radiation (both ultraviolet-A [UVA] radiation and UVB), most notably topical sunscreen preparations.

If solar radiation is a primary risk factor for malignant melanoma, it is reasonable to conclude that reducing sun exposure via topical sunscreen use would be associated with reduced disease risk. However, the available epidemiological data are contradictory. In fact, the majority of studies suggest that sunscreen use is associated with an *increased* melanoma risk (see, e.g., studies cited in references 9 and 10). To address this uncertainty, we designed the present study to systematically evaluate the available data using rigorous meta-analytic techniques.

By pooling data from 11 case–control studies meeting protocol inclusion criteria (yielding a statistically nonsignificant summary odds ratio of 1.11), we demonstrated that sunscreen use is not associated with an increased risk of developing malignant melanoma. Unfortunately, further evaluation showed the data to be highly heterogeneous (i.e., the available studies are not measuring an effect size of the same magnitude), thereby making

TABLE 3—Selection of Cases and Controls

Reference Authors	Hospital- vs Population-Based	Selection Criteria—Cases	Selection Criteria—Controls
Espinoza Arranz et al.[6]	Hospital	Patients referred from Dermatology and Plastic Surgery Service to Medical Oncology Service	Age and sex matched—patients who attended the hospital due to emergencies not related to neoplasms or dermatological diseases
Holly et al.[7]	Population	Women aged 25–59. Derived from SEER cancer registry for San Francisco Bay area.	Women who lived in the same county as cases using random digit dialing. Age frequency matched.
Westerdahl et al.[8]	Population	Patients identified using Regional Tumor Registry for South Swedish Health Care region	Aged and sex matched—identified by "random sampling" from the same Regional Tumor Registry
Wolf et al.[9]	Hospital	Patients presenting to Dept. of Derm. at Univ. of Graz. between 6/93 and 7/94	Same as cases except without history of skin cancer
Klepp et al.[10]	Hospital	Melanoma patients admitted to the Norwegian Radium Hospital with diagnosis of melanoma between 1/74 and 5/75	Same as cases except patients had diagnoses of lymphoma, testicular cancer, and bone and soft tissue tumors
Graham et al.[11]	Hospital	Consecutive patients with melanoma seen between 1974 and 1980	Patients with nonmelanoma cancers seen over the same time period
Holman et al.[12]	Population	Patients aged less than 80 years in Western Australia diagnosed with melanoma between 1/80 and 11/81	Same source. Matched by sex, 5-year birth period, and electoral subdivision.
Osterlind et al.[13]	Population	Patients identified via national population register of residents of East Denmark	Sex and age matched from same source
Beitner et al.[14]	Hospital	Patients seen at the Dept. of Dermatology Karolinska Hospital from 2/78 to 12/83	Age and sex matched. Derived from a population register covering Stockholm county
Autier et al.[15]	Hospital	Consecutive patients seen at 5 hospitals between 1/91 and an unspecified time in 1994	Derived from same hospital registries
Rodenas et al.[16]	Hospital	All patients diagnosed with melanoma seen at Univ. of Grenada Hospital between 1989 and 1993	"Random" selection of controls from visitors to patients at same hospital without acute disease

the validity of the summary odds ratio questionable. We then explored reasons for the observed heterogeneity.

Several constitutional factors are accepted as important risk factors for melanoma; these include presence of nevi, having red or fair hair color, freckling, and having blue eye color.[14] Failure to control for possible confounders could certainly contribute to the observed statistical heterogeneity. Two of the studies used in the meta-analysis[10,11] did not adjust for such factors. Nonetheless, our sensitivity analysis indicated that heterogeneity remained even when the data from Klepp et al. and Graham et al. were dropped from the pooled analysis.

Our careful review of study designs and selection criteria for case and control subjects suggested that the source of study subjects might contribute to a biased estimate of effect (i.e., individual study odds ratios). We found that data from the 4 studies that used population registry–derived subjects[7,8,12,13] were statistically homogeneous compared with data from studies that used hospital-derived databases. This result provided strong evidence that selection bias is an important factor contributing to the spurious finding, seen in much of the literature, of a positive association between sunscreen use and melanoma development.

Hospital-derived data are problematic because referral patterns differ widely depending on hospital location, type of facility (e.g., university vs community hospital), and practice patterns, among other factors. In addition, some studies did not provide adequate information on control patient selection. For instance, Rodenas et al.[16] reported that "controls were selected from the visitors to the hospital on a random basis" without providing details of the "random" selection process. Autier et al.[15] selected case subjects from 5 collaborating hospitals; they noted that "controls were randomly chosen in the same municipality as the cases." Again, no further details are provided on what constituted "random" selection.

Referral patterns may influence study results. If referral patterns among hospitals in a given city or region differ, the overreferral of exposed cases to one hospital implies an underreferral of exposed cases to the others. Due to "differential referral," a factor may be associated with increased disease risk in one hospital-based study and may be protective in another. In an individual study, pooling data across hospitals helps to eliminate bias from differential admission of cases. Pooling data from several sources in a meta-analysis, as done in the study reported here, has partially accomplished this. Although many individual hospital-based studies showed a positive association between sunscreen use and melanoma risk, the pooled analysis indicated that this finding was spurious.

Other factors that may affect outcome in case–control studies include "ascertainment

bias" and misclassification of exposure status (in this case, sunscreen use). One factor not considered in the available studies is the possible influence of socioeconomic status (SES). Melanoma tends to affect white-collar, educated, and urban individuals. SES is known to affect recall of some types of information and could play a role in the studies examined in our analysis in which SES was not generally accounted for.[20]

These factors may all contribute to the wide variation in outcome observed across studies that used hospital registries. In contrast, data from more than 4800 patients enrolled in population-based case–control studies showed no such variation (i.e., the data were not heterogeneous and could reliably be combined in a meta-analysis). The resulting summary RR of 1.01 (95% CI=0.46, 2.28) provides strong evidence for a lack of any positive association between sunscreen use and increased melanoma risk.

CONCLUSIONS

The relationship between sunlight and melanoma is complex. Existing data suggest that the effect of solar radiation on melanoma development is more complex than is the case for other types of skin cancer.[6] Many unanswered questions remain regarding factors that may influence melanoma development, including the type of sunlight exposure most associated with melanoma etiology, interaction with host factors possibly important in disease risk, sunburn history, and tanning ability, among others. Because of this complexity, it has been difficult to separate the effects of sun exposure per se from the effects of host factors.

Nonetheless, because sunlight remains the most important recognized etiological factor in this disease, methods to reduce exposure (including use of topical sunscreens) appear to be a rational approach to disease prevention and risk reduction. We undertook the present meta-analysis to address the counterintuitive findings of multiple case–control studies that suggest sunscreen use as a risk factor for malignant melanoma. The largely positive association seen in the existing literature appears to be due to bias inherent in study designs and uncontrolled confounding.

It is our hope that the results of the present analysis will contribute to the design of future studies addressing this issue. Until more conclusive data are available, recommending use of sunscreens as a cancer prevention strategy would appear to be prudent. ■

About the Authors

Michael Huncharek is with the Division of Radiation Oncology, Department of Clinical Oncology, Marshfield Clinic Cancer Center, Marshfield, Wis; the Meta-Analysis Research Group, Stevens Point, Wis; and St. Michael's Hospital Cancer Center, Stevens Point. Bruce Kupelnick is with the Meta-Analysis Research Group, Stevens Point.

Requests for reprints should be sent to Michael Huncharek, MD, MPH, FACA, Director, Meta-Analysis Research Group, 2740 Sunset Blvd, Stevens Point, WI 54481 (e-mail: metaresearch@hotmail.com).

This article was accepted September 24, 2000.

Contributors

M. Huncharek planned the study, collected and analyzed the data, and wrote the article. B. Kupelnick designed the literature search strategy, conducted electronic and manual literature searches, performed literature retrieval, and participated in writing the article.

Acknowledgments

Partial funding for this project was provided by the Marshfield Medical Research Foundation and by a grant from Warner-Lambert Company to the Meta-Analysis Research Group.

The authors thank Dr Henry Klassen and Lamar Wheeler for assistance with data extraction.

References

1. Donawho C, Wolf P. Sunburn, sunscreen, and melanoma. *Curr Opin Oncol.* 1996;8(2):159–166.

2. Osterlind A. Epidemiology of malignant melanoma in Europe. *Acta Oncol.* 1992;31(8):903–908.

3. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis.* New York, NY: Academic Press; 1985.

4. Cooper H, Hedges LV, eds. *The Handbook of Research Synthesis.* New York, NY: Russell Sage Foundation; 1994.

5. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev.* 1987;9:1–30.

6. Espinoza Arranz J, Sanchez Hernandez JJ, Bravo Fernandez P, et al. Cutaneous malignant melanoma and sun exposure in Spain. *Melanoma Res.* 1999;9(2):199–205.

7. Holly EA, Aston DA, Cress RD, et al. Cutaneous melanoma in women, I: exposure to sunlight, ability to tan, and other risk factors related to ultraviolet light. *Am J Epidemiol.* 1995;141(10):923–933.

8. Westerdahl J, Olsson H, Masback A, et al. Is the use of sunscreens a risk factor for malignant melanoma? *Melanoma Res.* 1995;5(1):59–65.

9. Wolf P, Quehenberger F, Mullegger R, et al. Phenotypic markers, sunlight-related factors and sunscreen use in patients with cutaneous melanoma: an Austrian case-control study. *Melanoma Res.* 1998;8(4):370–378.

10. Klepp O, Magnus K. Some environmental and bodily characteristics of melanoma patients. A case-control study. *Int J Cancer.* 1979;23(4):482–486.

11. Graham S, Marshall J, Haughey B, et al. An inquiry into the epidemiology of melanoma. *Am J Epidemiol.* 1985;122(4):606–619.

12. Holman CD, Armstrong BK, Heenan PJ. Relationship of cutaneous malignant melanoma to individual sunlight-exposure habits. *J Natl Cancer Inst.* 1986;76(3):403–414.

13. Osterlind K, Tucker MA, Stone BJ, et al. The Danish case-control study of cutaneous malignant melanoma, II: the importance of UV-light exposure. *Int J Cancer.* 1988;42(3):319–324.

14. Beitner H, Norell SE, Ringborg U, et al. Malignant melanoma: aetiological importance of individual pigmentation and sun exposure. *Br J Dermatol.* 1990;122(1):43–51.

15. Autier P, Dore JF, Schifflers E, et al. Melanoma and use of sunscreens: an EORTC case-control study in Germany, Belgium and France. The EORTC Melanoma Cooperative Group. *Int J Cancer.* 1995;61(6):749–755.

16. Rodenas JM, Delgado-Rodriguez M, Herranz MT, et al. Sun exposure, pigmentary traits, and risk of cutaneous malignant melanoma: a case-control study in a Mediterranean population. *Cancer Causes Control.* 1996;7(2):275–283.

17. Herzfeld PM, Fitzgerald EF, Hwang SA, Stark A. A case-control study of malignant melanoma of the trunk among white males in upstate New York. *Cancer Detect Prev.* 1993;17(6):601–608.

18. Autier P, Dore JF. Influence of sun exposures during childhood and during adulthood on melanoma risk. EPIMEL and EORTC Melanoma Cooperative Group, European Organisation for Research and Treatment of Cancer. *Int J Cancer.* 1998;77(4):533–537.

19. Ries LAG, Kosary CL, Hankey BF, et al. *SEER Cancer Statistics Review: 1973–1994.* Bethesda, Md: National Cancer Institute; 1997. NIH publication 97-2789.

20. International Agency for Research on Cancer. *Solar and Ultraviolet Radiation.* Monographs on the evaluation of the carcinogenic risk of chemicals to humans. Vol. 55. Lyon, France: IARC: 1992.

21. Glass R, Johnson B, Vessey M. Accuracy of recall of histories of oral contraceptive use. *Br J Prev Soc Med.* 1974;28(4):273–275.

The Perceived Impact of Privatization on Local Health Departments

Christopher Keane, ScD, John Marx, PhD, Edmund Ricci, PhD, and Gerald Barron, MPH

Objectives. This article presents nationally representative data on the effects of privatization on local health departments (LHDs).

Methods. A stratified representative national sample of 380 LHDs was drawn from a national list of 2488 departments. Telephone interviews were conducted with 347 LHD directors.

Results. One half of the directors of LHDs with privatized services reported that privatization helped the performance of core functions. Privatization often resulted in increased time needed for management and administration. More than a third of LHD directors reported concern about loss of control over the performance of privatized functions and services.

Conclusions. Privatization is part of a broader shift toward "managing" rather than directly providing public health services, yet privatization often reduces LHDs' control over the performance of services. (*Am J Public Health.* 2002;92:1178–1180)

In a previously published national survey of privatization in local health departments (LHDs), we showed that 73% of all LHDs have privatized at least 1 service.[1] Advancing prior discussions of privatization,[2,3] we identified 2 types of contracting out: (1) contracting out to a private provider a service formerly performed directly by the LHD and (2) contracting out the performance of a new service from its inception. Both types raise questions about the effect of contracting out on a government agency's ability to retain organizational authority and control.

Many governmental organizations face ideological and financial pressures to privatize services.[2,4,5] Because privatization is enthusiastically endorsed,[4–10] government department directors might expect that it will improve their department's image and standing in the community. However, privatization often requires that more resources be devoted to overseeing contracts and monitoring performance.[2,11,12] Although contracting out has often been undertaken to sharpen the focus on core organizational functions,[2,4,13] there is concern that privatization will undermine the regulatory authority of governmental agencies.[2,9] In the context of these issues, this report explores the effects of privatization on the performance of the core public health functions of assessment, assurance, and policymaking[14]; on staffing; on the image of the department; and on changes in the time spent on management and administration.

METHODS

A prior report described our methods in detail.[1] Briefly, we drew a stratified random sample of 380 LHDs from the 2488 LHDs listed in the 1997 database of the National Association of County and City Health Officials. We completed interviews with 347 LHD respondents (91.3% response rate), either directors (95% of respondents) or persons closest to that position.

RESULTS

Impact of Privatization on Core Functions

Of directors whose departments privatized services, 50% claimed that privatization helped their department's performance of core functions. Such directors believed that by not directly providing certain services, their department could better focus on the core functions. For example, one director said that privatization "has freed us up to focus on these core functions rather than spending too much money and time on clinical care."

Another 38% either believed that privatization hindered performance of the core functions or were unsure of the effect. Many of these directors were concerned about the loss of control by LHDs over various aspects of the services (Table 1). One director stated that privatization hinders performance of core functions and "would erode the infrastructure of public health. If contracted, those who actually provide the services and do hands on work would not be able to participate as well in the assessment and assurance functions."

Impact of Privatization on Time Devoted to Management and Administration

Of LHDs privatizing 1 or more services, 41% reported an increase in time spent on administration and management; open-ended responses revealed that this consisted of monitoring contractors' performances, meetings, and serving on committees with providers. The increase was highest among LHDs privatizing 2 or more services (Table 2).

Impact of Privatization on Staffing Levels

Two thirds of directors reported no change in staffing (Table 2). One fifth experienced a decrease in staffing and 13% saw an increase, often to manage contracted programs. In some cases, those reporting no overall difference had added program management staff, while reducing staff engaged in direct service provision.

Effect of Privatization on the Image of the LHD

The majority (59%) of directors believed privatization improved their department's image; open-ended responses revealed that public, business, and foundation officials often look favorably on privatization. Almost a third of directors (30%) reported no change, often commenting that "Joe Public" is unaware of LHD operations.

TABLE 1–Perceptions of Privatization's Effects on Performance of Core Functions: Selected Comments of Local Health Department (LHD) Directors[a]

Helps or Would Help LHD's Performance of the Core Functions

- "It has freed us up to focus on these core functions rather than spending too much money and time on clinical care."
- "It helps as long as the health department maintains overall control. It is positive to involve the community as stakeholders and this enhances performance of the core functions."
- "Services we have contracted out allows us [*sic*] to do the core functions. This also allows us to have funding we're saving to go into those other areas."
- "I think it can help if you're able to convince the taxpayers that funding should continue to be put toward the core functions. The problem is that the public isn't cognizant of the true importance of the core functions if so many functions are delegated."
- "[Privatization] allows us to distill our function. Direct delivery takes 99% of focus. Elevates focus to core functions, especially assurance part."
- "As long as health departments oversee the functions they delegated out, delegating should allow us to re-focus our energies on the core functions."

Hinders or Would Hinder LHD's Performance of the Core Functions

- "We wouldn't have control over services and functions, therefore performance of the core functions would be hindered."
- "I think the health department should have the responsibility for the services because it's easier to keep track of things when one entity is providing the services."
- "Hinders access in our situation. Access is a big problem. Delegation would make assurance of access difficult."
- "Privatization would erode the infrastructure of public health. If contracted, those who actually provide the services and do hands on work would not be able to participate as well in the assessment and assurance functions. Health departments need to keep their fingers on the pulse."
- "We lose our identity and accountability in the process of contracting. Other agencies take over our services. Where does it leave us?"
- "We'd have to shut our doors if the health department contracted out. Revenue would be lost as would contact with our clients. As a result performance of the core functions would be compromised."

Not Sure

- "[Privatization] helps in that it allows the health department to focus on the core functions by freeing up time. It also fosters community collaboration/involvement. It hinders in that monitoring private sector performance can be difficult and can undermine the health department input in certain area."
- "The core functions would be hindered if there was excessive privatization. The private sector tends to worry too much about the almighty dollar and this focus does hinder our core public health responsibilities."
- "It can free up time for the core functions yet too much delegation can threaten the authority of a HD [health department]. It depends on how the contracts are carried out."
- "Helps assurance, policy making, extra service providers; hinders assessment, data reporting is bad, one more channel for data to filter through."
- "I am concerned about fragmentation of public health and maintaining a certain level of staff to respond to crisis. If you have functions being performed in-house the staff are there to take on emergencies."
- "Helps in that it frees us from doing personal health, giving us more time for core functions. It hinders performance of maternal and child health and the public's perception is blurred. We are accountable but we have no authority."

[a]Comments are by directors of both LHDs with privatized services and LHDs without privatized services. Responses are to the question, "How does/would the delegation affect your local health department's performance of the core public health functions of assessment, assurance, and policymaking?"

Privatization and the Ability to Control Service Provision

A central theme in directors' discussions of the effects of privatization on core functions involved concerns about control over service provision. In addition, 36% reported loss of control as a negative outcome. Specifically, 27% of directors claimed that privatization resulted in some loss of control over functions and services and 14% found control more difficult. Open-ended explanations revealed that many directors felt it was easier to keep track of services provided in-house: privatization added another administrative layer, which made obtaining reliable information quickly more difficult. Some directors observed that although LHDs are held accountable for the performance of services, they have little control over contractors.

Loss of control over the provision of services is associated with spending increased time on administrative or managerial tasks ($P<.005$). Forty-six percent of directors who reported spending more time on administration or management mentioned some loss of control over, or difficulty in controlling, the performance of services. By contrast, loss of control is mentioned by 26% of those reporting no change in time devoted to managing and administering and by 35% of those reporting a decrease.

DISCUSSION

Health departments that privatized services were increasingly engaged in "managing" rather than directly providing public health services, especially those that privatized at least 2 services. The new management activities included monitoring, contract negotiations, and attending meetings with private sector boards or coalitions. However, half of all directors who spent more time on administration as a result of privatizing reported difficulty or problems in exercising control over service performance. Control problems included difficulty monitoring the performance of staff employed by private organizations, difficulty specifying quality indicators in contracts, scarcity of staff with adequate public health values and skills, problems obtaining timely medical information, problems coordinating services, and weakened enforcement mechanisms.

A large proportion of directors believed that privatization enhanced their LHD's performance of core functions. However, the findings of this study raise questions about how LHDs can exercise authority over core public health functions if they are losing control over service performance. The Institute of Medicine report *The Future of Public Health* stated that "carrying out the assurance function requires the exercise of authority. This is not a responsibility that can be delegated to the private sector."[14] Privatization threatens an LHD's ability to carry out not only assurance but also policy making and assessment. For example, one respondent stated that privatization hindered policymaking by "relinquishing control over public health functions," thereby undermining the LHD's "authority, credibility and pull at the policy making table." Diminished control over data reporting could also impair an LHD's ability to carry out assessment. One director's experience was that privatization "hinders assessment" because "data reporting is bad, [with yet] one more channel for data to filter through."

TABLE 2–Administrative Effects of Privatization Expressed as Percentages of Local Health Departments (LHDs), by Size of Jurisdiction

	Size of Jurisdiction of LHD					
	<25 000 (n = 68)	25 000–49 999 (n = 68)	50 000–99 999 (n = 67)	100 000–349 999 (n = 71)	>350 000 (n = 73)	Total (n = 347)[a]
Change in total time LHD spends on administration and program management[b]						
Of those privatizing 1 service						
Increased	30.0	44.4	25.0	20.0	16.7	31.3
Stayed the same	50.0	44.4	50.0	40.0	50.0	47.9
Decreased	20.0	0.0	25.0	40.0	16.7	17.5
Of those privatizing 2 or more services						
Increased	50.0	39.0	34.8	38.6	55.9	43.2
Stayed the same	26.7	51.2	50.0	45.6	27.1	40.0
Decreased	23.3	9.8	13.0	14.0	11.9	15.7
Effect of privatization on LHD staffing[c]						
Reduced staff	15.4	18.7	17.0	29.5	35.9	20.6
Added staff	15.4	10.4	13.2	11.5	12.5	13.0
No overall change in staff	69.2	70.8	69.8	59.0	51.6	66.5
Effect of privatization on overall image of LHD[d]						
Improved	61.5	46.9	64.8	59.7	64.6	58.7
Worsened	5.1	2.0	3.7	3.2	3.1	3.7
No change	23.1	42.9	27.8	30.6	23.1	29.7
Don't know/inconclusive: some positive, some negative	9.3	8.1	3.8	6.5	9.2	7.9

[a]Sample sizes reported in this column are unweighted. All percentages reported in the table are weighted.
[b]Responses to the question, "How has privatization changed the proportion of time your local health department has to spend on administration and program management? Has it increased, decreased or stayed the same?"
[c]Responses to the question, "What effect has privatization had on your staffing? Have you reduced staff? Added staff? Or kept staffing at about the same size?"
[d]Responses to the question, "How has privatization affected the overall image of the local health department?"

Unlike the situation that has confronted clinical medicine, direct corporate managerial takeover of basic public health functions is improbable because there is little comparable opportunity for dramatic profit. LHDs are experiencing a different form of managerial incursion; namely, they are devoting more time and resources to managerial and administrative activities as they contract out services to private organizations. ■

About the Authors

Christopher Keane, Edmund Ricci, and John Marx are with the Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pa. John Marx is also with the Department of Sociology, University of Pittsburgh. Gerald Barron is with the Allegheny County Health Department, Pittsburgh, Pa.

Requests for reprints should be sent to Christopher Keane, ScD, 211 Parran Hall, 130 DeSoto St, Pittsburgh, PA 15261 (e-mail: crkcity@pitt.edu).

This article was accepted March 23, 2001.

Contributions

C. Keane designed the study questions and methods, drafted the questionnaire, analyzed the data, designed and supervised data collection, and wrote the initial drafts of the article. J. Marx initially conceived of studying the privatization of public health, worked on the study design, revised the questionnaire, and revised the manuscript. E. Ricci secured initial funding and significantly edited the manuscript. G. Barron participated in pilot studies, enlisted the participation of health departments, and reviewed the study instrument and manuscript.

Acknowledgments

This research was funded by the Centers for Disease Control and Prevention through a cooperative agreement through the Association of Schools of Public Health.

We thank the National Association of City and County Health Officials for their support and advice. We also want to express our gratitude to the reviewers at the Journal for their careful reading and useful suggestions.

References

1. Keane C, Marx J, Ricci E. Privatization and the scope of public health: a national survey of local health department directors. *Am J Public Health.* 2001;91:611–617.

2. Sclar E. *You Don't Always Get What You Pay For: The Economics of Privatization.* Ithaca, NY: Cornell University Press; 2000.

3. Halverson PK, Kaluzny AD, Mays GP, Richards TB. Privatizing health services: alternative models and emerging issues for public health and quality management. *Qual Manag Health Care.* 1997;5:1–18.

4. Osborne D, Gaebler T. *Reinventing Government.* New York, NY: Plume; 1993.

5. Poole RW. *Cutting Back City Hall.* New York, NY: Universe Books; 1980.

6. Friedman M, Friedman R. *Free to Choose.* New York, NY: Harcourt Brace Jovanovich; 1980.

7. Savas ES. *Privatization: The Key to Better Government.* Chatham, NJ: Chatham House Publishers; 1987.

8. Fitzgerald R. *When Government Goes Private: Successful Alternatives to Public Services.* New York, NY: Universe Books; 1988.

9. Starr P. The new life of the liberal state: privatization and the restructuring of state–society relations. In: Ezra N. Suleiman EN, Waterbury J, eds. *The Political Economy of Public Sector Reform and Privatization.* San Francisco, Calif: Westview Press; 1990.

10. Sappington DE, Stiglitz J. Privatization, information and incentives. *J Policy Anal Manage.* 1987;6:567–582.

11. Ramanadham VV, ed. *Privatization and After: Monitoring and Regulation.* New York, NY: Routledge; 1994.

12. Belluck P. As more prisons go private, states seek tighter controls. *New York Times.* April 15, 1999:A1.

13. Wall S. Transformations in public health systems. *Health Aff.* 1998;17:65–89.

14. Committee for the Study of the Future of Public Health, Institute of Medicine. *The Future of Public Health.* Washington, DC: National Academy Press, 1988.