SEX-SPECIFIC REPORTING OF SCIENTIFIC RESEARCH

A WORKSHOP SUMMARY

Theresa M. Wizemann, Ph.D.
Rapporteur

Board on Population Health and Public Health Practice

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, NW • Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

This study was supported by Contract N01-OD-4-2139, TO 246 between the National Academy of Sciences and the National Institutes of Health. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number-13: 978-0-309-22524-3
International Standard Book Number 10: 0-309-22524-8

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America.

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine). 2012. *Sex-specific reporting of scientific research: A workshop summary.* Washington, DC: The National Academies Press.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."
—Goethe

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PLANNING COMMITTEE ON SEX-SPECIFIC REPORTING OF SCIENTIFIC RESEARCH: A WORKSHOP[1]

NANCY E. ADLER (*Chair*), Professor of Medical Psychology, Vice Chair of the Department of Psychiatry, Director of the Center for Health and Community, University of California, San Francisco
MARIETTA ANTHONY, Director of Women's Health, Critical Path Institute
FLOYD BLOOM, Executive Director of Scientific Communications, Professor Emeritus, Scripps Research Institute
JEROME P. KASSIRER, Distinguished Professor of Medicine, Tufts University School of Medicine
JON LEVINE, Director of the Wisconsin National Primate Research Center, Professor of Neuroscience, University of Wisconsin–Madison
HAROLD C. SOX, Editor Emeritus, *Annals of Internal Medicine*

Study Staff

MICHELLE C. CATLIN, Study Director
TREVONNE WALFORD, Research Assistant
ROSE MARIE MARTINEZ, Director, Board on Population Health and Public Health Practice
HOPE HARE, Administrative Assistant
NORMAN GROSSBLATT, Senior Editor

[1]Institute of Medicine planning committees are solely responsible for organizing the workshops, identifying topics, and choosing speakers. Responsibility for the published workshop summaries rests with the workshop rapporteur and the institution.

Reviewers

This workshop summary has been reviewed in draft form by persons chosen for their diverse perspectives and technical expertise in accordance with procedures approved by the National Research Council Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published workshop summary as sound as possible and to ensure that the summary meets institutional standards of objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We thank the following for their review of the summary:

Nancy E. Adler, Ph.D., Lisa and John Pritzker Professor of Psychology, University of California, San Francisco and Chair, Workshop Planning Committee

Jeanne Miranda, Ph.D., Professor, Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles

Rita F. Redberg, M.D., Director, Women's Cardiovascular Services, University of California, San Francisco Medical Center

John B. Wong, M.D., Professor of Medicine and Chief, Division of Clinical Decision Making, Informatics, and Telemedicine, Tufts Medical Center, Tufts University School of Medicine

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the workshop summary before its release. The review of the summary was overseen by **Kristine M. Gebbie, Dr.P.H., R.N.,** Adjunct Professor,

Flinders University School of Nursing and Midwifery, Adelaide, South Australia. Appointed by the Institute of Medicine, she was responsible for making certain that an independent examination of the workshop summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the workshop summary rests entirely with the institution.

Contents

Workshop Summary

OVERVIEW

On August 30, 2011, the Institute of Medicine hosted a workshop, Sex-Specific Reporting of Scientific Research, sponsored by the Office of Research on Women's Health (ORWH) of the National Institutes of Health (NIH).[1] The workshop explored the need for sex-specific reporting of scientific results; potential barriers and unintended consequences of sex-specific reporting of scientific results; experiences of journals that have implemented sex-specific requirements, including the challenges and benefits of such editorial policies; and steps to facilitate the reporting of sex-specific results. Presenters and participants included current and former editors of scientific journals, researchers, and scientists and policymakers from government, industry, and nonprofit organizations. Presentations and discussions highlighted the importance to both women and men of having sex-specific data, the problems with sample size and financial constraints for conducting the research, the appropriateness of sex-specific analyses, and the limitations of journal policies to change experimental designs. During closing remarks, the planning committee chair summarized some of the individual suggestions discussed for advancing sex-specific reporting as: identifying the sex of

[1]The workshop was planned in collaboration with the Institute of Medicine's Board on Health Sciences Policy and was organized by an independent planning committee whose role was limited to identification of topics and speakers. The present summary was prepared by the rapporteur as a factual summary of the presentations and discussions that took place at the workshop. Statements, recommendations, and opinions expressed are those of individual presenters and participants and are not necessarily endorsed or verified by the National Academies, and they should not be construed as reflecting any group consensus.

populations in journal populations, sharing of sex-identified raw data, giving "extra credit" in review to manuscripts that include sex-specific information, and requiring sex-stratified analyses where applicable.

INTRODUCTION

The number of women participating in clinical trials has increased over the last two decades, but women are still underrepresented in clinical trials in general. Some of the overall increase can be attributed to the greater number of women-only trials (of therapies for diseases that affect only women). Even when women are included in clinical trials, the results are often not analyzed separately by sex.

On August 30, 2011, the Institute of Medicine (IOM) Board on Population Health and Public Health Practice hosted the workshop Sex-Specific Reporting of Scientific Research. Nancy Adler, professor of medical psychology at the University of California, San Francisco, and workshop chair, cited a recent review of high-impact publications of clinical studies, including clinical trials and prospective cohort studies, of non–sex-specific cancers. It found that women constituted less than 40% of participants (Jagsi et al., 2009). Other research indicates that studies of cardiovascular disease are particularly male-biased. A review of 19 randomized controlled cardiovascular trials found that only 27% of the participants were female and that only 13 of the studies presented sex-based analyses of the data (Kim and Menon, 2009). That bias is often unintentional, Adler noted. In designing inclusion criteria, for example, early age of onset of myocardial infarction and chest pain as a presenting symptom will favor the enrollment of men over women (Bairey Merz et al., 2006; Canto et al., 2007; Gurwitz et al., 1992). Similarly, the end points selected can lead to bias. Unstable angina, stroke, and unrecognized myocardial infarction are more common in women, and if a study does not include these as end points, cardiovascular disease in female participants may be underestimated. Bias in inclusion criteria in human immunodeficiency virus (HIV) studies has also been reported (Gandhi et al., 2005). In addition, there is a substantial male bias in animal models of disease, even for diseases that are more prevalent in women.

A recent IOM consensus report, *Women's Health Research: Progress, Pitfalls, and Promise* (IOM, 2010, p. 12), found that

> limitations in the design, analysis, and scientific reporting of health research have slowed progress in women's health. In-

> adequate enforcement of recruitment of women and of reporting data by sex has fostered suboptimal analysis and reporting of data on women from clinical trials and other research. That failure has limited possibilities for identifying potentially important sex or gender differences. New methods and approaches are needed to maximize advances in promoting women's health.

On the basis of that finding, the IOM committee recommended (IOM, 2010, p. 13) that

> the International Committee of Medical Journal Editors and other editors of relevant journals should adopt a guideline that all papers reporting the outcomes of clinical trials report on men and women separately unless a trial is of a sex-specific condition. . . . The National Institutes of Health should sponsor a meeting to facilitate the establishment of the guidelines.

To address the recommendation, the NIH ORWH requested that the IOM convene a 1-day workshop to explore the benefits of and barriers to sex-specific reporting of scientific data. The workshop brought together representatives of academe, industry, government, and research-advocacy organizations and editors of leading scientific and medical journals to consider

- the need for sex-specific reporting of scientific results;
- potential barriers to and unintended consequences of sex-specific reporting of scientific results;
- experiences of journals that have implemented sex-specific requirements, including the challenges and benefits of such editorial policies; and
- steps to facilitate the reporting of sex-specific results.

The present report summarizes the presentations and discussions by the expert panelists. The first session focused on why sex-specific reporting is important. Panelists highlighted historical and current events that have hindered or helped to advance the study of women. In the next session, panelists in academe discussed the challenges of collecting, analyzing, and reporting sex-specific data from the researcher's perspective. That was followed by two panels of leading journal editors who shared their experiences in developing and implementing editorial policies and the implications of sex-specific reporting policies for journals. In the

BOX 1
A Brief History of Inclusion of Women in Clinical Research Funded by the National Institutes of Health

- **Late 1980s:** Concerns were first raised that clinical research on conditions that affect both women and men was being conducted primarily in a homogeneous white male population but that the results were being applied in medical practice to both men and women of all races.
- **1990:** ORWH was established in NIH to ensure that women are included in NIH-funded clinical studies.
- **1993:** NIH policies on the inclusion of women and minorities in clinical research became law as a result of the NIH Revitalization Act of 1993 (PL 103-43). The act included four major requirements. NIH must
    - ensure that women and members of minority groups and their subpopulations are included in all human-subjects research;
    - ensure that in phase 3 clinical trials, women and minorities and their subpopulations are included in such a way that valid analyses of differences in intervention effect can be performed;
    - not allow cost to be used as an excuse for excluding these groups; and
    - initiate programs and support for outreach efforts to recruit these groups into clinical studies.
- **2000:** The U.S. General Accounting Office (GAO; now the Government Accountability Office) reported that NIH had made substantial progress in strengthening and implementing its policy on inclusion of women in clinical trials.

SOURCE: Clayton, 2011.

final session, members of the workshop planning committee and others reflected on the discussion and summarized the individual suggestions made over the course of the day for advancing sex-specific reporting of scientific research.

INCLUSION OF WOMEN IN CLINICAL TRIALS FUNDED BY THE NATIONAL INSTITUTES OF HEALTH

On behalf of the workshop sponsor and as background for the discussions, Janine Clayton, deputy director of ORWH, provided a brief history of the inclusion of women in NIH-funded clinical studies (Box 1). Despite the success of NIH efforts to enhance enrollment of women,

the law and policy apply only to NIH-funded studies, not to studies done by or supported by other agencies or entities. In addition, NIH cannot require editors and journals to mandate inclusion of analysis by sex in reports of studies. As a result, key health data are not reaching other researchers and the public.

To begin to address that situation, ORWH established a working group of scientific-journal editors as an ad hoc subgroup of the Advisory Committee on Research on Women's Health. In 2001, the group issued a statement calling on scientific journals to require that, where appropriate, clinical and epidemiologic studies be analyzed to see whether there is an effect of sex; if there is no effect, that should also be reported. Any statistical limitations of such analyses should be made clear. To date, however, very few journals have adopted such a policy. Clayton cited the *Journal of the National Cancer Institute* (*JNCI*) as an example of journals that address sex-specific analysis in their instructions for authors.

The continuing challenge, Clayton concluded, is to get sex-differences research accomplished *and* the results reported, from basic through applied research.

WHY SEX-SPECIFIC REPORTING IS IMPORTANT

Early History

Ameeta Parekh, director of research and development in the Office of Women's Health of the U.S. Food and Drug Administration (FDA), reminded participants that the severe birth defects associated with thalidomide use by pregnant women in the 1960s led to a conservative approach to testing of new drugs in women. In 1977, FDA issued *General Considerations for the Clinical Evaluation of Drugs,* which stated that "women of childbearing potential should be excluded from the earliest dose-ranging studies." Although the guidance went on to state that such women could be included in further studies if additional evidence had been amassed on the safety or preclinical toxicity of a drug, that exclusion inadvertently led to the underrepresentation or exclusion of women from *all* clinical trials. The exclusion of women from clinical research was not generally questioned, because sex was not recognized as a variable in health research and was not considered to be a factor that could affect health and illness. In addition, investigators believed that women were more difficult to study because they introduced more variables (for example, hormonal cycles) and were difficult to recruit. The result of not studying women is gaps in our knowledge and understand-

ing of the differences between men and women with regard to treatments and response.

Carolyn Clancy, director of the Agency for Healthcare Research and Quality (AHRQ), added that the first randomized trial of estrogen to prevent heart disease was conducted in the early 1960s in men. The effect of estrogen on heart disease in women was not studied in a randomized trial until the Women's Health Initiative, 35 years later.

The Need to Study Both Sexes

Martha Nolan, vice president for public policy of the Society for Women's Health Research (SWHR), said that there is a great need to identify biologic and physiologic differences between men and women and to understand the implications of the differences for diagnosis and treatment. She noted, for example, that more women than men take antidepressants; women respond more slowly and are less likely to achieve an optimal response to treatment for depression; and women are more likely to stop using the medication because of adverse events. There are many other examples of differences that are not fully understood. Female athletes, especially those in contact sports, sustain a higher percentage of concussions during play than male athletes do, but virtually all the literature and mass-media attention is on male football and ice-hockey players. Transplantation of donor organs from females is less successful than transplantation from males. Boys are more likely than girls to receive a diagnosis of peanut allergy early in life, but by the age of 24, more women than men are receiving the diagnosis.

Parekh provided further support for the need to study both sexes. Women make up more than 50% of the U.S. population (50.7% according to the 2010 U.S. Census) and on the average outlive men (80.7 years vs 74.8 years). Many diseases place a heavier burden on women than on men (consider, for example, heart disease, cancer, rheumatoid arthritis, lupus, and osteoporosis); however, treatment guidelines are based largely on data on men. Women also rely more on medical systems than men do and are likely to seek treatment sooner.

Jesse Berlin, vice president of epidemiology at Johnson & Johnson Pharmaceutical Research and Development, said that sex-specific reporting helps to define the most appropriate population for treatment and to determine whether benefits or harms differ by sex. Differences between the sexes are more than just pharmacokinetic, however. For example, Berlin cited a recent report that describes sex-specific differences in cell regulatory processes (Mittelstrass et al., 2011).

Sex-specific analysis and reporting are not just "women's health" issues. Better data on women would be better data for everyone, Clancy said. Sex-specific data could allow guidelines to be more specific and allow clinicians to better tailor care to individuals.

Speakers also presented examples of the importance of sex-specific differences. One example of critical differences between males and females is drug-induced electrocardiographic changes. Parekh explained that several drugs withdrawn from the market were associated with prolongation of the QT interval (a measure of cardiac repolarization) and torsades de pointes (a potentially fatal form of polymorphic ventricular tachycardia). Women have a longer baseline QT interval and a higher propensity for drug-induced QT prolongation, and they are two to three times more likely to develop torsades than men. The effects of drugs being studied for cardiotoxicity, Parekh said, need to be looked at and understood in both men and women.

A more recent example of the importance of sex-based data is A Diabetes Outcome Progression Trial (ADOPT), a randomized controlled trial (RCT) that compared rosiglitazone with metformin and glyburide over several years. The overall fracture rate associated with rosiglitazone use was higher than that associated with glyburide and metformin, but analysis by sex showed that women had a rate of fractures twice that of men (Kahn et al., 2008). As a result, the label for rosiglitazone includes data on the increased fracture risk for women.

Raising Awareness of Sex Differences

A 1992 GAO review of FDA policies and pharmaceutical-industry practices found that women were not adequately included in clinical studies and that data were not analyzed for sex differences with any consistency, and that consequently there was a lack of understanding of sex differences (GAO, 1992). As a result, Parekh said, FDA issued several new guidance documents and regulations. The 1977 policy that mentioned exclusion of women of childbearing potential was reversed through the 1993 guideline *Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs*, which recommended collection and analysis of data on sex differences in effectiveness, adverse effects, pharmacokinetics, and pharmacodynamics.

The 1998 investigational new drug (IND) application and new drug application (NDA) regulation, also called the demographic rule, requires NDA submissions to provide safety and effectiveness data and IND submissions to tabulate numbers of participants according to age,

race, and sex. In 2000, FDA issued the clinical-hold rule, which permits FDA to stop IND studies of treatments for life-threatening diseases if women are excluded because of their reproductive potential.

Parekh noted that data reported in poster sessions at a recent Drug Information Association meeting indicated that analysis of safety and efficacy data by sex has been increasing—around 75% of clinical trials in 2007–2009 reported analysis by sex—and a review of approved product labels found that nearly all included pharmacokinetic information by sex.

With regard to reporting in the literature, Nolan said that a decade ago the NIH ORWH, in collaboration with SWHR, convened a meeting of scientific-journal editors to discuss the development of specific instructions for authors and reviewers about the analysis of clinical-trial data by sex. However, in an informal survey of 11 science journals[2] conducted by SWHR in 2010, only *JNCI* and *Circulation* required reporting of sex differences; the others did not set any sex-specific requirements for authors.

Nolan cited several recent articles that draw attention to the need to consider sex differences. In March of 2010, an article in *Science* reported on sex bias in animal models and predicted that reporting would change if journals adopted a common set of guidelines for manuscripts to provide details on the sex of the animals used and required authors to state their rationale for studying only one sex and the implications of not studying the other (Wald and Wu, 2010). A June 2010 editorial in *Nature* suggested that funding agencies should require researchers to justify sex inequalities in grant proposals and should favor proposals that include both sexes; that FDA should ensure that physicians and the public are aware of sex differences in drug reactions and dosages; and that medical schools should train physicians in how diseases, symptoms, and drug responses can differ by sex (Putting gender on the agenda, 2010). The editorial also noted that *Nature* was considering whether to require authors to document the sex of animals in published papers. Finally, an article in the *New England Journal of Medicine* in June 2010 noted how the global H1N1 influenza pandemic disproportionately affected pregnant women and stressed the need for inclusion of pregnant women in clinical trials (Goldkind et al., 2010).

---

[2] *Journal of the National Cancer Institute (JNCI)*, *Circulation*, *JAMA*, *New England Journal of Medicine*, *Endocrinology*, *American Journal of Physical Medicine & Rehabilitation*, *BMJ*, *Lancet*, *Immunology*, *Gastroenterology*, and *Urology*.

The Department of Defense and the Department of Veterans Affairs are beginning to examine sex differences, such as how psychologic and physical health conditions affect female soldiers and veterans, Nolan said. These agencies are also reporting when research shows no difference between the sexes.

Great strides have been made in raising public awareness about sex-based differences in cardiovascular, muscular, skeletal, and behavioral health and disease, but only rarely are medical-care options tailored to the patient's sex, Nolan said. She suggested that it could take less time for research to be translated into medical practice if major journal publishers required analysis by sex and reporting of differences found or the lack thereof.

Barriers to Studying Sex Differences

There are both technical and political barriers to advancing knowledge of sex differences. Clancy described an imbalance between the fear of not knowing what the health-related differences between men and women are and the fear that identifying such differences is somehow impolitic or inappropriate. When the fear and concern associated with not knowing overpower concerns about the influence of politics on science, studying sex differences will become straightforward, she said.

There are methodologic challenges to studying population subgroups, such as males and females. A primary issue in breaking down data by sex is sample size. Berlin asked, Are two separate, adequately powered studies, one in each sex, needed? Or can a single study have sufficient statistical power to detect interaction? Separate studies of men and women risk confounding. Separate studies of men and women might use different doses as in, for example, studies of aspirin and myocardial-infarction prevention. It could then be difficult to tell whether differences in outcome were due to different doses, sex, or other factors. Instead, conducting two studies, each with both men and women, might allow stratification of both studies by sex and provide replication for sex-specific findings. Alternatively, meta-analytic principles could be applied to a program of development and testing.

A barrier to meta-analysis is availability of data. Clancy noted that the opportunity to conduct meta-analyses often rests on the goodwill of investigators in sharing data from clinical trials. As data collection has moved from paper to electronic form, the technical barriers to data-sharing have diminished. The unanswered question is who owns the data, particularly when studies have been funded with taxpayer dollars. With

regard to sample size, meta-analysis of clinical data can be a valuable exercise before investment in a large clinical trial—it can help in designing trials strategically.

Other Groups

Although the focus of the workshop was on sex-specific analysis and reporting, some panelists pointed out that race and ethnicity may also be clinically relevant, as may other clinically, genetically, or socially defined characteristics. Berlin cited Freedman and colleagues (1995), who discussed the possibility of finding clinically unimportant but statistically significant differences or clinically important but statistically nonsignificant differences and argued against separate results in the absence of a priori evidence of subgroup differences. Berlin argued, however, that such clinical-trial results can point to basic science and the needs for further elucidation.

Clancy referred participants to a 2009 Kaiser Family Foundation report, *Putting Women's Health Care Disparities on the Map: Examining Racial and Ethnic Disparities at the State Level*.[3] The principles being discussed in the present workshop do not refer only to definitions of gender and sex but extend to other population groups as well, she stressed.

Managing the Data

Parekh highlighted several current FDA initiatives, including one focused on standardizing the data that are electronically submitted to FDA so that analysis of data on women and other populations is easier.

Clancy raised the concept of a learning health care system whereby medical knowledge is advanced by making use of the substantial amounts of data and other information collected every day in the provision of health care. The implementation of electronic health-record systems is a key component of a learning health care system. Many professional societies and other organizations have created patient-level registries, which offer another method of collecting data. Clancy added that AHRQ is using American Recovery and Reinvestment Act funds designated for comparative-effectiveness research and patient-centered outcomes research to develop "a registry of registries" that will be

[3]See http://www.kff.org/minorityhealth/7886.cfm (accessed August 26, 2011).

functionally interoperable with the clinical-trial registration database, ClinicalTrials.gov.

THE RESEARCHER PERSPECTIVE: COLLECTING, ANALYZING, AND REPORTING SEX-SPECIFIC DATA

Researchers encounter barriers to the reporting of sex-specific biomedical research results well before the publication stage, said session moderator Jon Levine, director of the Wisconsin National Primate Research Center and editor-in-chief of *Frontiers in Neuroendocrinology*. Challenges emerge in designing experiments, applying for grants, and making the most of limited funding inasmuch as these activities build on the existing knowledge base, which is historically biased toward males.

Collecting the Data: Sex in Biomedical Research

The Politics of Sex Differences

Biases against studying females are embedded in the research culture, and there are numerous misconceptions, said Larry Cahill, professor of neurobiology and behavior at the University of California, Irvine. In neuroscience, for example, some think that if there is no behavioral difference between the sexes, there is no brain difference. It is known, however, that identical behaviors can be manifested through different neurobiologic mechanisms. Others assert that consideration of sex differences makes things more complicated. But analyzing data by sex can sometimes provide clarity.

Cahill offered an example of sex differences in brain function from his work on emotional memory. He discovered that the amygdala operates differently in men and the women when they watch the same emotional event; activity in the left-hemisphere amygdala is more predictive of memory of a given event in women, while activity in the right-hemisphere amygdala is more predictive of memory of the same event in men.

The greatest obstacle to moving forward, Cahill said, is the profound biases that exist against the consideration of sex differences. Such biases may be even greater in studies of the brain. Sex differences in the liver or kidneys are not particularly controversial, but sex differences in the brain can become a political issue. Cahill said that researchers need

to be bold and assert that male-only studies are not good enough anymore. How many false conclusions have been published as a result of failure to consider sex differences? he asked.

Male Bias in Animal Studies

One argument for the preferential use of males in animal studies, said Rae Silver, Kaplan Professor of Natural and Physical Sciences at Barnard College and Columbia University, is that females are more variable than males, partly because of cyclic reproductive hormones. There is evidence that some behaviors exhibit cycle-related variations, but in most instances there is little or no evidence that such variations make female models inappropriate.

But the arguments persist. One commentary cited by Silver described how a particular rat model of arthritis was more reproducible in male rats and that therefore far fewer males than females were needed to achieve statistically significant results. The researcher asserted, however, that the results were applicable to both sexes.

One argument that is true is that the cyclic nature of female sex hormones necessitates larger samples and more test groups in rodent work. Studying females requires more time, is more labor-intensive, and is more expensive than studying only males. Researchers must often justify the cost, as well as the increased use of animals, to their administration or institutional animal-care-and-use committee.

Silver questioned whether it would be possible to require the animal-research community to include both males and females when appropriate, as has been done for humans. Workshop participant Vivian Pinn, director of ORWH, responded that it takes great effort for NIH to monitor the mandated inclusion of women and minorities in clinical trials, and it could become overwhelming to monitor the sex of animals in studies in the same way. It would be more practical, and probably as effective, if researchers knew that information on the sex of animals was desired or required when submitting the results of studies for publication.

Sex Differences Across the Full Spectrum of Research

Denise Faustman, director of the Immunobiology Laboratory at Massachusetts General Hospital, noted that three large phase 3 clinical trials of type I diabetes products had recently failed; together, they were estimated to have cost over $3 billion to conduct. In two of those trials, she said, enrollment of males and females was fairly well balanced—

about 60–70% men. The preclinical data that informed the human trials, however, were obtained solely in female mice. She asked how these large, expensive trials might have been designed differently if pharmacokinetics or responsiveness or the stage of the disease had been studied in both male and female animals. The blame for failed clinical trials is shared equally by the clinical researchers who design and conduct the trials and the basic researchers who continue to publish data on only males or only females because it is easier. Sex differences must be considered and reported across the whole spectrum of research, Faustman said.

Subpopulations of Males and Females

A participant pointed out that males and females constitute broad subpopulations that can each be divided. For example, women in the follicular phase are different from women in a luteal phase; prepubertal women are different from postpubertal women; women taking hormone-replacement therapy are different from women who are not; and women taking estradiol and progesterone are different from women taking Premarin with hydroxyprogesterone acetate. Similarly, men taking androgens are different from men who are not. Those who understand or study reproduction or endocrinology are more aware of these issues, but researchers in other fields often are not. A challenge is how to make researchers more aware. Cahill concurred, noting that in his early work on emotional memory he simply divided subjects into men and women, but he later discovered that the division had led to false conclusions. He failed to find enhancing effects of stress hormones on memory in women, he explained, because he had not accounted for menstrual cycle or the use of hormonal contraception.

Analyzing the Data: Methods of Subgroup Analysis

Analysis and Interpretation of Subgroups

Clinical-trial data reflect groups of participants, explained John B. Wong, chief of the Division of Clinical Decision Making at Tufts Medical Center, but each patient that a physician sees is a unique individual with unique risk factors, genetic profile, experiences, and medications. The question is which of the participants in a randomized controlled trial is the same as the patient about to be treated. That is the driving force for subgroup analysis.

Wong offered a cautionary tale about subgroup analysis. The International Study of Infarct Survival, a randomized controlled trial of thousands of patients, found an overall statistical benefit of aspirin over placebo in prevention of death (ISIS-2, 1988). Sleight (2000) conducted an analysis of 12 subgroups and identified two that had a nonsignificant adverse effect. Those two subgroups, Wong revealed, were participants whose astrologic signs were Gemini and Libra. That is amusing at first, but Sleight, a noted statistician, stressed in his publication that "when clinicians believe such subgroup analyses, there is real danger of harm to the individual patient" (Sleight, 2000, p. 25).

Frequentist Statistics and Null-Hypothesis Errors

The frequentist statistical perspective, sometimes called the null hypothesis, begins with the position that a drug and a placebo are equal. Given that assumption, any observed differences in results would be due to chance. Given the alternative hypothesis that the drug and the placebo are different, observed differences in results would be due to differences between the drug and the placebo, but the null hypothesis is easier to test.

Wong pointed out the problems of type I and type II errors, and the often greater concern about the former, and the problem of statistical power where an inadequate sample size increases the chance of a type II error. Wong further explained that two types of errors can occur in association with a hypothesis that there is no difference between drug and placebo (Table 1): either the drug is truly beneficial or not, and the study either suggests that the drug is beneficial or not. A type I error occurs when the study results show that the drug is beneficial but in fact it is not—a false positive. There is less than a 0.05 probability ($\alpha = 0.05$) that this would be the case if it were assumed that the drug was equivalent to the placebo. A type II error occurs when the study results show that the drug is not beneficial but, in fact, it is—a false negative. There is usually a probability of 0.1–0.2 ($\beta = 0.2$ or $\beta = 0.2$) that this would be the case if it were assumed that the drug was equivalent to the placebo.

The consequence of these two kinds of errors in subgroup analysis is multiplicity. For a type II error, if a drug is truly beneficial (the unknown truth is that it works), the probability that the study will erroneously find the drug to be not beneficial is about 20% [1 − 80% = 20%]. Assuming that each subgroup is independent, and two subgroups are analyzed, the probability of erroneously finding the drug to be not beneficial in at least one subgroup increases to 36% [1 − (80%)(80%) = 36%]. With

TABLE 1 Errors of Hypothesis Testing

		Truth	
		Drug Beneficial	Drug Not Beneficial
Study Result	Drug Beneficial	$1 - \beta = 0.80$ Power	$\alpha = 0.05$ Type I error
	Drug Not Beneficial	$\beta = 0.20$ Type II error	$1 - \alpha = 0.95$

SOURCE: Wong, 2011, Slide 6.

12 subgroups, there is a 93% chance of an erroneous finding that the drug is not effective in at least one subgroup [$1 - (80\%^{12}) = 93\%$].

For a type I error, if the drug is truly not beneficial, the probability that the study will erroneously find it to be beneficial is 5% if there are no subgroups [$1 - 95\% = 5\%$], 10% if there are two independent subgroups [$1 - (95\%)(95\%) = 10\%$], and 46% if there are 12 subgroups [$1 - (95\%^{12}) = 46\%$].

Having described the general concerns subgroup analysis, Wong suggested Bayesian statistical inference as one possible approach to reporting of sex-based subgroups. Bayesian inference is a method of showing how knowledge or belief is altered by data (for further background, see Goodman, 1999). It provides a framework for combining prior belief or evidence with current evidence. The FDA guidance on using Bayesian methods for medical-device clinical trials, Wong said, describes it as "learning from evidence as it accumulates" (FDA, 2010, p. 5).

To illustrate the use of Bayesian inference, Wong asked: What is the probability that an asymptomatic woman 40–50 years old with a positive mammogram has breast cancer? Prior knowledge is that about 0.8% of asymptomatic 40- to 50-year-old women have breast cancer. In other words, of 1,000 asymptomatic women, based on prior knowledge of prevalence, eight (0.8%) would have breast cancer. Seven of those eight (90%) would have positive mammograms. However, 69 (7%) of the remaining 992 women who do not have breast cancer would have positive mammograms. The Bayes rule, or a Bayesian interpretation, Wong explained, would suggest that the probability of breast cancer in those with positive mammograms is 7 of the total positive mammograms (7 + 69), or 9%, because so many more women do not have breast cancer

than have breast cancer.[4] Most physicians, Wong noted, guess that the likelihood is over 90%.

Another way to look at the data is with what Wong referred to as a likelihood ratio. If a patient has a positive mammogram, the likelihood that she has breast cancer is 90%, and the likelihood that she does not is 7%. Hence, the patient is 13 times as likely to have breast cancer as not if she has a positive mammogram (90% ÷ 7% = 13).

Wong also described the Bayes factor, which compares how well a hypothesis predicts the data (for further background, see Goodman, 1999). All information from a clinical trial is taken into account in the Bayes factor, Wong noted; the Bayes factor indicates the likelihood of an effect discussed above. In essence, it is the probability of the data given the null hypothesis vs the probability of the data given the alternative hypothesis. As opposed to the frequentist statistical perspective discussed above, there is a separation between the probability of error, which is the null hypothesis, and the weight of the evidence from a particular clinical trial, which is the Bayes factor. In other words, a Bayesian integration gains strength from prior information whereas a frequentist approach cannot.

A Bayesian approach formally integrates prior knowledge with data ("sequential learning"). However, it requires a subjective prior belief or evidence; conclusions depend on the prior evidence, and different investigators may use different prior evidence (which may actually help to determine how robust the conclusions are). A Bayesian approach can be used for hierarchical modeling, which combines results or "borrows strength" from different studies. For example, if the national prevalence of diabetes in the United Kingdom is 2% with a standard deviation of 0.5% and, in a local sample of 1,000 patients in a given city, 1.5% have diabetes, with the Bayesian framework the national and local data could be integrated to estimate that 1.7% of the patients in the city have diabetes (with a 95% credible interval of 1.2–2.4%). In contrast, the frequentist approach could not integrate the national data and would estimate that 1.5% of the patients in the city have diabetes (with a 95% confidence interval of 0.8–2.5%). It has also been suggested that a Bayesian approach can be used in the design and conduct clinical trials and would facilitate flexibility, including adaptive randomization and stopping criteria (Berry, 2005).

[4]Wong referred participants to *Calculated Risks* by Gerd Gigerenzer (2002) for further discussion.

Wong pointed out, however, how assumptions about prior evidence can affect interpretation of a new study and have large effects on the conclusions drawn. Berlin suggested the need for a "research czar" that could help to facilitate some level of consistency among similar studies, for example, in common terminology and definitions. Wong noted that the Patient-Centered Outcomes Research Institute has a methodology committee that is attempting to address some of the issues, such as methodological standards, that would help facilitate assessment of data among studies.

From an industry perspective, Berlin said, a barrier to sharing clinical-trial data is that participants sign an agreement that dictates whether and how their information can be shared. He said that sponsors should develop participant agreements to facilitate sharing.

Risk Stratification

Frank Davidoff, editor emeritus of *Annals of Internal Medicine*, suggested risk-stratification analysis as an alternative to Bayesian statistics for applying clinical-trial results to an individual patient, and he referred participants to the work of Kent and Hayward (2007a,b). When multiple risk factors are used to segregate a sizable study population into risk subgroups, the difference in rates of outcomes can be as great as a factor of 50, he said. For example, a drug that demonstrates an overall beneficial effect may have virtually no beneficial effect in some subgroups, probably because their baseline risk is small to begin with. At the other extreme, the intervention may have a large clinical effect in patients who have a high baseline risk. For many researchers, risk stratification is less difficult to grasp than Bayesian analysis, and Davidoff suggested that it is statistically robust. Risk-stratification analysis can be applied to existing trials to look for differences in intervention effects among different groups, including sex. There are methodologic challenges to risk stratification, he noted, including the need for an independent determination of the risk groups, and there is a potential for type I and type II errors.

Reporting the Data

A Role for Journals

Silver referred participants to the report of a 2010 IOM workshop, Sex Differences and Implications for Translational Neuroscience

Research, which focused on defining roles for industry, government, academe, and journals in the translation of sex differences in neuroscience from bench to bedside. One of the suggestions raised at that workshop was that journal publishers set standards "for the inclusion of sex-related subject information in all publications, including sex of origin of tissues, cell lines, etc." and "establish guidelines to encourage authors to analyze data by sex and to report sex differences, or the lack thereof" (IOM, 2011, p. 77). It was noted, however, that it is not possible to study everything all the time, and one of the challenges raised at the 2010 workshop was to set priorities.

Silver cited the work of Beery and Zucker (2011), who analyzed the distribution of animal and human male and female subjects in published studies in journals in diverse biologic disciplines. The sex of subjects was not specified in a number of journals; in many cases in which sex was noted, there was a male bias. Silver noted that in nearly every discipline that was the case more often in nonhuman studies than in human studies. Silver quoted five of the recommendations of Beery and Zucker (2011, p. 570) that were based on their findings:

> If male and female models are thought to differ in response to an intervention, then the study must be designed with adequate sample size to answer the question for each sex.
>
> If prior research strongly indicates that there are no significant sex differences between male and female animals, then sex is not required in subject sex selection, but study of both males and females is both feasible and encouraged.
>
> If information about the existence of sex differences is absent or equivocal, then both sexes should be studied in numbers sufficient to permit valid analysis.
>
> Outreach training activities offering practical suggestions and additional sources of information should be made available by the NIH to help investigators design studies that fully incorporate female animals. . . .
>
> The review process for extramural funding should treat inclusion of females as a matter of scientific merit that affects funding.

Journal policies determine manuscript reporting requirements, Silver said, and if journal editors believe that it is important to know the sex of origin of a cell type that is being studied or the sex of animal or human participants, investigators will have to include that information.

Cahill suggested that for studies (of a non–sex-specific issue) in which only one sex has been used, journal editors should make the last two words of the article title "in males" or "in females." In addition to providing immediate clarity to basic researchers as they refer to the literature, this truth-in-advertising policy would raise awareness and would be a powerful statement that sex matters. Davidoff noted that that is similar to what was done in the mid-1990s in publications of randomized controlled trials. Such publications were not always easily identifiable, partly because "randomized controlled trial" was not included in the title and partly because the articles were not indexed as trials in U.S. National Library of Medicine's online database (MEDLINE®/PubMed®). Proper titling and indexing of papers allow researchers to study the frequency with which types of studies are published, and allow meta-analyses to be done more quickly, easily, and completely.

Sex-Based Comparisons vs Reporting of Participant Sex

Judith Lichtman, associate professor in the Department of Epidemiology and Public Health at Yale University School of Medicine, suggested that in considering standardization of journal policies for sex-specific reporting, it is important to remember that there are studies that are designed to assess sex-based differences, or of which such assessment is a natural extension, and studies in which sex-related data would be interesting to know but are not necessarily the focus. Studies designed to analyze by sex and studies that simply note the sex of participants as an observation present different methodologic issues. The extent to which sex is considered affects the focus of the work, the analyses, and often the length of the resulting paper. She suggested that requiring sex-based analysis takes study-design decisions out of the hands of the authors and peer reviewers and that comparisons drawn from studies that were not designed to assess sex differences may not be robust and could be misleading.

Sex-specific analysis presents methodologic and analytic challenges. For example, sample size is important. There must be enough data for adequate statistical power and useful comparisons. When the events being studied are very rare, there can be unintentional bias in enrollment or a disproportionate blend of women or men among study sites. There may also be differences in prevalence or risk factors between males and females, and differences in psychosocial factors may come into play in comparisons. Lichtman added that older datasets that do not have the desired distribution of men and women can still be of value

even though they may not have adequate power: relationships may be apparent, and they can help in generating hypotheses.

Lichtman described her quick survey of August 2011 issues of the *Annals of Internal Medicine*, *JAMA*, and the *New England Journal of Medicine*. Of 11 original contributions, four included some level of sex stratification of data, five that she thought probably should have included sex-specific analysis did not, and in the remaining two it was not clear whether stratification would have been appropriate (for example, an investigation of a nationwide outbreak of *Salmonella* infections associated with peanuts). She stressed that it is important to consider when sex-specific analysis makes sense and when it does not.

Other Subgroups: Race and Age

There is no question that sex is an important difference and one that has been underreported in the literature, Lichtman said, but differences are also associated with race and age, and perhaps reporting policies will need to be extended to those categories—although when sex, age, and race are considered, data presentation and interpretation can become complicated, and what the most useful comparisons are need to be considered.

Workshop participant Pinn pointed out that the law requires NIH to include women and minorities and their subpopulations in clinical research. Analysis by race can be challenging, and researchers are often confused about how to address subpopulations. Although ORWH focuses primarily on women, the NIH National Institute for Minority Health and Health Disparities (NIMHD) focuses on minorities and other health-disparity populations. Both ORWH and NIMHD report data by race and by sex.

Data on Sex-Specific Reporting

Pinn stressed that in looking at data on sex-specific reporting, it is important to know what studies the data are based on, for example, whether the data are only for clinical trials, or for clinical trials and observational studies, or whether the data are for studies funded by NIH or for all studies. She noted that NIH has been conducting analyses of clinical research and in looking at 12,000 protocols in FY 2010 found that 56% of the 23.3 million participants were women. When sex-specific studies of diseases that affect only women or only men were excluded from the analysis, 51.6% of the participants in NIH-funded extramural

research were women (NIH, 2011). Pinn pointed out that the overall percentage of female participants in NIH studies has varied "over the years from 66 down to 49 percent depending upon what large studies may be underway." That presents a somewhat different picture of women in clinical trials from that presented in the report of Jagsi and colleagues (2009),[5] which found that women made up less than 40% of participants in a set of cancer clinical trials.

Beyond Journals

Workshop participant Nancy Lee, deputy assistant secretary of health for women's health and director of the Office on Women's Health of the U.S. Department of Health and Human Services, said that issues of sex-specific reporting vary among disciplines, for example, in experimental studies, epidemiologic studies, and surveillance reporting. There has not been enough separation by sex in surveillance reporting, she said, and she stressed the importance of holding surveillance and epidemiologic publications to the same standards as experimental and clinical reports. Lichtman agreed that it is important to report data on men and women in surveillance reports and noted that the challenge is to draw comparisons between men and women, making sure that there is adequate power to ensure confidence in conclusions. Consideration of sex-based differences is important not only in peer-reviewed literature but in grant applications, Lichtman added.

THE EDITOR PERSPECTIVE: IMPLEMENTING JOURNAL EDITORIAL POLICIES

Editorial Policy-Making

Journal of the National Cancer Institute: *Trailblazing the Way on Sex-Specific Reporting*

Barnett Kramer, editor-in-chief of *JNCI*,[6] noted that *JNCI* was the first journal to include instructions for addressing the effects of sex as

[5]See discussion by Adler on page 2.
[6]*JNCI* is owned and published by Oxford University Press and is not affiliated with the National Cancer Institute or the federal government.

part of its manuscript-preparation policy. Specifically, the *JNCI* instructions for authors state that "where appropriate, clinical and epidemiologic studies should be analyzed to see if there is an effect of sex or any of the major ethnic groups. If there is no effect, it should be so stated in Results."

The inclusion of that statement, Kramer explained, was the result of a telephone call from ORWH director Pinn. By the end of the conversation, he said, he viewed the lack of such a statement as an oversight (not something that needed to be debated), and *JNCI* simply developed language and updated the instructions. Five in-house PhD-level senior editors at *JNCI* edit all manuscripts for content and ensure that the quality of the science meets journal standards, including appropriate sex-specific reporting.

Peer reviewers look at content and statistics and may also note whether there are sex-specific analyses, but the responsibility for ensuring rigor and determining what falls under "where appropriate" in the policy for sex-specific reporting falls to the in-house editors at *JNCI*. It is their interpretation of the instructions for authors that determines when sex-specific reporting is appropriate, Kramer clarified. It is not left to the authors to decide.

The *JNCI* policy emphasizes human studies, Kramer said, because it is sufficiently challenging to extrapolate data from in vitro and animal models to human applications. Authentication of cell lines used in in vitro studies is a particular concern. For example, the cross-contamination of cell cultures with HeLa cells is a substantial problem for laboratories, and researchers may not even know that they have female-derived HeLa cells in their cultures. *JNCI* tries to ensure that the lineage of cell lines is accurately identified in manuscripts; it asks authors to provide evidence that a line was obtained from a reliable source or that DNA "fingerprinting" has confirmed the lineage. Another issue is the predictive value of animal studies, and Kramer noted that animal models have not served the field of cancer research very well.

Another concern is analysis. As discussed by Wong, the more groups compared, the greater the likelihood of false-positive or false-negative findings and of spurious statistical interactions. If the number of events is small, the results may be less reliable, and authors or editors will often state this as a caveat.

There have been no challenges to implementing the *JNCI* policy, Kramer said. The editorial board agreed to the changes, and authors were made aware; there has been no pushback since the institution of the policy.

The International Committee of Medical Journal Editors: Many Editors, One Policy

Over the last decade or so, effective changes in editorial policies have generally been instituted journal by journal, Davidoff explained. Actions by associations of editors have been minimal. One exception is the International Committee of Medical Journal Editors (ICMJE), which started as the Vancouver Group, a small group of journal editors with no budget and no staff but a common interest. They began to consider some of the issues that editors face and together, as the ICMJE, developed and announced the *Uniform Requirements for Manuscripts Submitted to Biomedical Journals*.[7] For a manuscript of a clinical trial to be considered for publication in a journal that adheres to these requirements, for example, the authors must declare that all investigators had access to all the data. Another requirement is that the trial have been registered in a public registry. The ICMJE is an example of how the editorial community can function and have an impact as a community. Davidoff stressed that policy changes created by a representative group should not be mandatory. The *Uniform Requirements* are voluntary; journals choose to follow them and state in their instructions for authors that they do so. In addition, policies should be authoritative rather than authoritarian.

Jerome Kassirer, former editor-in-chief of the *New England Journal of Medicine*, agreed, noting that the clinical journals by and large adhere to the policies of the ICMJE, but they can and do diverge if they decide that their own policies are more appropriate for them. He added that although the ICMJE developed solid, evidence-based policies for how medical journals should operate, there is no comparable group of basic-science journal editors.

Kramer noted that sex-specific reporting is also part of the ICMJE *Uniform Requirements*, which state that "where scientifically appropriate, analyses of the data by such variables as age and sex should be included."

Sex and Gender Medicine vs Women's Health

Early on, the push to gather more information about differences between men and women through enrollment of women in clinical trials was thought of as a feminist issue, said Marianne Legato, founding editor

[7]See http://www.icmje.org/urm_main.html (accessed October 1, 2011).

of *Gender Medicine*. Some hospitals compounded that notion by using women's health as a marketing tactic. Decades later, there is still pushback, and Legato opined that the academic community is still not completely convinced that sex-specific medicine is relevant and necessary. One of the issues that remains for providers is whether sex-specific practical medicine makes a difference in outcome for a given patient sitting in the office. Another issue is whether it is economically feasible not only to delineate the differences between men and women but to apply them to a clinical population.

Objections to including women are still rampant, Legato said. Many investigators and practitioners believe that "gender medicine" is a politically motivated and scientifically indefensible discipline. Including both sexes increases the expense of studies and strains on already restricted research funding. The trajectory of a given disease (such as coronary arterial disease) is often quite different in men and women in timing, characteristics, and symptoms, and inclusion of both sexes is fraught with such issues as concomitant conditions in older patients, hormonal differences, and differences in the physiology of aging. Many researchers think that including premenopausal woman in clinical studies is potentially dangerous to their reproductive function and is potentially dangerous to any child conceived during the course of the trial, and women of childbearing age are often fearful and difficult to recruit.

Despite the resistance to studying women, there is growing acknowledgment that men and women have substantial and widespread biologic differences. There is more awareness of the extent and complexity of the sex-specific and gender-specific properties of living organisms.

In addition, there is a growing understanding of the human genome and increasing incorporation into research of ever-more effective ways of measuring the genome's effect on human biology. Some of the important questions to be addressed, given the complexity of how phenotype is determined (not just by genes but by environment and experience), are: How useful will delineation of a person's genome be in predicting disease and choosing therapy? Is it ever possible to separate what is hard-wired into the organism by virtue of biologic sex? What is the result of the effect of other factors on the phenotype? What are the effects of biologic sex on gene expression?

Finally, the emerging field of synthetic biology is transforming our understanding of what constitutes life and raising questions about what it means to be human. To what extent will the human phenotype be augmented or changed? Legato asked.

Legato explained that the first iteration of *Gender Medicine,* then called the *Journal of Gender-Specific Medicine*, was not successful and was discontinued. As the editor, Legato faced the challenge of explaining to researchers that gender-specific medicine is not women's health, but a genuine comparison between the sexes. The relaunched journal, *Gender Medicine*, is now in its 5th year of publication and is more successful in impact and content.

Preclinical vs Clinical Studies

Kramer said that it is not possible to have identical guidelines for preclinical animal studies and (human) clinical studies, but *JNCI* strives to maintain the same level of rigor in every field in which it publishes, namely, the entire spectrum of research that is fundable in the National Cancer Institute. About one-third of the papers report bench-oriented research in the laboratory, one-third epidemiologic research and observational studies, and one-third clinical studies of therapeutics, prevention, or screening. Every study must be reported in such a way that someone else could verify the results by repeating the study exactly as reported. Thus, one must know the sex of the animals used in a study, and reporting of sex should be an absolute requirement, Kramer said.

Jeffrey Blaustein, editor-in-chief of *Endocrinology* and a neuroendocrinologist, commented on the numerous aspects of animal studies that one could monitor: randomization; blinding; adverse environment, such as nearby construction; age of animals at shipment; interruptions in circadian cycles; and so on. It is not now possible to monitor submitted manuscripts for discussion of all variables in animal studies, and in many cases one has to assume that randomization or blinding, for example, has been done.

Kramer responded that the "rules of engagement" are more mature for clinical studies, and there are standard checklists for required elements, such as randomization. The *JNCI* in-house editors have checklists for both preclinical and clinical studies. In the clinical literature, when editors are rating the quality of reporting in a manuscript and, for example, how the randomization was done is not reported, the manuscript loses points on the checklist. There is no assumption that the authors did it if they do not report it. The same thing should be done for reports of animal studies, Kramer asserted. If blinding and randomization are not mentioned, he said, he suspects that the author did not know enough about the process and about where biases can be built in to report it.

Moving Toward Broader Sex-Specific Reporting

During his tenure at the *New England Journal of Medicine*, Kassirer said that many policy decisions were made and implemented, for example, the Ingelfinger Rule regarding duplicate publication and conflict-of-interest policies. Researchers are eager to have their papers published in high-profile journals, such as the *New England Journal of Medicine, JAMA,* and *Annals of Internal Medicine*. As a result, editorial policies implemented by those journals can be effective in modifying behavior. If the major clinical journals banded together to promote policies that foster sex-specific reporting, it seems reasonable to assume that investigators would take notice. Kassirer said that there is no bar to journals' introducing policies that require sex-specific reporting, as was done at *JNCI*.

Davidoff pointed out that until about 10–15 years ago, nearly all clinical-journal editors were men. In recent years, however, four of the five major general clinical journals have had female editors (today, two of the five have female editors). That is an important change, he said.

Legato suggested that editors put out a request for papers that directly address sex differences, perhaps for publication in a supplemental issue. Robert Golub, deputy editor of *JAMA,* said that his journal publishes theme issues on topics that are of immediate relevance, and sex differences could be considered as one of those topics. The goal of a theme issue is to highlight research or current thinking in a field. Floyd Bloom, former editor-in-chief of *Science*, supported the idea of editors' commissioning special issues on themes as a way to stimulate sex-specific science that could then be appropriately reported in journals; he noted that this has been successful in the neurosciences.

IMPLICATIONS FOR JOURNALS OF SEX-SPECIFIC REPORTING POLICIES OF JOURNALS

Nonclinical Science Journals

In basic-science research, it is very rare that male and female animals are studied side by side, said Levine. Even more problematic is the fact that male animals are often the default choice to avoid confounding effects of the estrous cycle or other female-specific physiologic circumstances. Levine suggested that basic-science journals consider a policy whereby authors are asked to justify the use of males vs females and

have reviewers check to see whether authors have considered the implications of their choice of males vs females in the design, conduct, and interpretation of their experiments.

Blaustein supported the idea of including the justification of the sex of animals as one of the items that reviewers and associate editors should look for, noting that it would help to increase awareness. He added that in many cases, the default choice is actually immature animals because researchers think that this avoids the effects of the estrous cycle and mix immature males and females.

Blaustein cited an article on sex bias in biomedicine by Zucker and Beery (2010, p. 690), who suggested that

> to correct the sex bias in animal research, we need stringent, strictly enforced measures, not voluntary appeals. Journal editors and reviewers should require authors of research studies that use only male or only female animals to state this in the title of their papers. This would highlight sex biases and spur researchers to balance the numbers of males and females that they use. Funding agencies should refuse to consider grant proposals that do not properly acknowledge the sex of the animals to used, and favor those that include males and females, and analyze data by sex.

Guidance on conducting sex-differences research is available, Blaustein noted. For example, an article that was the product of a series of discussions sponsored by the Society for Women's Health Research outlined strategies and methods for research on sex differences in behavior and the brain (Becker et al., 2005). The article was followed by a book, *Sex Differences in the Brain*, that discusses the subject in more depth (Becker et al., 2008). Guidance on designing and analyzing experiments to consider sex differences can be provided, Blaustein said, but that will not force researchers to look at both sexes.

Interdisciplinary Science Journals

Katrina Kelner, editor of the new journal *Science Translational Medicine* and former deputy editor for biology at *Science*, said that the issue of sex-specific reporting "is not on the radar screen" of the interdisciplinary journals (such as *Proceedings of the National Academy of Sciences*, *Science*, and *Nature*). Major issues that editors at *Science* have been working to address over the last 10 years include image manipulation, conflict of interest, and availability of the data and materials for the

analysis being reported, whether climate change, physics, biomedicine, or another discipline.

A challenge for any journal is to decide how stringently to enforce an editorial policy, Kelner said. High-profile journals must balance the large volume of manuscripts submitted against the limited time of reviewers and staff. A particular challenge for editors at an interdisciplinary journal is that one's expertise in any particular topic is limited. *Science* has around 20 editors, about half of whom are physical scientists and half of whom are molecular biologists. If *Science* adopted the sex-specific reporting recommendation of the 2010 IOM report *Women's Health Research: Progress, Pitfalls, and Promise*, the journal would have to rely, to some extent, on outside reviewers, Kelner said. There are limitations in relying on outside reviewers to ensure that appropriate analysis has been done. It can be difficult to find suitable reviewers—most are quite busy—and, because papers can be heterogeneous, editors do not always get the needed advice from a review.

Science Translational Medicine has adopted some of the ICMJE recommendations—including the institutional review board, authorship, and clinical-trial registration requirements—but has not adopted the recommendation about sex-specific reporting (although Kelner noted that she intends to revisit the issue after the workshop).

To aid editors of interdisciplinary journals, Kelner suggested that a role for the IOM and the editorial organizations could be to ensure that educational resources are available to editors, including information that they can refer to when assessing sex-specific reporting in papers.

Medical Journals

Study Design, Analysis, and Reporting

Golub shared the results of his informal audit of the 50 most recent randomized controlled trials published in *JAMA*. Of the 50, 21 reported single-sex results (on sex-specific topics) or presented results analyzed and stratified by sex. Of the 29 studies that did not report results by sex, 19 included at least 40% women, 22 at least 30% women, and 27 at least 20% women. Golub opined that only one of the 29 studies was possibly adequately powered to do subgroup analyses; he reiterated that statistical power is a recurring problem. Adler suggested that having data similar to what Golub presented published annually would be helpful.

The issue of reporting results by sex is related not simply to the percentage of female participants in studies but to fundamental questions of adequate power and of whether subgroup analysis is appropriate. The harm in separating males and females in the absence of sufficient statistical power is the risk of errors, such as a type II error in publishing comparisons by sex that show no significant difference. Negative results do not necessarily mean that there is no difference.

Women's Health Research states that "in the absence of a compelling reason not to, it should be assumed that there are sex differences in conditions" (IOM, 2010, p. 233). If the intent is to report raw data by sex for future research, that assumption is not likely to be an issue. For analysis of data by sex, however, Golub said that unless there is a priori evidence to support the assumption, the analysis may be flawed, and a Bayesian statistical approach may lead to false associations. In other words, a starting assumption that there is a sex difference for any association being studied could lead to the publication of false conclusions.

Gregory Curfman, executive editor the *New England Journal of Medicine*, concurred that the real challenge in reporting clinical-trial results separately for males and females is whether there is adequate statistical power for subgroup analysis. The rationale for reporting data by sex is indisputable, he said, but if a clinical trial has not been adequately powered to look at males and females separately, the conclusions are not going to be statistically sound. In many cases, achieving statistical significance for subgroup analyses would require unattainable or unjustifiable numbers of participants. Curfman therefore cautioned against editorial policies that require trials to be designed to reach valid statistical conclusions for males and females separately. Such editorial policies would create a "steep mountain to climb for investigators and for funding agencies," he said. In addition, as has been discussed, it may not be enough to report results only by sex. Other important demographics, such as race and age, add to the complexity of reporting results and of sample-size calculations for large clinical trials. Curfman also raised a concern that reporting results by sex in studies underpowered for valid subgroup analysis may be misleading and may be subject to misinterpretation by the health care community.

Science and *Science Translational Medicine* have a strong emphasis on data access, and Kelner endorsed the idea that even if a study has not been properly powered for male and female subgroup analysis, the resulting data should still be made available for others, perhaps in an appendix (with the appropriate warning that, because of statistical limitations, sex-specific data do not necessarily indicate a sex-specific differ-

ence). Ensuring that all data are available to everyone once a study report is published is an important role for journals, Kelner said. Berlin added that appropriate study design and use of common definitions would improve the ability to conduct meta-analyses of such archived data.

Current NIH policy on the participation of women in clinical trials has not achieved the desired effect of having enough women enrolled in all studies for sufficient statistical power, Golub said. And statistical power is an ever-increasing problem. For example, improvements in standards of care mean that control groups have better outcomes and that more participants must be enrolled to find a significant difference between control and test treatments. In some studies, composite outcomes (such as major adverse cardiac events) are already necessary for achieving adequate power, and this limits the possibility of valid subgroup analysis on individual outcomes. Composite outcomes can hide important effects, Golub noted, and often the events that are driving a difference are the less important events, such as rehospitalization, whereas the number of deaths may not be statistically different or may be underpowered for analysis. Given those challenges, is it feasible to have enough women enrolled to facilitate statistical analyses of the sexes separately? Will a Bayesian statistical trial design enable studies with a lower aggregate enrollment? Once the primary hypothesis and primary outcomes are established, the necessary enrollment is determined on the basis of power analysis and expected attrition. Adding more patients is expensive, Golub said, and researchers cannot simply over-enroll in anticipation of unplanned subgroup analyses.

Blaustein questioned the extent to which the number of participants would need to be increased to see an effect of sex. In animal studies, he noted, sex differences are often apparent even in small studies. Kassirer responded that as the difference in effect between two therapies becomes smaller and smaller, larger and larger studies are needed to be able to identify it. There are differences between animal studies and human studies and between different types of human studies, Golub added. For example, in clinical trials with major cardiac outcomes, severe events, including death, have become much more rare, and the effect size that the investigator is willing to accept may be fairly small compared with a more basic study in which one would be able to detect effect differences with 10 subjects in each group. Sufficient power for subgroup analysis may be easier to achieve in some types of studies and harder to achieve in large clinical trials with rare outcome events.

A participant raised a concern about the reliance on a statistical *p* value of 0.05 for accepting or rejecting a hypothesis. In some cases, a

finer *p* value, such as 0.000001, might be appropriate, and in others, a *p* value of 0.20 might be acceptable. It may be necessary sometimes to allow more leeway in interpretation (which, it was noted, is not the same as exploratory data analysis).

Christine Laine, editor-in-chief of *Annals of Internal Medicine,* said that publishing a large trial without sex-specific results does not necessarily mean that there *are* no sex-specific results; when discussing the limitations of a study, authors should point out that there could be sex differences. Absence of evidence is not evidence of absence, Davidoff agreed. He added that statisticians are often involved in clinical trials at a late stage and probably should be engaged much sooner.

One of the quandaries that an editor faces, Golub said, is the need to cater to two different readerships: researchers and clinicians. Researchers are interested in hypothesis findings to inform the design of their next study, but the mission of medical journals is to publish studies that will affect clinical care, and it is increasingly difficult for clinical readers to understand the articles. New methods, including Bayesian analysis and adaptive clinical trials, will allow studies to be done without the need to anticipate everything from the start. The downside is that editors will need to re-educate themselves so that they can understand the methods. More statisticians will be required as it becomes more difficult to find reviewers who can review the statistical design of a study. And again, there is the issue of whether readers will be able to understand the reports.

A dilemma in drug development, said Marietta Anthony, director of women's health programs at the Critical Path Institute, is that clinical studies are powered to show that a therapy is safe and effective in general, not necessarily safe and effective specifically in females or males. However, FDA guidelines mandate that medical products be demonstrated to be safe and effective in the populations that will use them, so there is some question of interpretation. Parekh explained that FDA guidance documents for preclinical animal studies recommend studying both male and female animals. Phase 1 clinical studies look for safety and pharmacokinetic differences among subpopulations of healthy volunteers, including women and men. If there are significant differences, they are monitored carefully in early phase 2 safety and efficacy studies. Early phase 2 studies in patients almost always include women. Phase 3 clinical trials are hypothesis-driven studies with prospectively defined end points and involve large numbers of patients. Sometimes the overall data will show no effect of a product but subgroup analysis reveals a significant effect in a particular population. Such a finding is hypothesis-

generating, Parekh said, and another prospectively designed trial is then conducted to confirm the finding.

Exploratory Subgroup Analyses Kramer said that in his experience, the issue with subgroup analyses in observational studies is not getting authors to do sex-specific analyses but preventing overinterpretation of results of exploratory analyses. Exploratory analyses—whether they are sex-specific or specific for comorbidity, age, race, or socioeconomic status—are different from prospective primary end-point analyses that are protected by randomization. *JNCI* does not publish post hoc analyses or exploratory analyses in article abstracts, and such analyses are clearly labeled in the results.

Golub concurred in the need to exercise caution on exploratory analyses. When the primary outcome of a study is negative, authors will often stress a positive secondary or post hoc finding. Identifying and addressing it is part of the peer-review process. Like *JNCI*, *JAMA* does not include exploratory analyses in abstracts and requires that they be clearly described as exploratory in the text. Golub noted that there is a difference between a prespecified secondary analysis and a post hoc exploratory analysis.

Sex Subgroups Sex hormones influence virtually all cells, and stage of reproductive life and development should be considered in designing and reporting studies, Blaustein said. For instance, children are not the same as adults, and estrous-cycling females are not the same as acyclic females. He acknowledged that considering such factors requires some knowledge of reproductive endocrinology or the involvement of a reproductive endocrinologist, which is unlikely in most cases. However, he pointed out that it is important to begin thinking about variations within the sexes.

Influence of Journal Editorial Policies

ICMJE and journal policies can influence researcher behavior. For example, Golub said, it is now rare to receive a report of a clinical trial that was not registered in a public database in accordance with the ICMJE policy, and all submissions are accompanied by author disclosure forms. It is important to note, however, that although a trial must be registered before the study begins enrollment, most policies (for example, requiring financial disclosure forms and race and ethnicity statements), unlike a sex-specific reporting policy, require little of the authors until

the study is complete. Even though the trial-registration policy changed researcher behavior, its implementation was straightforward and did not require any major changes in study design or any substantial investment of resources. Kramer added that journals can assert important influence on the quality of reporting. For example, journal editors have become more attuned to looking for discrepancies between a primary end point in the original study as designed and a primary end point that the authors discuss in their paper.

Golub questioned the extent to which an editorial policy could affect study design relative to sex-specific reporting, in light of the study-design issues involved (such as adequate power). Changing editorial policy might send a signal to researchers who are beginning to design a study, but would they be able to make the changes needed and design studies with adequate power? Larger studies would require additional funding, which may not be available. Changing standards and publishing analyses that are not likely to be valid is not a good solution either. It comes down to the design of the studies, and, inasmuch as study design has not changed substantially in the last 10 years, it is not clear what would make it change now. If editorial policies required sex-based analysis, would the funding follow?

As noted in *Women's Health Research* (IOM, 2010), not considering sex and gender differences in the design, analysis, and reporting of studies has limited understanding of important sex differences and slowed progress in women's health. Laine stressed that problems in the design and analysis of a study cannot be fixed simply by changing reporting requirements. Journal editors can ask authors to reanalyze their data but cannot ask them to redesign their studies and redo them. Journals could reject papers that do not report sex-specific results, but that is unlikely to happen.

Journals do not provide research funding, and Blaustein suggested that changes in how experiments are done start with funding agencies. Kelner concurred that funding agencies need to be partners in encouraging good research practices. There needs to be a culture shift within science. Questions about what can be accomplished by editors and publishers through setting standards for authors, whether these be recommendations or mandates, versus the role of federal agencies and other funders in shaping research culture to embrace consideration of sex differences as part of sound study design, were raised in a number of comments by participants.

As discussed earlier, the ICMJE policy is specific to medical journals. Laine listed several other editorial associations and published

guidelines that do not address sex-specific reporting at all, including the Council of Science Editors (which covers science broadly, not only biologic science but the physical sciences), the World Association of Medical Editors, the Guideline on Good Publication Practice of the pharmaceutical industry, and the EQUATOR (Enhancing the QUAlity and Transparency Of health Research) network, an umbrella organization that catalogs numerous reporting guidelines, such as CONSORT (Consolidated Standards of Reporting Trails) for randomized trials and STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) for observational studies. (The CONSORT and STROBE guidelines cover the majority of the clinical research published in the major journals.)

Annals of Internal Medicine does not have a specific policy on sex-specific reporting but follows the ICMJE policy, Laine said, and encourages authors to follow reporting guidelines, including CONSORT and STROBE. She added that for many years, *Annals* has indicated in the title and abstract when a study includes only men or only women and indicates in the limitations of the study if data are insufficient to examine potentially relevant sex differences or racial or ethnic differences. *Annals* does not ask authors to report sex-specific results when the study design is insufficient to enable useful reporting of such results.

A journal can put a policy into place, but there has to be a way to implement it. Laine offered clinical-trial registration as a case example: the ICMJE put its registration policy into place before there was anywhere for sponsors to register their trials. Similarly, journals could require that studies be powered for subgroup analysis, but that would entail the availability of resources to fund those types of studies. It will not work if the funders, researchers, and journal editors are not aligned. Editors can foster more accurate reporting, but must be careful about making requirements that are not feasible, Laine cautioned. As data-sharing advances, journals may be different a decade from now, and researchers whose studies do not meet some set criteria may move away from traditional journals and publish their results in an open-access setting. Berlin added that a motivation for trial registration was to eliminate publication bias—to make all results available regardless of whether they are positive or negative. He cautioned that a situation in which only studies with sex-specific results are published is not desirable.

Another issue is the long pipeline of current high-quality clinical trials, many with long-term followup. These will be coming to completion over the next decade or later, and panelists discussed how any editorial policy that affects study design would need to be phased in over an

extended period. In the interim, Golub said, it is unlikely that journals would forgo publishing a well-done, informative study that could affect patient care solely because it lacked enough power to permit valid sex-specific reporting. Conversely, journals are not willing to publish poor-quality studies or invalid or meaningless data or analyses.

As discussed earlier, the present workshop was designed to consider a recommendation in *Women's Health Research* (IOM, 2010) that "the International Committee of Medical Journal Editors and other editors of relevant journals should adopt a guideline that all papers reporting the outcomes of clinical trials report on men and women separately unless a trial is of a sex-specific condition" (IOM, 2010, p. 13). Laine raised several concerns about editorial policies that might be developed on the basis of that recommendation. First, it appears that observational studies are not included. Second, as discussed above, many trials include insufficient numbers of women or men to allow valid comparisons or within-group conclusions. Third, if randomization was not stratified by sex, the results should not be interpreted as causal relationships. Finally, simply reporting sex-specific results does not address the question of whether any of the observed sex differences are due to sex or to confounding factors.

Suggestions from the Editors

On the basis of their experiences in implementing editorial policies, the panelists offered a variety of suggestions regarding the inclusion of sex-specific information in scientific publications (summarized in Box 2).

It was also suggested that the ICMJE consider adopting a stronger sex-specific analysis and reporting statement similar to that of *JNCI*. Laine predicted, however, that ICMJE members would question why only sex was being addressed and not other key factors, such as age, race, ethnicity, and insurance coverage.

LOOKING FORWARD[8]

In the final session of the workshop, members of the workshop planning committee and others reflected on the recurring themes of the meeting. Clayton noted the intersecting roles of journal editors, government funding agencies, industry, basic researchers, and others in advancing the understanding of sex differences in health through careful study design, data-sharing, subgroup analysis, and sex-specific reporting of results.

Anthony pointed out that major goals of biomedical research are the translation of findings into clinical practice and informing the development of health policy. FDA's mandate, for example, is to ensure that the medical products that it regulates are safe and effective in the appropriate populations. In that regard, profound sex-based differences that affect health and disease in both males and females at the cellular, molecular, and physiologic levels should be considered. Adler pointed out that, despite progress over the last decade, there is still a need for more and better data on sex differences. The focus of the workshop was on how to bring information on sex differences in health to light and specifically how journal editors, through editorial policies, could influence how research is reported.

BOX 2
Summary of Suggestions by Individual Editors

Preclinical Studies

- The sex of the animals studied should be reported.
- If only one sex of an animal was studied, this should be indicated in the title of the article.
- In most cases, the sex of origin of cells used should be reported (excluding, for example, immortalized cell lines, which are highly transformed and for which the sex of the original cells may not be relevant).
- Both male and female animals should be studied when appropriate; and, when it is possible, both sexes should be studied in the same experiment.

[8]The topics highlighted in this section are based on closing remarks of members of the workshop planning committee and the session chairs and on the open discussions throughout the workshop. They should not be construed as reflecting any group consensus or endorsement by the IOM.

Clinical-Study Design[1]

- Studies should be designed with stratified randomization by sex; stratified analyses should not just be conducted post hoc. Simply mandating post hoc subgroup analyses today, on a study that was started 10 years ago, is not necessarily valid, because it will probably violate the randomization.
- Studies should be designed with adequate statistical power for subgroup analyses and to test for interactions.
- In the absence of adequate power, raw data should be archived by sex for future pooling and meta-analysis.
- One possible criterion for requiring the analysis and reporting of sex-specific results should include an a priori reasonable likelihood that sex-based associations might exist.

Clinical-Study Reporting

- The title and abstract should indicate whether a study involved only men or only women.
- If the study design allows identification of sex differences, journals should require authors to present these results.
- If there is an inability to identify sex differences, this should be reported in the discussion of the limitations of the study.

Researchers should be allowed to report inconclusive or descriptive sex-specific findings as raw data in electronic-only appendixes to meet NIH and FDA policies. As above, this will make the data available to researchers for conducting meta-analyses.

The Role of Editorial Policy

Levine reiterated earlier discussions that developing and implementing editorial policies regarding the analysis and interpretation of information that has already been obtained is fairly straightforward. Journal editors can set standards in their instructions for authors regarding what information is expected to be included in a manuscript.

More challenging and perhaps controversial is the development and implementation of editorial policies that ultimately influence how experiments are designed and conducted, including being appropriately powered to allow comparison between the sexes. Workshop participants expressed varied opinions regarding the extent to which a journal policy should stipulate what analyses authors must include. Some participants

also discussed how any clinical-journal editorial policy that affects study design may need to be phased in over a long period, inasmuch as clinical studies that are already under way may be years away from completion and publication.

Levine reminded participants that several editors stressed that the design and analysis of studies are not likely to change simply because editors change reporting requirements. Journals can encourage change by making very clear what they consider to be the standards for sex-specific reporting, but there needs to be a culture shift within science. In that regard, there is a key role for federal agencies and other funders in shaping research culture to embrace consideration of sex differences as part of sound study design. For example, when reviewing grant proposals, NIH and other funders could consider whether criteria for sex-specific analysis are met (for example, whether a study includes both males and females, is powered for valid subgroup analysis, or justifies the study of only one sex).

Adler pointed to earlier discussions that there are clearly other subgroups that may be relevant to consider, such as race and ethnicity. However, issues of sample size are more challenging with multiple groups and there is much greater evidence of the biological effects of sex than of race. The need for research that will allow for a better understanding of racial and ethnic differences in health and treatment effects does not diminish the need for sex-specific analyses, nor is it tied to it.

Levine discussed the issue, raised earlier by participants, of consideration of sex in studies that use animals, because these studies help to elucidate mechanisms and inform drug-development studies in humans. It was suggested that editorial policies for basic-science journals be "more of a carrot policy than a stick policy." That is, the value added in the review process by a stated policy of sex-specific reporting should be stressed, and reviewers should be advised to consider the inclusion of sex-specific information as a desirable attribute of a manuscript. The currency of scientific work is publication, and it was suggested that this approach will feed back to the design of experiments as researchers begin to understand that manuscripts that include sex-specific information and analyses, or a clear justification for studying only one sex, will be reviewed more favorably.

Levine also mentioned, as noted earlier by some participants, that there is no editorial body of basic-science journal editors comparable with the ICMJE for clinical journals.

Statistical Power for Subgroup Analysis

Anthony summarized earlier discussions pointing out that for clinical studies, issues of statistical power are paramount. But in a resource-limited research environment, larger samples do not constitute a feasible solution for enabling valid subgroup analysis. Instead, new study designs and advanced statistical methods may help reap the most reward from patient participation in clinical trials. It was suggested that when there is insufficient power to analyze sex differences within a study, it may be possible to combine data from various studies and conduct meta-analysis or apply advanced statistical methods, such as the use of Bayesian inference. Golub also cautioned about the potential for type II errors by publishing comparisons by sex that do not show a significant difference. If a study is not adequately powered to look for such differences, then a study showing no differences is meaningless.

Summary of Participants' Suggestions for Advancing Sex-Specific Reporting

Adler discussed four themes that she thought reflected the suggestions for advancing sex-specific reporting discussed by others over the course of the day.

First, Adler noted the earlier discussions related to identifying the sex of populations in journal publications, including listing the sex of origin of cells and tissues, the sex of animals in basic and preclinical research, and the sex of participants in observational research and in clinical trials. If only one sex is studied, noting that in the title of the paper would be helpful. Adler suggested that having summary data, similar to what Golub presented, published annually would be helpful.

Second, she reiterated the advantages of sharing sex-identified raw data, noting that if a study is not sufficiently powered for subgroup analysis, sex-identified raw data could be made available, either as a supplement or on a website, to facilitate meta-analyses (with the necessary caution to avoid overinterpretation of the raw data).

Third, referring back to the discussion of using a "carrot policy" by giving "extra credit" in review to manuscripts that include sex-specific information, Adler pointed out that editors could make it clear that including sex-specific information will enhance a paper's chances of publication.

Fourth, Adler mentioned earlier discussions about requiring sex-stratified analyses. She reminded participants that it was noted that requiring identification of sex-specific information is easier to implement than requiring sex-specific analyses, but possibly less effective than requiring them. Editorial policies that require sex-stratified analysis would affect how studies are designed and conducted, not only how they are reported. The ultimate goal would be a culture change in scientific research that embraces sex as a key variable for analysis.

CLOSING REMARKS

Clayton offered closing comments on behalf of the workshop sponsor. The purpose of research is to inform, she said, and, for those involved in health research, to increase knowledge about human biology and to foster development of evidence-based health policy and clinical care.

Journal editors, Clayton pointed out, are uniquely positioned as gatekeepers for much of the scientific knowledge that reaches the public domain. They have the power to advance appropriate consideration of sex differences, she said, acknowledging that the term *appropriate* is subject to interpretation and that "one size does not fit all." Journal editors and editorial bodies, such as the ICMJE, can set standards for the inclusion of sex-related information in manuscripts submitted to their publications, including the sex of origin of tissues and cells and the sex of animals and humans in preclinical and clinical studies. They are also in a position to set guidelines to encourage authors to think about analysis and reporting of sex differences.

Clayton reiterated that NIH requires the inclusion of women and minorities, as scientifically appropriate, in all clinical research that is supported by NIH. For a phase 3 clinical trial, if an evidence review reveals a likelihood of a sex-based difference, the study must be designed to allow comparisons between males and females, and the results must be provided to NIH in the final progress report. However, NIH does not have any control over what is published in the scientific literature. Together, the scientific community needs to find ways to ensure that this information gets out so that it can be helpful to researchers, clinicians, and policy-makers, she said. Funding is more limited than at other times, so scientists should also be efficient in collecting as many data as possible from studies. She concluded by noting that patients who participate in trials are relying on researchers to get the maximum information from clinical research.

REFERENCES

Bairey Merz, C. N., L. J. Shaw, S. E. Reis, V. Bittner, S. F. Kelsey, M. Olson, B. D. Johnson, C. J. Pepine, S. Mankad, B. L. Sharaf, W. J. Rogers, G. M. Pohost, A. Lerman, A. A. Quyyumi, and G. Sopko. 2006. Insights from the NHLBI-sponsored Women's Ischemia Syndrome Evaluation (WISE) study: Part II: Gender differences in presentation, diagnosis, and outcome with regard to gender-based pathophysiology of atherosclerosis and macrovascular and microvascular coronary disease. *J Am Coll Cardiol* 47(3 Suppl):S21-29.

Becker, J. B., A. P. Arnold, K. J. Berkley, J. D. Blaustein, L. A. Eckel, E. Hampson, J. P. Herman, S. Marts, W. Sadee, M. Steiner, J. Taylor, and E. Young. 2005. Strategies and methods for research on sex differences in brain and behavior. *Endocrinology* 146(4):1650-1673.

Becker, J. B., K. J. Berkley, N. Geary, E. Hampson, J. P. Herman, and E. Young, Eds. 2008. *Sex Differences in the Brain: From Genes to Behavior*. New York: Oxford University Press, Inc.

Beery, A. K., and I. Zucker. 2011. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev* 35(3):565-572.

Berry, D. A. 2005. Introduction to Bayesian methods III: Use and interpretation of Bayesian tools in design and analysis. *Clin Trials* 2(4):295-300; discussion 301-304, 364-378.

Canto, J. G., R. J. Goldberg, M. M. Hand, R. O. Bonow, G. Sopko, C. J. Pepine, and T. Long. 2007. Symptom presentation of women with acute coronary syndromes: Myth vs reality. *Arch Intern Med* 167(22):2405-2413.

Clayton, J. Presentation at Sex-Specific Reporting of Scientific Research: A Workshop. Institute of Medicine, Washington, DC.

FDA (U.S. Food and Drug Administration). 2010. *Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials*. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM071121.pdf (accessed November 14, 2011).

Freedman, L. S., R. Simon, M. A. Foulkes, L. Friedman, N. L. Geller, D. J. Gordon, and R. Mowery. 1995. Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993—The perspective of NIH clinical trialists. *Control Clin Trials* 16(5):277-285; discussion 286-289, 293-309.

Gandhi, M., N. Ameli, P. Bacchetti, G. B. Sharp, A. L. French, M. Young, S. J. Gange, K. Anastos, S. Holman, A. Levine, and R. M. Greenblatt. 2005. Eligibility criteria for HIV clinical trials and generalizability of results: the gap between published reports and study protocols. *AIDS* 19(16):1885-1896.

GAO (U.S. General Accounting Office). 1992. *Women's Health: FDA Needs to Ensure More Study of Gender Differences in Prescription Drug Testing*. GAO/HRD-93-17. http://archive.gao.gov/d35t11/147861.pdf (accessed November 2, 2011).

Gigerenzer, G. 2002. *Calculated Risks: How to Know When Numbers Deceive You*. New York: Simon & Schuster.

Goldkind, S. F., L. Sahin, and B. Gallauresi. 2010. Enrolling pregnant women in research—lessons from the H1N1 influenza pandemic. *N Engl J Med* 362(24):2241-2243.

Goodman, S. N. 1999. Toward evidence-based medical statistics. 2: The Bayes factor. *Ann Intern Med* 130(12):1005-1013.

Gurwitz, J. H., N. F. Col, and J. Avorn. 1992. The exclusion of the elderly and women from clinical trials in acute myocardial infarction. *JAMA* 268(11):1417-1422.

IOM (Institute of Medicine). 2010. *Women's Health Research: Progress, Pitfalls, and Promise.* Washington DC: The National Academies Press.

IOM. 2011. *Sex Differences and Implications for Translational Neuroscience Research: Workshop Summary*. Washington, DC: National Academies Press.

ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. 1988. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 2(8607):349-360.

Jagsi, R., A. R. Motomura, S. Amarnath, A. Jankovic, N. Sheets, and P. A. Ubel. 2009. Under-representation of women in high-impact published clinical cancer research. *Cancer* 115(14):3293-3301.

Kahn, S. E., B. Zinman, J. M. Lachin, S. M. Haffner, W. H. Herman, R. R. Holman, B. G. Kravitz, D. Yu, M. A. Heise, R. P. Aftring, and G. Viberti. 2008. Rosiglitazone-associated fractures in type 2 diabetes: An analysis from A Diabetes Outcome Progression Trial (ADOPT). *Diabetes Care* 31(5):845-851.

Kent, D., and R. Hayward. 2007a. When averages hide individual differences in clinical trials. *American Scientist* 95(1):60-68.

Kent, D. M., and R. A. Hayward. 2007b. Limitations of applying summary results of clinical trials to individual patients: The need for risk stratification. *JAMA* 298(10):1209-1212.

Kim, E. S., and V. Menon. 2009. Status of women in cardiovascular clinical trials. *Arterioscler Thromb Vasc Biol* 29(3):279-283.

Mittelstrass, K., J. S. Ried, Z. Yu, J. Krumsiek, C. Gieger, C. Prehn, W. Roemisch-Margl, A. Polonikov, A. Peters, F. J. Theis, T. Meitinger, F. Kronenberg, S. Weidinger, H. E. Wichmann, K. Suhre, R. Wang-Sattler, J. Adamski, and T. Illig. 2011. Discovery of sexual dimorphisms in metabolic and genetic biomarkers. *PLoS Genet* 7(8):e1002215.

NIH (National Institutes of Health). 2011. *Monitoring Adherence to the NIH Policy on the Inclusion of Women and Minorities as Subjects in Clinical Research: Comprehensive Report: Tracking of Human Subjects Research As Reported in Fiscal Year 2009 and Fiscal Year 2010*. Washington, DC: HHS and NIH.

Putting gender on the agenda. 2010. *Nature* 465(7299):665.

Sleight, P. 2000. Debate: Subgroup analyses in clinical trials: Fun to look at—but don't believe them! *Curr Control Trials Cardiovasc Med* 1(1):25-27.

Wald, C., and C. Wu. 2010. Biomedical research. Of mice and women: The bias in animal models. *Science* 327(5973):1571-1572.

Wong, J. B. 2011. *What are my chances?* Presentation at Sex-Specific Reporting of Scientific Research: A Workshop. Institute of Medicine, Washington, DC.

Zucker, I., and A. K. Beery. 2010. Males still dominate animal studies. *Nature* 465(7299):690.

A

Workshop Agenda

Tuesday, August 30, 2011

20 F St NW Conference Center
Washington, DC 20001

8:30–8:35 Introduction

Nancy E. Adler, Ph.D.; Lisa and John Pritzker Professor of Psychology, University of California, San Francisco; and Chair, Workshop Planning Committee

8:35–8:45 Opening Remarks—Office of Research on Women's Health

Janine A. Clayton, M.D.; Deputy Director, Office of Research on Women's Health, National Institutes of Health

8:45–9:30 Session 1—Why Sex-Specific Reporting Is Important

Session Moderator:
Nancy E. Adler, Ph.D.

Panelists:
Jesse Berlin, Sc.D., Vice President, Epidemiology, Johnson & Johnson Pharmaceutical Research and Development

Carolyn M. Clancy, M.D.; Director, Agency for Healthcare Research and Quality
Martha Nolan, J.D.; Vice President, Public Policy, Society for Women's Health Research
Ameeta Parekh, Ph.D.; Director, Research and Development, Office of Women's Health, U.S. Food and Drug Administration

9:30–11:00 Session 2—Sex-Specific Reporting Policies: Implications for Researchers

Session Moderator:
Jon E. Levine, Ph.D.; Director, Wisconsin National Primate Research Center; Professor, Department of Neuroscience, School of Medicine and Public Health, University of Wisconsin–Madison; Editor-in-Chief, *Frontiers in Neuroendocrinology*; Member, Advisory Committee, Office of Research on Women's Health; and Member, IOM Workshop Planning Committee

Panelists:
Larry Cahill, Ph.D.; Professor, Neurobiology and Behavior, University of California, Irvine
Denise L. Faustman, M.D., Ph.D.; Director, Immunobiology Laboratory, Massachusetts General Hospital; Associate Professor, Harvard Medical School
Judith H. Lichtman, M.P.H., M.Sc., Ph.D.; Associate Professor, Epidemiology and Public Health, Yale University School of Medicine
Rae Silver, Ph.D.; Helene L. and Mark N. Kaplan Professor of Natural and Physical Sciences, Barnard College and Columbia University
John B. Wong, M.D.; Professor of Medicine and Chief, Division of Clinical Decision Making, Informatics, and Telemedicine, Tufts Medical Center, Tufts University School of Medicine

10:45–11:00

Discussion with Audience

11:00–11:15 Break

11:15–12:30 Session 3—Experiences of Journal Editors Implementing Editorial Policies

11:15–12:00

Session Moderator:
Floyd Bloom, M.D.; Professor Emeritus Molecular and Integrative Neuroscience Department; The Scripps Research Institute; Former Editor-in-Chief, *Science*; and Member, Workshop Planning Committee

Panelists:
Frank Davidoff, M.D., MACP; Executive Editor, Institute for Healthcare Improvement; Editor Emeritus, *Annals of Internal Medicine*; Interim Chief Executive Officer, Physicians for Human Rights (via telephone)
Jerome P. Kassirer, M.D.; Distinguished Professor, Tufts University School of Medicine; Visiting Professor, Department of Medicine, Stanford University; Former Editor-in-Chief, *New England Journal of Medicine*; and Member, Workshop Planning Committee
Barnett S. Kramer, M.D., M.P.H.; Editor-in-Chief, *Journal of the National Cancer Institute*
Marianne J. Legato, M.D., FACP; Founder and Director, The Partnership for Gender-Specific Medicine, Columbia University; Professor of Clinical Medicine, Columbia University; Adjunct Professor of Medicine, Johns Hopkins University; and Founding Editor, *Gender Medicine*

12:00–12:30

Discussion with Audience (experiences of other journal editors)

12:30–1:30 Lunch

1:30–4:20 Session 4—Sex-Specific Reporting Policies: Implications for Journals

1:30–3:00

Session Moderator:
Marietta Anthony, Ph.D.; Director of Women's Health Programs, Critical Path Institute; and Member, Workshop Planning Committee

Panelists:
Jeffrey D. Blaustein, Ph.D.; Professor, University of Massachusetts Amherst; and Editor-in-Chief, *Endocrinology*
Gregory D. Curfman, M.D.; Executive Editor, *New England Journal of Medicine*
Robert M. Golub, M.D.; Deputy Editor, *JAMA*
Katrina L. Kelner, Ph.D.; Editor, *Translational Medicine*; Managing Editor, Research Journals *Science* Magazine (via telephone)
Christine Laine M.D., M.P.H., FACP; Clinical Associate Professor, Jefferson Medical College; Editor-in-Chief, *Annals of Internal Medicine*; and Senior Vice President, American College of Physicians

3:00–3:20 Break

3:20–4:20

Discussion with Audience

Discussion Moderators:
Marietta Anthony, Ph.D.
Jerome P. Kassirer, M.D.

4:20 – 5:00 Reflections and Looking Forward

Session Moderator:
Nancy E. Adler, Ph.D.

4:20–4:30

Janine A. Clayton, M.D.

4:30–5:00

Workshop Planning Committee Members:
Marietta Anthony, Ph.D.
Floyd Bloom, M.D.
Jerome P. Kassirer, M.D.
Jon E. Levine, Ph.D.

B

Speaker Biosketches

Jesse A. Berlin, Sc.D., is vice president of epidemiology at Johnson & Johnson Pharmaceutical Research and Development. Before moving to Johnson & Johnson, Dr. Berlin spent 15 years as a faculty member of the University of Pennsylvania in the Center for Clinical Epidemiology and Biostatistics. He is a fellow of the American Statistical Association. Dr. Berlin has served on the Institute of Medicine Committee to Review the Health Effects in Vietnam Veterans of Exposure to Herbicides and on the committee for the first biennial update. He has authored or coauthored over 230 publications in a wide variety of clinical and methodological areas, including papers on the study of meta-analytic methods as applied to randomized trials and epidemiology. He serves on the Scientific Advisory Committee to the Observational Medical Outcomes Partnership, a public–private partnership aimed at understanding methods for assessing drug safety in large, administrative databases. Dr. Berlin received his Sc.D. in Biostatistics from the Harvard School of Public Health.

Jeffery D. Blaustein, Ph.D., is editor-in-chief of *Endocrinology* and president of the Society for Behavioral Neuroendocrinology. A professor at the University of Massachusetts Amherst, Dr. Blaustein founded the Center for Neuroendocrine Studies and is director of the Neuroscience and Behavior Program. He was a charter member of the Society for Women's Health Research Interdisciplinary Studies in Sex-differences Network on Sex, Gender, Drugs and the Brain. He is on the editorial boards of *Journal of Neuroendocrinology* and *Frontiers in Neuroendocrinology*. Dr. Blaustein has been working in behavioral neuroendocrinology for almost four decades. His major research interests

are in the cellular processes underlying steroid hormone effects on behavior and reproductive physiology and the mechanisms by which stress and other environmental factors influence steroid hormone action in the brain. Although he has used female sexual behavior as a model for many years, he has more recently branched out into animal models of disorders of mental health. He has published over 140 articles, and his research has been funded by the National Institutes of Health and the National Science Foundation for over 30 years. Dr. Blaustein received his Ph.D. from the University of Massachusetts in 1977 and completed a postdoctoral program at Rutgers University from 1977 to 1979.

Larry Cahill, Ph.D., is a professor of neurobiology and behavior at the University of California, Irvine (UCI). He received his bachelor's degree from Northwestern University and his Ph.D. from UCI. He did postdoctoral work both at the Technical University in Darmstadt, Germany, and at UCI. He has investigated brain mechanisms of emotional memory in both animal and human subjects for over 30 years, the last 10 of which drew him into studies of sex influences on brain function, a topic that he now considers his most important field of research. He has twice been voted his school's Outstanding Professor by the students, a fact he considers to be the finest formal honor of his career.

Carolyn M. Clancy, M.D., was appointed director of the Agency for Healthcare Research and Quality (AHRQ) on February 5, 2003, and reappointed on October 9, 2009. Before her appointment, Dr. Clancy was director of AHRQ's Center for Outcomes and Effectiveness Research. Dr. Clancy, a general internist and health-services researcher, is a graduate of Boston College and the University of Massachusetts Medical School. After clinical training in internal medicine, Dr. Clancy was a Henry J. Kaiser Family Foundation Fellow at the University of Pennsylvania. Before joining AHRQ in 1990, she was an assistant professor in the Department of Internal Medicine of the Medical College of Virginia. Dr. Clancy holds an academic appointment at the George Washington University School of Medicine (clinical associate professor in the Department of Medicine) and serves as senior associate editor of *Health Services Research*. She serves on multiple editorial boards, including those of *Annals of Internal Medicine*, *Annals of Family Medicine*, the *American Journal of Medical Quality*, and *Medical Care Research and Review*. Dr. Clancy is a member of the Institute of Medicine and was elected a Master of the American College of

Physicians in 2004. In 2009, she was awarded the William B. Graham Prize for Health Services Research. Dr. Clancy's major research interests include improving health care quality and patient safety and reducing disparities in care associated with patients' race, ethnicity, gender, income, and education. As director of AHRQ, she launched the first annual report to Congress on health care disparities and health care quality.

Janine A. Clayton, M.D., is the deputy director of the Office of Research on Women's Health (ORWH) in the Office of the Director of the National Institutes of Health (NIH). Before joining ORWH, she was the deputy clinical director of the National Eye Institute (NEI). She is a board-certified ophthalmologist, and her research interests include immune-mediated diseases of the cornea and conjunctiva, women's eye health, and the standardization of outcome measures of diseases of the anterior segment. Dr. Clayton has been an attending physician and clinical investigator in cornea and uveitis at NEI since 1996, conducting research on inflammatory diseases of the anterior segment and providing medical and surgical uveitis fellowship training. Her clinical research has included randomized controlled trials of novel therapies for immune-mediated ocular diseases and studies of the development of digital imaging techniques for the anterior segment. Dr. Clayton has served on several committees at the NIH Clinical Center and currently serves on the Food and Drug Administration Advisory Panel for Ophthalmic Devices, the board of directors of Women in Ophthalmology, the executive committee of the Women's Eye Health.Org, the medical and scientific advisory board of Tissue Banks International, and the editorial boards of *The Ocular Surface* and *Oral Diseases*. Dr. Clayton received her undergraduate degree with honors from the Johns Hopkins University and her M.D. from Howard University College of Medicine. She completed a residency in ophthalmology at the Medical College of Virginia and fellowship training in cornea and external disease at the Wilmer Eye Institute of Johns Hopkins Hospital and in uveitis and ocular immunology at NEI.

Gregory D. Curfman, M.D., is the executive editor of the *New England Journal of Medicine (NEJM)*. In his leadership position at *NEJM*, he is responsible for setting journal editorial policies and for articles dealing with cardiovascular disease and health policy. Dr. Curfman has edited many articles on clinical trials. He developed and directs the Perspective section of *NEJM*, which addresses issues at the interface of medicine and

society. He has an interest in health law, has provided congressional testimony on health issues, and has written numerous editorials for *NEJM*. Dr. Curfman attended Princeton University and Harvard Medical School and is board-certified in internal medicine and cardiovascular medicine.

Frank Davidoff, M.D., MACP, is executive editor for the Boston-based Institute for Healthcare Improvement and a contributing writer for *JAMA*. He has also served on the faculty of Harvard Medical School and the University of Connecticut Medical School before becoming senior vice president for education of the American College of Physicians. He previously served as editor of *Annals of Internal Medicine* from 1995 to 2001. Dr. Davidoff has served on the Non-Prescription Drug Advisory Committee of the Food and Drug Administration, as vice-chair of the board of Physicians for Human Rights, as chair of the Journal Oversight Committee for *JAMA*, and as a member of the editorial boards of *Quality and Safety in Healthcare*, the *Journal of General Internal Medicine*, and the *Canadian Medical Association Journal.* Dr. Davidoff received his M.D. from Columbia University in 1959 and completed his residency training in internal medicine and endocrinology at the Massachusetts General Hospital in Boston. He has been the principal investigator of research grants from the National Institutes of Health, the National Fund for Medical Education, the Commonwealth Fund, the Pew Charitable Trust, and the American College of Physicians–American Society of Internal Medicine Foundation.

Denise L. Faustman, M.D., Ph.D., is an associate professor of medicine at Harvard Medical School and director of the Immunobiology Laboratories at the Massachusetts General Hospital. Dr. Faustman's research accomplishments include the first scientific description of modifying the antigens on donor tissues to change their foreignness, which is now being used in clinical trials. Currently Dr. Faustman works on strategies aimed at halting the established autoimmune disease process, such as that for type I diabetes, and the regeneration of the destroyed organs that form the basis of those autoimmune diseases. She is a member of the American Association for the Advancement of Science and has served as a member of many committees of the Institute of Medicine. Dr. Faustman earned her M.D. and Ph.D. from Washington University School of Medicine in 1982 and 1985, respectively. She completed her internship, residency, and fellowships in internal medicine

and endocrinology at the Massachusetts General Hospital. She started as an independent investigator at Harvard Medical School in 1987.

Robert M. Golub, M.D., is deputy editor of *JAMA* and associate professor of medicine at the Northwestern University Feinberg School of Medicine. His academic appointments are in the Division of General Internal Medicine and the Department of Preventive Medicine. Dr. Golub's research is in medical decision-making, including decision analysis, cost-effectiveness analysis, psychology of decision-making, and assessing patient preferences, and he has served on the board of trustees of the Society for Medical Decision Making. He served as chair of the Northwestern University Medical School curriculum committee and developed the curriculum on medical decision-making (which includes critical appraisal of the medical literature), serving as course director since 1992; for this curricular work, he received the Society of General Internal Medicine National Clinician–Educator Award for Teaching Innovation. Dr. Golub received his undergraduate degree from Princeton University and his M.D. from Columbia University College of Physicians and Surgeons. He completed his internship and residency at Northwestern University School of Medicine Northwestern Memorial Hospital.

Katrina L. Kelner, Ph.D. is the editor of *Science*'s new journal *Science Translational Medicine*. She started at *Science* as a manuscript editor for research papers in neuroscience over 20 years ago. Since then, she has held several other positions at *Science*: editor of Biology Perspectives, deputy editor for Commentary, and deputy editor for Life Sciences, overseeing the editorial staff who handle research papers in the life sciences. Dr. Kelner's interests revolve around application of basic-science advances and tools to clinical problems. She has spoken on numerous panels and at meetings on current advances in biology, the peer-review process, data-sharing, and conflict of interest in scholarly publishing. Dr. Kelner earned her undergraduate degree in biology at Reed College in 1975 and her Ph.D. in cell biology and neuroscience in 1981 at Baylor College of Medicine.

Barnett S. Kramer, M.D., M.P.H., is editor-in-chief of the *Journal of the National Cancer Institute*. He serves as chairman of the Physician Data Query (PDQ) Editorial Board on Screening and Prevention and is a member of the PDQ Treatment Editorial Board. Dr. Kramer has served on the Cancer Prevention Committee of the American Society of Clinical

Oncology and was the committee chair from 2006 to 2007. He has extensive experience in cancer-treatment studies, primary-prevention studies, and clinical screening trials of lung, ovarian, breast, and prostatic cancers. He is an investigator and on the steering committee for two large cancer screening trials sponsored by the National Cancer Institute: the Prostate, Lung, Colorectal, Ovarian Trial and the National Lung Screening Trial. He has a strong interest in weighing and reporting the strength of medical evidence and runs an annual Medicine in the Media Workshop to help working journalists to develop methods of reporting medical evidence. Dr. Kramer received his medical degree from the University of Maryland Medical School and completed his internship and residency in internal medicine at Barnes Hospital in St. Louis, MO. He completed a medical-oncology fellowship at the National Cancer Institute. He is board-certified in internal medicine and medical oncology and has received a master's degree in public health from Johns Hopkins University Bloomberg School of Public Health.

Christine Laine, M.D., M.P.H., FACP, is editor-in-chief of *Annals of Internal Medicine*. She is board-certified in internal medicine and remains active in patient care and teaching at Jefferson Medical College in the Division of Internal Medicine. Dr. Laine first joined *Annals of Internal Medicine* in 1995 as an associate editor and became a deputy editor in 1998 and senior deputy editor in 2008. In 2009, Dr. Laine became the editor and a senior vice president at the American College of Physicians. She is active in medical journalism and holds leadership positions in the International Committee of Medical Journal Editors, the Council of Science Editors, and the World Association of Medical Editors Ethics and Policy Committee. She has been instrumental in the development of editorial policy about such issues as authorship, conflicts of interest, and data-sharing in medical research. Dr. Laine graduated summa cum laude with a double major in biology and writing from Hamilton College in Clinton, NY. She received her medical degree from the State University of New York at Stony Brook and completed residency training in internal medicine at the New York Hospital (Cornell University) and a fellowship in general internal medicine and clinical epidemiology at Beth Israel Hospital (Harvard University). Dr. Laine earned her MPH with a concentration in quantitative methods and clinical epidemiology at Harvard University.

Marianne J. Legato, M.D., FACP, is the founder and editor-in-chief of *Gender Medicine*, professor of clinical medicine at Columbia University

College of Physicians and Surgeons, and adjunct professor of medicine at Johns Hopkins Medical School. Dr. Legato founded the Partnership for Gender-Specific Medicine at Columbia University in 1997 and is the editor of the first textbook on gender medicine, *Principles of Gender-Specific Medicine*. She spent her research career in cardiovascular research on the structure and function of the cardiac cell with the support of the American Heart Association and the National Institutes of Health. She has written numerous books for the lay public on women and cardiovascular disease and on gender-specific medicine. Dr. Legato received her medical degree from New York University School of Medicine in 1962.

Judith H. Lichtman, M.P.H., M.Sc., Ph.D., is associate professor of epidemiology and public health at the Yale University School of Medicine. Dr. Lichtman's research focuses on heart disease and stroke outcomes, using large administrative databases and observational studies, and on biologic, social, and environmental factors that influence the presentation and outcomes of young women with heart disease. Since her faculty appointment at Yale in 2001, Dr. Lichtman has been the prinicipal investigator in studies funded by the Centers for Disease Control and Prevention, the National Institute on Aging, the Agency for Healthcare Research and Quality, the National Heart, Lung, and Blood Institute (NHLBI), the National Institute of Neurological Disorders and Stroke (NINDS), and several private foundations. She is currently the principal investigator in two NINDS-funded projects that examine disease trends and outcomes for elderly stroke patients and co–principal investigator in an NHLBI-funded prospective observational study designed to examine the care and outcomes of young acute-myocardial-infarction patients (the VIRGO study). Dr. Lichtman has served on numerous national committees related to heart disease and stroke, including the American Heart Association (AHA) Patient Education System Task Force, the AHA Peer Review Evaluation Design Task Force, the AHA Stroke and Epidemiology Councils, and the AHA Quality of Care and Outcomes Research Expert Panel. She has been a member of the Program Committee for the AHA Conference on Cardiovascular Disease Epidemiology and Prevention and a cochair of two National AHA Writing Committees, and she is a member of the AHA Council on Epidemiology and Prevention Stroke Statistics Committee. She is also a member of the American Stroke Association Advisory Committee and the Quality Improvement Working group for

the AHA Get With the Guidelines Program. Dr. Lichtman received her Ph.D. in epidemiology from Yale University.

Martha Nolan, J.D., is vice president of public policy for the Society for Women's Health Research. She is responsible for the development and implementation of the society's government-relations and public-policy programs. She also provides advice and counsel to the society president on public-policy goals and strategies. Dr. Nolan joined the society staff in September 2003 and has over 17 years of experience in working for the health-insurance industry. Her most recent experience was as assistant vice president for federal affairs at MetLife. Before joining MetLife, she was counsel for state affairs for United Health Group, where she oversaw lobbying, coordinated advocacy, and managed state legislative and regulatory issues for over half the country. She has also worked for CIGNA and the Health Insurance Association of America. A lawyer by profession, Dr. Nolan earned her J.D. at Suffolk University Law School. She received a bachelor's degree in American history from Harvard University.

Ameeta Parekh, Ph.D., is the director of research and development in the Office of Women's Health (OWH) of the Food and Drug Administration (FDA). She leads the science program in OWH through collaborations and partnerships with other scientists in the FDA centers and with external leaders advancing women's health. As the OWH R&D director, Dr. Parekh represents FDA and OWH in national and international organizations and meetings to advance the scientific understanding of sex differences and provides FDA regulatory updates on the participation of women in clinical trials and the regulations, policies, and review practices around this topic. Dr. Parekh's background is in clinical pharmacology, pharmacokinetics, pharmacodynamics, and biopharmaceutics, and she has extensive regulatory experience at FDA, where she has worked for 25 years. Before joining OWH, she was a clinical-pharmacology lead in drug development with the Center for Drug Evaluation and Research in cardiovascular, gastrointestinal, reproductive, and urologic drugs. She is the FDA regulatory expert scientist for food effects and bioavailability of drugs. She has published extensively in peer-reviewed journals and book chapters and presented widely on women's health, drug development, clinical pharmacology, subgroup populations in drug development, and the drug-approval process.

Rae Silver, Ph.D., is Helene L. and Mark N. Kaplan Professor of Natural and Physical Sciences and holds joint appointments in arts and sciences at Barnard College, Columbia University, and the Department of Pathology and Cell Biology of the Columbia University Medical Center. Dr Silver's laboratory engages in two lines of research focusing on understanding neuroimmune system interactions and the brain clock. She created the undergraduate program in quantitative reasoning at Bernard College and, with colleagues, published studies of mathematical learning. She initiated the undergraduate major in neuroscience and served as its first program director. She also served as director of the graduate program in psychology at Columbia University. Dr Silver is a fellow of the American Academy of Arts and Sciences, American Association of Arts and Sciences. She has participated extensively in scientific and educational activities, including serving as cochair of the National Aeronautics and Space Administration's Research Maximization and Prioritization Committee, and as chair of the autonomic and limbic system theme of the Society for Neuroscience Program Committee. As senior adviser at the National Science Foundation, she worked with staff in all the scientific directorates to create a series of workshops to examine opportunities for the next decade in making advances in neuroscience through the joint efforts of biologists, chemists, educators, mathematicians, physicists, psychologists, and statisticians. She is a member of the Institute of Medicine Forum on Neuroscience and Nervous System Disorders. Dr. Silver received her Ph.D. in biopsychology from Rutgers University.

John B. Wong, M.D., is the chief of the Division of Clinical Decision Making, Informatics and Telemedicine in the Department of Medicine of Tufts Medical Center and the Clinical and Translational Science Institute of the Tufts University School of Medicine. He is a fellow of the American College of Physicians, a past president of the Society for Medical Decision Making, the statistical editor in decision and cost-effectiveness analysis for *Annals of Internal Medicine* at the American College of Physicians, and a consulting research member of the Tufts Evidence-Based Practice Center funded by the Agency for Healthcare Research and Quality (AHRQ). Besides serving on study sections for AHRQ and the National Institutes of Health, Dr. Wong has been a member of guideline committees for the American Association for the Study of Liver Disease Practice, the European League Against Rheumatism, OMERACT (Outcome Measures in Rheumatology), and the American College of Chest Physicians Antithrombotic Therapy. He

is the course director for evidence-based medicine at the Tufts University School of Medicine, the fellowship codirector for the National Library of Medicine–sponsored fellowship training program in medical informatics at Tufts Medical Center, and the medical informatics concentration leader for the Clinical Research Graduate Program of the Tufts University Sackler School of Biomedical Sciences. Dr. Wong's research focuses on the application of decision analysis to medical issues to help patients, physicians, and policy-makers to choose among alternative tests, treatments, and policies and thereby to promote rational evidence-based efficient and effective patient-centered care that reflects individualized risk assessment and patient preferences. Dr Wong received his M.D. from the University of Chicago and had postgraduate training in internal medicine at Tufts Medical Center, including a National Library of Medicine–sponsored medical informatics fellowship in clinical decision-making.

C

Planning Committee Biosketches

Nancy E. Adler, Ph.D. (*Chair*), is professor of medical psychology in the Department of Psychiatry and the Department of Pediatrics, vice chair of the Department of Psychiatry, and director of the Center for Health and Community at the University of California, San Francisco. She is a social psychologist by training. Her research interests include the effects of risk perception on reproductive and sexual health decision-making and identification of mechanisms by which socioeconomic status (SES) influences health. In the field of risk perception, she has studied how adolescents' perceptions of risk of sexually transmitted diseases and pregnancy influence sexual behavior and use of contraceptives. Dr. Adler's research on SES and health has focused on how social, psychologic, and biologic factors associated with SES act together to determine the onset and progression of disease and how the relationship of SES and health may depend on sex and ethnicity. She is the author of over 150 articles, books, and book chapters and is currently a member of the editorial boards of *Annals of Behavioral Medicine*, the *Journal of Health Psychology*, and the *Journal of Applied Social Psychology*. Dr. Adler was elected to the Institute of Medicine (IOM) in 1994. She served as a member of the IOM Committee on Prevention and Control of Sexually Transmitted Diseases (1995–1997) and chaired the Committee on Psychosocial Services to Cancer Patients/Families in a Community Setting (2006–2007). Dr. Adler received her Ph.D. in psychology from Harvard University.

Marietta Anthony, Ph.D., is director of women's health programs at the Critical Path Institute, which builds collaborative partnerships to support a new model of drug development. She is the associate director of the Arizona Center for Education and Research on Therapeutics (CERT), which focuses on drug safety. She was a senior health-policy analyst at the Agency for Healthcare Policy and Research (now the Agency for Healthcare Research and Quality). Dr. Anthony was the deputy director of the Office of Women's Health of the Food and Drug Administration and director of research programs in the Office of Research in Women's Health of the National Institutes of Health. She was in the Department of Pharmacology at Georgetown University and later vice president for health sciences in women's health the University of Arizona, where she founded and directed a National Center of Excellence in Women's Health. Dr. Anthony served on the Institute of Medicine panel on Women's Health Research: Progress, Pitfalls, and Promise.

Floyd Bloom, M.D., is the executive director of scientific communications and professor emeritus and former chairman of the Department of Neuropharmacology at the Scripps Research Institute. He is a past editor-in-chief of *Science* and served as president of the American Association for the Advancement of Science from 2002 to 2003 and as chairman of its board of directors from 2003 to 2004. Dr. Bloom is the recipient of numerous prizes for his contributions to science, including the Janssen Award in the Basic Sciences and the Pasarow Award in Neuropsychiatry, and is a member of the Royal Swedish Academy of Sciences. A member of the National Academy of Sciences since 1977 and a member of the Institute of Medicine since 1982, Dr. Bloom has participated on over 35 National Academies committees, including his current appointments on the Committee on Publications and the Report Review Committee.

Jerome P. Kassirer, M.D., served as the editor-in-chief of the *New England Journal of Medicine* from 1991 to 1999. He is currently Distinguished Professor of Medicine at Tufts University School of Medicine, where he has also served as vice chairman of the Department of Medicine. Dr. Kassirer's current interests are in clinical decision-making, teaching of clinical cognition, assessment of the quality of health care, professionalism, ethical scientific conduct, and financial conflict of interest. He has been highly critical of for-profit medicine, abuses of managed care, and political intrusion into medical decision-making. He has served on the American College of Physicians Board of

Governors and Board of Regents, chaired the National Library of Medicine's Board of Scientific Counselors, and chaired the American Board of Internal Medicine. He is a member of the Association of American Physicians, the American Academy of Arts and Sciences, and the Institute of Medicine. He is cochair of the Committee on Science for Judges—Development of the Third Edition of the *Reference Manual on Scientific Evidence*: Phase II.

Jon Levine, Ph.D., is the director of the Wisconsin National Primate Research Center and a professor in the Department of Neuroscience at the University of Wisconsin–Madison. He is also editor-in-chief of *Frontiers in Neuroendocrinology* and a member of the Steering Council for the Office of Research on Women's Health of the National Institutes of Health. Before going to Madison, Dr. Levine was a professor of neurobiology and physiology at Northwestern University, where his research focused on neuroendocrine regulation of gonadotropin-releasing hormone neurons and the cellular and molecular mechanisms of action of steroid hormones in the brain. Dr. Levine received his BA from Oberlin College and his Ph.D. from the University of Illinois at Urbana–Champaign. He completed postdoctoral training at the Oregon National Primate Research Center and the Oregon Health & Science University. Dr. Levine is an active member of numerous professional societies, including the Endocrine Society, the Society for Neuroscience, the Society for the Study of Reproduction, the American Neuroendocrine Society, and the Society for Behavioral Neuroendocrinology.

Harold C. Sox, M.D., MACP, is an editor emeritus of *Annals of Internal Medi*cine and has served on the editorial boards of three medical journals, including the *New England Journal of Medicine*. He is the principal author of *Medical Decision Making* and of the first and second editions of *Common Diagnostic Tests*, in addition to many other books, book chapters, editorials, and original research articles. Before becoming editor of *Annals of Internal Medicine*, he was the Joseph M. Huber Professor of Medicine and chair of the Department of Medicine of Dartmouth-Hitchcock Medical Center after spending 15 years on the faculty of Stanford University School of Medicine, where he served as chief of the Division of General Internal Medicine and as a director of ambulatory care at the Palo Alto Veterans' Administration Medical Center. He served as president of the American College of Physicians and chaired the U.S. Preventive Services Task Force and the Medicare Coverage Advisory Committee. A member of the Institute of Medicine

(IOM) since 1993, Dr. Sox served as chair of the IOM Committee to Study HIV Transmission Through Blood Products and the IOM Committee on Health Effects of Exposures in the Persian Gulf War. Most recently, he was a member of three IOM committees involved in the emergence of comparative-effectiveness research.