Volume 356, February 1, 2007, Number 5, pp. 437-536

**THIS WEEK IN THE JOURNAL**

**Article Summaries**

**PERSPECTIVE**

D. M. Nathan

Doctors and Soccer Players — African Professionals on the Move
F. Mullan

Pakistani Physicians and the Repatriation Equation
S. Shafqat and A. K.M. Zaidi

Focus on Research: JAK2 Mutations in Polycythemia Vera — Molecular Mechanisms and Clinical Applications
A. Tefferi

**CLINICAL PROBLEM-SOLVING**

Anchors Away
C. S. Calfee and Others

**EDITORIALS**

Cardiovascular Risks from Fine Particulate Air Pollution
D. W. Dockery and P. H. Stone

The Healing Power of Listening in the ICU
C. M. Lilly and B. J. Daly

Pay for Performance at the Tipping Point
A. M. Epstein

**ORIGINAL ARTICLES**

Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women
K. A. Miller and Others

**CLINICAL IMPLICATIONS OF BASIC RESEARCH**

Huntington's Disease — Making Connections
J. T. Greenamyre

JAK2 Exon 12 Mutations in Polycythemia Vera and Idiopathic Erythrocytosis
L. M. Scott and Others
A Communication Strategy and Brochure for Relatives of Patients Dying in the ICU
A. Lautrette and Others

Brief Report: Prepubertal Gynecomastia Linked to Lavender and Tea Tree Oils
D. V. Henley, N. Lipson, K. S. Korach, and C. A. Bloch

SPECIAL ARTICLES
Public Reporting and Pay for Performance in Hospital Quality Improvement
P. K. Lindenauer and Others

CORRESPONDENCE
Rituximab for Pemphigus Vulgaris
Effect of Ramipril on the Incidence of Diabetes
Prevention of Meningococcal Disease

Medical Mystery: Abnormal Abdominal Radiograph — The Answer
Case 32-2006: A Girl with Fever after a Visit to Africa
Intoxication of a Prison Inmate with an Ethyl Alcohol–Based Hand Sanitizer
Intoxication of a Hospitalized Patient with an Isopropanol-Based Hand Sanitizer

BOOK REVIEWS
Psychiatric Aspects of HIV/AIDS
The Clinical Neuropsychiatry of Stroke: Cognitive, Behavioral, and Emotional Disorders Following Vascular Brain Injury
McAlpine's Multiple Sclerosis

Textbook of Neural Repair and Rehabilitation, Volume I: Neural Repair and Plasticity; Volume II: Medical Neurorehabilitation

CORRECTIONS
Prevention of Meningococcal Disease
Release from Prison — A High Risk of Death for Former Inmates

CLINICAL THERAPEUTICS
Inhaled Insulin for Diabetes Mellitus
G. T. McMahon and R. A. Arky

Prevention of Meningococcal Disease

David M. Nathan, M.D.

Two modern-day epidemics, HIV–AIDS and type 2 diabetes mellitus, have inspired impassioned calls for more effective interventions. In the 1980s, the rapid spread of HIV, with its associated severe, acute illness and high mortality, prompted activist groups and others to call for the accelerated approval of medications that showed promise of efficacy. There was no treatment available, and people were dying quickly. More recently, pressure to develop new drugs for type 2 diabetes has been stimulated by the remarkable worldwide increase in the incidence of this disease (54% in the past 7 years in the United States1) and the recognition that intensive glycemic control is highly effective in reducing the development and progression of its long-term complications.

Of course, type 2 diabetes today differs from HIV–AIDS in the 1980s in at least two relevant ways: it is already treatable with fairly safe and effective agents and interventions, and it is a chronic illness, not a swiftly fatal one. Thus, patients who adhere to the current treatments for diabetes can live quite a long time, during which they may be subject to the adverse effects of any new medications they receive; this makes the risk–benefit equation dramatically different from that for the early AIDS drugs.

The pace of the development of new medications for diabetes may seem frustratingly slow to patients, but it has arguably never been faster. Nine classes of medications are now available for the treatment of type 2 diabetes, as compared with four barely a decade ago (see table). Although the Food and Drug Administration (FDA) has been criticized for being too slow in approving new medications, it has also been accused of being haphazard in its approvals. The history of troglitazone — the first thiazolidinedione to be approved, only to be withdrawn 2 years later because of severe hepatotoxicity — should serve as a cautionary tale. With more than 30 available medications that can be used either as monotherapy or in numerous combinations, it is fair to ask how many drugs are necessary to combat the epidemic of type 2 diabetes, and at what cost.

In theory, newer classes of antidiabetes medications might
be welcome additions to the existing armamentarium; however, those that have been developed recently are generally no more potent, and often less effective in lowering glycemia, than the three oldest classes (insulin, the sulfonylureas, and the biguanides), all of which are more than 50 years old (see table). Moreover, the newer classes are uniformly more expensive and are associated with adverse effects — some that are shared by the older drugs, but others that are new. Ironically, the two oral antidiabetes medicines that are most effective in lowering glycemia are also the oldest and were discovered accidentally, without the benefit of our contemporary understanding of their mechanisms of action or of the pathophysiology of type 2 diabetes. They were also discovered without high-throughput screening and other drug-development tools that we now have at our disposal.

With the approval of sitagliptin (Januvia, Merck) on October 17, 2006, and the pending approval of vildagliptin (Galvus, Novartis), we now have a ninth class of antidiabetes medications. The concept behind these dipeptidyl peptidase IV (DPP-IV) inhibitors, known as the gliptins, derives from the recognition that glucagon-like peptide (GLP) 1,37 — a naturally occurring gastrointestinal peptide that stimulates insulin secretion, suppresses glucagon levels, and slows gastric emptying — is rapidly inactivated by DPP-IV. Given that the main therapeutic effectiveness of the gliptins is mediated by their ability to increase levels of GLP, which is not a very effective glucose-lowering agent, it is not surprising that these agents are relatively ineffective in lowering glycated hemoglobin levels. Moreover, although they have been developed to be relatively specific for the GLP substrate and not to increase the levels of the many other DPP-IV substrates, none of them is so selective as to preclude alteration of the other substrates, including proteins involved in immunity and other hormones. The potential for unexpected consequences thus remains relatively high.

What is surprising is that despite these concerns and limitations, and despite the paucity of published data from long-term clinical trials on its efficacy and safety, sitagliptin was approved by the FDA. At the time of its ap-
approval, there was only one published, peer-reviewed, moderately large clinical trial; it included 392 treated patients who were followed for 18 weeks to judge the efficacy and safety of the drug. Results of three other clinical trials were published within 1 month after the approval. The 24-week trials included a monotherapy study involving 467 treated patients, a study of sitagliptin added to pioglitazone therapy in 175 patients, and a study in which sitagliptin was added to metformin therapy in 453 patients. All the studies were placebo-controlled, and in all of them, sitagliptin achieved a reduction in the glycated hemoglobin value of 0.5 to 0.9 percentage point as compared with placebo.

Such results indicate that sitagliptin is one of the less effective glycemia-lowering drugs introduced in recent years. Moreover, the drug has no obvious extraglycemic benefits. The GLP analogues, for example, are associated with some weight loss, whereas sitagliptin does not appear to affect weight. Although sitagliptin seems relatively safe, causing no increase in severe adverse events, the published data reflect testing in only a limited number of patients for a limited period (641 patient-years in total). Since nearly 20 million people in the United States have type 2 diabetes, there is potential for extensive use of new diabetes medications. For example, within 1 year after the approval of troglitazone in 1997, an estimated 600,000 U.S. patients were receiving it. Although severe hepatotoxicity had not been identified in preliminary testing in about 5000 people, cases of severe idiosyncratic liver disease began to appear within 6 months after approval, ultimately reaching a prevalence of approximately 1 in 15,000 patients who were receiving the drug. In the first 6 weeks after the approval of sitagliptin, business reports suggested that it accounted for 14% of new prescriptions for antidiabetes medications.

What is surprising is that despite the paucity of published data from long-term clinical trials, sitagliptin was approved by the FDA.

The criteria that the FDA uses in approving antidiabetes medications are based primarily on safety and on effectiveness in lowering glycated hemoglobin levels. Considering the potential for unanticipated adverse events, especially with medications that have as many biochemical effects as the DPP-IV inhibitors, and the plethora of antidiabetes medications that are currently available, many of which are more effective than sitagliptin, one wonders: why the rush to approve the gliptins?

The ability of clinicians to judge the merits of new medications is already limited — most receive their information about them from drug companies’ representatives and promotional materials. The dearth of peer-reviewed, published studies on sitagliptin makes it difficult for physicians to weigh the benefits and risks of the medication or to describe them to their patients.

No one wants to slow the development of effective new drugs, especially for a disease with a prevalence that is reaching epidemic proportions. However, the FDA’s approval process for new antidiabetes medications should take into account their additional and unique contributions, especially when their glucose-lowering efficacy is similar to or less than that of currently available medications. A host of medications that are already available are effective as monotherapy or in combination with metformin or one of the thiazolidinediones (the approved uses of sitagliptin). The fact that these medications achieve better glycemic control than sitagliptin suggests the need for caution in approving a new medication that has received limited testing.

The failure of clinicians and their patients with diabetes to implement currently available interventions aggressively and effectively is, I suspect, the major barrier to good care. This problem will not be fixed by making more medications available. Ensuring the effective and cost-effective use of the medications that have already been established by high-quality clinical trials to control glycemia or prevent diabetes should be a higher priority than flooding the market with ever more medications.

Dr. Nathan reports receiving lecture fees from Pfizer, GlaxoSmithKline, and Novartis and grant support from Novo Nordisk and Sanofi-Aventis.
On June 22, 2006, the nation of Ghana erupted. SUVs flew through the streets of Accra with flag-waving celebrants jammed through sunroofs. Crowds led by shirtless drummers banging garbage-can tops snaked down major roads, picking up revelers as they went. Hundreds of thousands of people took to the streets, shouting jubilantly. Ghana, playing in its first World Cup, had beaten the United States and earned a berth in the final stage of the global soccer pageant. It was a paroxysm of national pride that Ghana had rarely experienced.

“It’s the same for football players as it is for doctors,” I was told by Tsiri Agbenyega, dean of the medical school in Kumasi, Ghana. “We have to train a lot more than will end up in Ghana, because they all leave. The football players go to Europe, and the doctors to America and the U.K.” Agbenyega spoke with a mixture of frustration, pride, and resignation. He was pleased that Ghanaian athletes and physicians were competitive internationally, but their success meant a loss to the country — a loss more problematic in medicine than in football.

The World Bank considers Ghana a low-income country, but its 20 million people enjoy...
natural resources (gold, timber, and cocoa) and a relatively stable recent political history. Ghana has a strong tradition of education, a public health system that has resulted in greater longevity and lower infant mortality than in much of West Africa, and a prevalence of HIV infection among adults of 2.3% — lower than the sub-Saharan African average of 6.1% and far lower than southern African levels exceeding 20%.

So the country would seem to be in a good position to build and sustain a health care workforce that could rapidly reduce loss of life among infants and parturient women in Ghana (both mortality rates are more than 10 times those in high-income countries) and initiate widespread antiretroviral treatment to stem its AIDS epidemic. If Ghana could show the way, one might think, other African countries might be able to follow.

But not so. For much of the past decade, health improvement in Ghana has been at a standstill, and health statistics in many sub-Saharan African countries are sliding backward. AIDS is a culprit, but so is the exodus of doctors and nurses who are lured by U.S. training and employment opportunities. According to the Ministry of Health, Ghana has about 13 physicians per 100,000 population (as compared with 256 in the United States) and about 92 nurses per 100,000 (as compared with 937 in the United States). Today, there are 532 Ghanaian doctors practicing in the United States. Although they represent a tiny fraction of the 800,000 U.S. physicians, their number is equivalent to 20% of Ghana’s medical capacity, for there are only 2600 physicians in Ghana. An additional 259 Ghanaian physicians are in practice in the United Kingdom and Canada — and this group includes only those who have successfully been licensed after leaving Ghana. In other countries, the situation is even worse:

For years, we have been educating about three quarters of the doctors we need and relying on the rest of the world to supply the balance.

60% of Liberia’s physicians are in practice in the United States or Britain.

“Our only recourse is to try to train more in the hopes we will keep more,” explained Yaw Boasiako of Ghana’s Ministry of Health, who outlined an ambitious plan for doubling the number of physicians and nurses educated in the next few years. Ghana, like many English-speaking developing countries, is caught in an educational conundrum: the better the quality of their universities and the more health professionals they train, the more they lose to the United States and the United Kingdom. They have a leaky bucket now. In desperation, they’re building a bigger leaky bucket.

But that’s not all they’re doing. As in most developing countries, the private medical sector is small, and most physicians work for the government health service, which staffs the public hospitals and clinics where most people receive care. Although the salaries of Ghanaian doctors are better than those in many African countries, doctors are quick to point out that their pay is still modest. “A trained physician can make more in London in two months than we can make in a year in Ghana,” I was told frequently. Struggling with a limited budget and against the lure of Western incomes, the government has embarked on some creative strategies to retain physicians. These include pay increases, cheap car loans for doctors in “hardship posts,” and a plan to subsidize staff housing in rural areas. To address the desire of medical graduates to obtain specialty training, the government has launched an expanded program of in-country medical residencies.

To augment physicians’ services, the ministries of health and education are expanding training opportunities for community health nurses, technical officers, and “medical assistants” — mid-level practitioners who substitute for doctors in shortage areas. For many years, the Rural Health Training School in Kintampo has provided experienced nurses with a year of advanced training and 6 months of internship to enable them to function independently as medical assistants. The school is doubling its class size to 200 but is changing to a non-nurse model, since the loss of nurses to emigration has depleted the ranks of program candidates. In the future, medical assistants will be secondary-school graduates who will receive 3 years of didactic
PERSPECTIVE

Doctors and Soccer Players — African Professionals on the Move

Pakistanis Physicians and the Repatriation Equation

Saad Shafqat, M.B., B.S., Ph.D., and Anita K.M. Zaidi, M.B., B.S.

In Pakistan, students who are accepted into medical school are congratulated — only half-jokingly — on three counts: that they will become doctors, that they will become certified by the American Board of Medical Specialties, and that they will soon be living in the United States.

Pakistan has contributed approximately 10,000 international medical graduates (IMGs) to the United States, even though it faces a shortage of physicians.

Take the case of Aga Khan University Medical College in Karachi. By 2004, it had produced 1100 graduates, 900 of whom had gone on to graduate medical training in the United States — despite the fact that doing so costs up to $20,000 (a fortune for most Pakistanis) and means leaving the comforts of one’s home and culture.

The United States represents an overpowering lure: a rigorous system of graduate medical education, a merit-based structure of professional rewards, and a culture of academic nurturing. And, of course, material rewards.

In Pakistan, an intern earns approximately $150 per month (the same salary as an unskilled, illiterate worker), whereas a U.S. intern can afford to live independently — and expect a better quality of life after residency.

Information from Pakistani medical institutions indicates that about 300 of the 10,000 U.S.-trained Pakistani physicians have resettled back home. Why did this minority choose to return? Aga Khan’s experience is instructive: the majority of the medical school’s 40 or so alumni who have repatriated from the United States have joined its faculty.

Motives for returning include aging parents and family ties, a desire to raise children in a familiar culture, and an emotional need to be home. But for many Aga Khan returnees, the attributes of the university and its hospital were key: teaching, research, and clinical care are patterned after the U.S. model, and salaries permit a comfortable lifestyle. Ultimately, attractive career prospects have to be the draw.

The challenge is local capacity to absorb highly trained physicians. U.S.-trained physicians represent a small fraction of Pakistan’s 116,000 doctors, but they return with ambitions to set new standards for clinical practice, education, and research and to influence academic medicine, health policy, and public health. To do so, they must negotiate local circumstances for which they are unprepared: exhausting clinical demands, an impoverished population, an environment in which malnutrition is a significant cause of death, collapsed health care delivery systems, and patients who respond to an uncritically authoritarian style.

Discussions with expatriate physicians indicate that many more wish to return but cannot find suitable jobs. Like many poor countries, Pakistan has...
both severe shortages of health care professionals and a high level of unemployment among physicians — a paradox caused by inadequate and inappropriate investment in local health care systems. Elite medical academies in developing countries are frequently derided as manufacturers of a product that, out of place in its environment, enters a workforce supply chain leading to the West.\textsuperscript{1,2}\textsuperscript{1,2} The answer, however, is not to lament the irrelevance of these institutions but to advocate for more — for they can attract back highly trained professionals who have the potential to assume leadership roles. Repatriated Aga Khan graduates have won grants from major international agencies, established nonprofit research organizations, joined hospitals serving refugee populations, and led disease-control programs. Such academic institutions can play pioneering roles if they reorient their priorities to match their countries’ needs — producing professionals with a strong public health ethic, establishing rigorous graduate programs in which trainees are paid good wages, and developing relationships with alumni that can help sustain rewarding careers in challenging environments.

Exhorting physicians to serve in environments to which their skills are ill-suited will not lure IMGs home. Barriers to immigration in individual countries are almost meaningless in a globalized world. For example, as immigration laws in Western countries are tightened, Pakistani physicians are seeking jobs in the Middle East. We believe that developed countries that import physicians to meet their own demands have a moral obligation to invest in improving health care systems in countries that train substantial segments of their workforce. Such investments provide employment opportunities for the diaspora of health care professionals, benefiting health in developing countries.

As a first step, the U.S. medical community can support IMGs who want to repatriate. U.S. academic medical centers could work with institutions in developing countries to develop training programs oriented toward global health,\textsuperscript{4} availing themselves of growing funding opportunities for such endeavors.\textsuperscript{5}

One approach is to offer motivated IMGs mentoring to equip them with skills needed in their home countries. The scheme could be formalized through international cross-appointments for mentor and mentee at each other’s institutions and a bilaterally recognized role for the mentor. Such initiatives are desperately needed; properly done, repatriation of IMGs can help diminish vast disparities in health care.

Dr. Shafqat is an associate professor of neurology, and Dr. Zaidi an associate professor of pediatrics and microbiology, at Aga Khan University Medical College, Karachi, Pakistan.

JAK2 Mutations in Polycythemia Vera — Molecular Mechanisms and Clinical Applications

Ayalew Tefferi, M.D.

Janus kinase 2 (JAK2) is a cytoplasmic tyrosine kinase that transduces signals, especially those triggered by hematopoietic growth factors such as erythropoietin, in normal and neoplastic cells. In March and April 2005, four groups of investigators reported finding an acquired JAK2 mutation (termed JAK2 V617F) in association with polycythemia vera and related myeloproliferative disorders. These seminal reports have already been cited many times, and JAK2 is now a target for the development of new treatments for the myeloproliferative disorders. JAK2 V617F is detectable in more than 95% of patients who have polycythemia vera, as defined according to conventional criteria. The report by Scott et al. in this issue of the Journal (pages 459–468) indicates that the remaining 5% of patients have other JAK2 mutations with functional effects similar to those of V617F.

Normally, JAK2 activation involves tightly regulated cytokine-induced phosphorylation of tyrosine residues in the kinase and negative regulatory JAK homology domains — JH1 and JH2, respectively. JAK2 V617F, resulting from a point mutation (1849G→T) in exon 14, causes the substitution of phenylalanine for valine at codon 617 in the JH2 domain. This change in a single amino acid renders the JAK2 enzyme constitutively active. The four JAK2 mutations described by Scott et al. are in-frame deletions or tandem point mutations in exon 12. Both JAK2 V617F and the exon 12 JAK2 mutations induce cytokine-independent proliferation of cell lines that express erythropoietin receptors and cause these cells to become hypersensitive to cytokines. Moreover, the transfection of hematopoietic stem cells with a mutant JAK2 gene causes a polycythemia vera–like phenotype in mice.

These observations implicate JAK2 mutations as essential elements in the development of polycythemia vera. Whereas the exon 12 JAK2 mutations seem specific for polycythemia vera and idiopathic thrombocytosis, the exon 14 JAK2 V617F mutation occurs not only in polycythemia vera, but also in essential thrombocytopenia and primary myelofibrosis. Often, only one allele carries the JAK2 mutation, but the loss of heterozygosity through mitotic recombination converts the affected cell from a heterozygous to a homozygous carrier of the mutation. This event occurs most frequently in polycythemia vera. Homozygosity endows the cell with a proliferative advantage, and the increase in the allele burden probably contributes to full expression of the polycythemia vera phenotype.

No similar mechanism of allelic reinforcement was seen by Scott et al. with the new exon 12 JAK2 mutations, which are heterozygous but associated with stronger abnormal JAK2 activation than is V617F. V617F-negative clones with the potential for abnormal growth have been found in V617F-positive polycythemia vera, which suggests that the acquisition of the V617F mutation is not the primary event in the origin of the disease. This finding, however, does not undermine the role of deregulated JAK–signal transducer and activator of transcription (STAT) signaling as the principal contributing lesion in myeloproliferative disorders. Instead, it suggests subtle interactions between JAK2 V617F and potentially competitive or collaborative mutations, such as those affecting the thrombopoietin receptor (MPL W515L/K, an activating mutation in myelofibrosis).

An opportunity to change current diagnostic approaches to polycythemia vera arises from the fact that virtually all patients with polycythemia vera carry JAK2 mutations, whereas patients with secondary polycythemia do not. In principle, it is now possible to initiate the diagnostic workup of a patient with suspected polycythemia vera by testing peripheral-blood cells for JAK2 V617F (see flow chart). A concomitant measurement of the serum erythropoietin level, which is abnormally low in more than 90% of patients with polycythemia vera, would reinforce the molecular result, as well as capture the uncommon patient with V617F-negative disease. Such patients may have other JAK2 mutations, and most will have a prominent feature of polycythemia vera — cells in peripheral blood that form erythroid colonies in vitro in the absence of erythropoietin. If an exon 12 mutation is suspected but cannot be found in peripheral-blood cells, it should be sought in bone marrow cells, since if relatively few cells in the peripheral blood carry the mutation, it can escape detection. Patients with an exon 12 mutation can present with only an increased
red-cell mass (idiopathic erythrocytosis), without meeting the clinical criteria for polycythemia vera or displaying characteristic morphologic changes in bone-marrow megakaryocytes.

The presence of the JAK2 V617F mutation has not been convincingly demonstrated in reactive myeloproliferation, solid tumors, or lymphoid disorders, but the mutation can be detected in some healthy people by very sensitive polymerase-chain-reaction (PCR) methods. The presence of the mutation cannot be used to distinguish one myeloproliferative disorder from another. For these reasons, it is reasonable to test for JAK2 V617F when a clonal myeloproliferative disorder is suspected, regardless of the presence or absence of an increased hematocrit.

The reliable molecular diagnosis of polycythemia vera and related disorders will require the development of reproducible laboratory testing procedures. Since the current JAK2 mutation-screening tests are not standardized and quality control is not guaranteed, molecular testing should not substitute for sound clinical judgment. It is important to recognize not only that highly sensitive assays can detect low levels of JAK2 V617F in healthy people, but also that inadequately sensitive assays lead to false negative results in diseases with a low allele burden. In general, quantitative PCR assays are preferred, because they enable accurate assessment of allele burden as well as molecular monitoring of treatment response and residual disease.

The recent discoveries of exon 14 JAK2 V617F and the exon 12 JAK2 mutations are paving the way toward a molecular classification of the myeloproliferative diseases. The World Health Organization’s committee on hematologic cancers is planning to revise its criteria in the coming year, and the new document is expected to include information on JAK2 mutations. The prognostic relevance of these mutations and the allele burden are under investigation, with emphasis on the risk of thrombosis and on overall survival. As for new treatments, drug targets of interest include not only JAK2, but also other JAK–STAT signaling molecules. While the story is unfolding and being translated into clinical practice, it is prudent to refrain from assigning “distinct” phenotypes to newly discovered mutations and from drawing definitive conclusions regarding their clinical relevance.

Dr. Tefferi is a professor in the Department of Internal Medicine, Division of Hematology, Mayo Clinic College of Medicine, Rochester, MN.

This Week in the Journal

Original Article
Air Pollution and Cardiovascular Events
In this prospective study of more than 65,000 women, fine particulate air pollution was found to be associated with an increased risk of cardiovascular events and death from cardiovascular causes. These observations add to the growing evidence that air pollution, especially fine particulate matter, has important adverse health consequences.

See p. 447; Editorial, p. 511; CME, p. 539

Original Article
New JAK2 Mutations and Erythrocytosis
A V617F mutation in the Janus kinase 2 gene (JAK2) occurs in most patients with polycythemia vera and in many with essential thrombocytopenia or idiopathic myelofibrosis. The mutation causes unregulated signaling by the cytoplasmic tyrosine kinase of the JAK2 protein. The authors describe new mutations in exon 12 of JAK2 in patients with polycythemia vera or idiopathic erythrocytosis who do not have the V617F mutation. The exon 12 mutations, which also induce unregulated signaling by JAK2, appear to define a distinctive myeloproliferative syndrome.

See p. 459; Perspective, p. 444

Original Article
A Communication Strategy and Brochure for Relatives of Patients Dying in the ICU
The death of a loved one in an intensive care unit is an emotionally trying experience. These investigators compared a proactive end-of-life conference with family members, including the provision of an informational brochure, with a customary conference; outcomes were reported by family members 90 days after the loved one’s death. Family members who participated in the intervention conference had improved outcomes, as compared with those who participated in the standard conference.

See p. 469; Editorial, p. 513; CME, p. 538

Brief Report
Prepubertal Gynecomastia and Lavender and Tea Tree Oils
Three otherwise healthy prepubertal boys with normal endogenous steroid levels had gynecomastia coincident with the topical application of products containing lavender and tea tree oils. Gynecomastia resolved after the use of these products was stopped. Studies in human cell lines indicated that both oils exhibited estrogenic and antiandrogenic activities, suggesting that repeated topical exposure probably caused prepubertal gynecomastia.

See p. 479

Special Article
Public Reporting and Pay for Performance in Hospital Quality Improvement
This study compared hospitals engaged in public reporting alone with hospitals engaged in both public-reporting and pay-for-performance programs. Performance on quality measures improved in both groups, but improvements were modestly larger for hospitals participating in pay for performance.

See p. 486; Editorial, p. 515

Clinical Therapeutics
Inhaled Insulin for Diabetes Mellitus
A 52-year-old man with type 2 diabetes mellitus requires insulin therapy. The possible role of inhaled insulin is considered. Inhaled insulin is a short-acting insulin that has been shown to have an efficacy similar to that of subcutaneous insulin in clinical trials. However, severe hypoglycemia has occurred more frequently with inhaled insulin than with subcutaneous insulin in some trials, and the long-term safety of this form of therapy is unknown. Inhaled insulin is not recommended for smokers or for patients with underlying lung disease.

See p. 497; CME, p. 537

Clinical Problem-Solving
Anchors Away
A 50-year-old Asian woman presented with a papulonodular, erythematous rash on her legs below the knees. The skin lesions were nontender and nonpruritic and were accompanied by paresthesias. She had no fever, arthralgias, or other systemic symptoms.

See p. 504

Clinical Implications of Basic Research
Mechanism of Huntington’s Disease
A protein that regulates gene expression may be the key to the neurodegeneration observed in Huntington’s disease.

See p. 518
Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women

Kristin A. Miller, M.S., David S. Siscovick, M.D., M.P.H., Lianne Sheppard, Ph.D., Kristen Shepherd, M.S., Jeffrey H. Sullivan, M.D., M.H.S., Garnet L. Anderson, Ph.D., and Joel D. Kaufman, M.D., M.P.H.

ABSTRACT

BACKGROUND
Fine particulate air pollution has been linked to cardiovascular disease, but previous studies have assessed only mortality and differences in exposure between cities. We examined the association of long-term exposure to particulate matter of less than 2.5 μm in aerodynamic diameter (PM$_{2.5}$) with cardiovascular events.

METHODS
We studied 65,893 postmenopausal women without previous cardiovascular disease in 36 U.S. metropolitan areas from 1994 to 1998, with a median follow-up of 6 years. We assessed the women’s exposure to air pollutants using the monitor located nearest to each woman’s residence. Hazard ratios were estimated for the first cardiovascular event, adjusting for age, race or ethnic group, smoking status, educational level, household income, body-mass index, and presence or absence of diabetes, hypertension, or hypercholesterolemia.

RESULTS
A total of 1816 women had one or more fatal or nonfatal cardiovascular events, as confirmed by a review of medical records, including death from coronary heart disease or cerebrovascular disease, coronary revascularization, myocardial infarction, and stroke. In 2000, levels of PM$_{2.5}$ exposure varied from 3.4 to 28.3 μg per cubic meter (mean, 13.5). Each increase of 10 μg per cubic meter was associated with a 24% increase in the risk of a cardiovascular event (hazard ratio, 1.24; 95% confidence interval [CI], 1.09 to 1.41) and a 76% increase in the risk of death from cardiovascular disease (hazard ratio, 1.76; 95% CI, 1.25 to 2.47). For cardiovascular events, the between-city effect appeared to be smaller than the within-city effect. The risk of cerebrovascular events was also associated with increased levels of PM$_{2.5}$ (hazard ratio, 1.35; 95% CI, 1.08 to 1.68).

CONCLUSIONS
Long-term exposure to fine particulate air pollution is associated with the incidence of cardiovascular disease and death among postmenopausal women. Exposure differences within cities are associated with the risk of cardiovascular disease.
Exposure to air pollution has been associated with death and hospitalization from cardiovascular causes. Uncertainty remains about the magnitude of these associations, the mechanisms, and the effects of long-term exposure to pollutants, as compared with short-term exposure. Although previous studies of daily increases in exposure to pollution have assessed both fatal and nonfatal events, studies investigating long-term exposure — estimating average exposure during years of follow-up — have evaluated mortality only on the basis of death certificates. The increase in mortality associated with long-term exposure to air pollution is larger than that seen in studies of short-term exposure, and long-term effects on death rates serve as the current basis for fiercely challenged environmental regulations in this country.

In previous studies of the long-term effect of air pollution on cardiovascular disease, investigators have averaged exposures across a city and then compared health effects between cities. However, gradients of exposure to pollutants within cities also affect the risk of death from cardiovascular causes and may be associated with subclinical atherosclerosis.

We evaluated long-term exposure to air pollution and the incidence of cardiovascular disease in the Women’s Health Initiative (WHI) Observational Study, a prospective cohort study with medical-record review and classification procedures designed to document specific first cardiovascular events. We also examined how between-city and within-city gradients of exposure to particulate matter of less than 2.5 μm in aerodynamic diameter (PM$_{2.5}$) are associated with first cardiovascular events.

**METHODS**

**STUDY SUBJECTS**

The WHI enrolled postmenopausal women between the ages of 50 and 79 years in the study from 1994 to 1998. The study design and characteristics of the subjects have been described in detail elsewhere. All subjects lived within commuting distance of one of 49 WHI clinical centers and satellite clinics in 36 U.S. Metropolitan Statistical Areas (referred to throughout as “cities”). Eligible subjects were those who planned to remain in the area and were free from conditions (including alcoholism, mental illness, and dementia) that might have precluded their participation in follow-up surveys. Baseline questionnaires assessed demographic and lifestyle characteristics, cardiovascular risk factors, medical history, diet, and medications. Written informed consent was obtained from all subjects. Anthropometric and blood-pressure measurements were performed at baseline.

We restricted our study population to subjects without a history of physician-diagnosed cardiovascular disease, including previous myocardial infarction, congestive heart failure, coronary revascularization, and stroke. To establish a stable primary residence during follow-up, we included women who lived within 150 mi (241 km) of a clinic (and had not changed clinics) before either death or the year 2002. Institutional review boards at the University of Washington and the Fred Hutchinson Cancer Research Center approved the study.

**DATA ON AIR POLLUTION EXPOSURE**

We obtained data on the monitoring of air pollution from the Environmental Protection Agency’s Aerometric Information Retrieval System with the use of AirData. Such data are recorded for PM$_{2.5}$ and particulate matter of less than 10 μm in aerodynamic diameter (PM$_{10}$), sulfur dioxide, nitrogen dioxide, carbon monoxide, and ozone. We selected monitors on the basis of monitoring objectives and scale to represent ambient community-scale exposure and excluded those with data available from less than 50% of intended samples (see the Supplementary Appendix). On the basis of a five-digit ZIP Code centroid, the nearest monitor to the location of each residence was identified and used to assign an average of annual pollutant concentrations to each study subject. Only women linked to a monitor within 30 mi (48 km) of their residence were included. The long-term average PM$_{2.5}$ concentration was the exposure of interest, and the annual average concentration in the year 2000 was the primary exposure measure, owing to the substantially increased network of monitors in place in that year, as compared with previous years.

**CARDIOVASCULAR OUTCOMES**

The WHI determined events on the basis of subjects’ responses on annual questionnaires and...
review of medical records (including hospital discharge summaries and diagnostic codes, results of electrocardiography, and reports on diagnostic tests and procedures) by physician adjudicators, following an established protocol. Deaths were identified by proxy reports or by a review of the National Death Index. For deaths, WHI adjudicators reviewed all available records, including those from emergency, outpatient, and inpatient departments, and emergency medical services; autopsy and coroner records; and death certificates. We included outcomes adjudicated through August 2003.

The first cardiovascular event was the first occurrence of any of the following: myocardial infarction, coronary revascularization, stroke, and death from either coronary heart disease (categorized as “definite” or “possible”) or cerebrovascular disease. These events were considered to be most consistent with an atherosclerotic disease process and most reliably verified by the WHI protocol. Death from coronary heart disease (both definite and possible diagnoses) required documented myocardial infarction or angina or an antecedent procedure related to coronary artery disease during the follow-up period. Possible deaths from coronary heart disease were those consistent with the condition but without an identifiable nonatherosclerotic cause. Definite deaths from coronary heart disease were those with chest pain within the previous 72 hours or a history of chronic ischemic heart disease, without valvular disease or nonischemic cardiomyopathy. For some analyses, events were classified as coronary events (myocardial infarction, revascularization, and death from coronary heart disease) or cerebrovascular events (stroke and death from cerebrovascular causes). Although some women had more than one type of event, no analysis included multiple events per subject.

Our principal hypothesis concerned levels of PM$_{2.5}$. Single-pollutant and multipollutant models were fit to investigate possible independent or joint effects of other pollutants (see the Supplementary Appendix).

**SENSITIVITY ANALYSES**

We performed sensitivity analyses that excluded subjects living more than 10 mi (16 km) from a monitor or residing for fewer than 20 years in their current state. Other sensitivity analyses included and excluded coronary revascularization and other outcomes; examined a larger group of cardiovascular events, with the addition of angina pectoris, congestive heart failure, transient ischemic attack, other carotid artery disease, and deaths from all cardiovascular causes; and excluded cities with the highest variation of within-city exposure or the lowest exposure concentration (see the Supplementary Appendix).

**STATISTICAL ANALYSIS**

We used Cox proportional-hazards regression to estimate hazard ratios and 95% confidence intervals (CIs) for the time to the first cardiovascular event associated with an elevation of 10 μg per cubic meter in the level of long-term exposure to PM$_{2.5}$. In all models, we included factors that we hypothesized a priori could potentially confound the relationship between air pollution and cardiovascular disease. These factors included age, body-mass index (BMI), smoking status, the number of cigarettes smoked per day, the number of years of smoking, systolic blood pressure, educational level, household income, race or ethnic group, and presence or absence of diabetes, hypertension, or hypercholesterolemia. Models were stratified with use of separate baseline hazards according to current treatment for diabetes, age, and BMI. We also evaluated other characteristics previously associated with the risk of cardiovascular disease — including presence or absence of environmental tobacco smoke, occupation, physical activity, diet, alcohol consumption, waist circumference, waist-to-hip ratio, medical history, medications, and presence or absence of a family history of cardiovascular disease — as possible confounders in extended models.

Interactions between exposure and factors that could modify the association between air pollution and the incidence of cardiovascular disease were evaluated with partial-likelihood ratio tests. For tests of linear trend using the partial-likelihood method, potential effect modifiers measured as continuous variables (such as BMI, waist-to-hip ratio, and waist circumference) or as ordered categorical variables (such as household income, educational level, and years lived in state) were grouped into quintiles.

We created exposure variables to estimate between-city and within-city effects. Exposures for all women in a metropolitan area were averaged into a weighted citywide exposure. Two approaches were used to estimate the within-city effects as part of the overall exposure–effect relationship.
One approach fit indicator variables for each metropolitan area, which we term “city-adjusted.” The other approach subtracted the weighted citywide mean exposure, which we termed “within-city” (see the Supplementary Appendix for details). Data were analyzed with the use of SAS software (version 8.0, SAS Institute) and Stata software (version 8.0, Stata).

RESULTS

STUDY SUBJECTS

Of the 93,676 subjects, 72,569 had no cardiovascular disease at baseline. Of those women, 65,893 (90.8%) returned a follow-up questionnaire, met our residence criteria, and were assigned PM$_{2.5}$ exposure data. We recorded 349,643 women-years of follow-up for the 58,610 women with complete information for the main analytical variables (88.9% of those who were eligible).

Most of the subjects were white (83.1%), and the median age at enrollment was 63 years. The characteristics of the subjects were similar in most respects across categories of PM$_{2.5}$ exposure (Table 1). Race and ethnic group and socioeconomic measures were distributed somewhat unevenly across exposure categories. Current smoking was rare (reported by 6.1% of the subjects), and half of the cohort reported never having smoked. Stable long-term residential location was typical; 85.7% of the subjects had lived for 20 years or more in their current state.

EXPOSURE TO POLLUTION

We linked each woman in the study to one of 573 PM$_{2.5}$ monitors operating in the year 2000, with

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Quintile of Level of PM$_{2.5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.4–10.9 μg/m$^3$ (N=12,906)</td>
</tr>
<tr>
<td>Cardiovascular and cerebrovascular events†</td>
<td>353</td>
</tr>
<tr>
<td>Age — yr</td>
<td>63.1±7.3</td>
</tr>
<tr>
<td>Race or ethnic group — %‡</td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>0.6</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>10.4</td>
</tr>
<tr>
<td>Black</td>
<td>2.1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>8.1</td>
</tr>
<tr>
<td>White</td>
<td>77.4</td>
</tr>
<tr>
<td>Other</td>
<td>1.5</td>
</tr>
<tr>
<td>Education — %</td>
<td></td>
</tr>
<tr>
<td>Not high-school graduate</td>
<td>5.6</td>
</tr>
<tr>
<td>Graduate of high school or trade school or GED</td>
<td>27.8</td>
</tr>
<tr>
<td>Some college or associate degree</td>
<td>28.0</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>38.5</td>
</tr>
<tr>
<td>Household income — %</td>
<td></td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>15.2</td>
</tr>
<tr>
<td>$20,000–49,999</td>
<td>43.2</td>
</tr>
<tr>
<td>≥$50,000</td>
<td>38.5</td>
</tr>
<tr>
<td>Respondent did not know</td>
<td>3.2</td>
</tr>
<tr>
<td>Married — %</td>
<td>65.8</td>
</tr>
<tr>
<td>BMI</td>
<td>26.9±5.6</td>
</tr>
<tr>
<td>Smoking history — %</td>
<td></td>
</tr>
<tr>
<td>Former smoker</td>
<td>40.9</td>
</tr>
<tr>
<td>Current smoker</td>
<td>5.7</td>
</tr>
</tbody>
</table>
a median of 20 monitors per city (range, 4 to 78) (Table 2). Most women lived within 6 mi (10 km) of a monitor. The overall median concentration of fine particle pollution was 13.4 μg per cubic meter (interquartile range, 11.6 to 18.3). The minimum concentration (3.4 μg per cubic meter) was observed in Honolulu, and the maximum (28.3 μg per cubic meter) in Riverside, California.

CARDIOVASCULAR EVENTS
A total of 1816 women had one or more cardiovascular events during the study (Table 3). An increase in exposure of 10 μg per cubic meter in the level of PM$_{2.5}$ was associated with an adjusted hazard ratio of 1.24 (95% CI, 1.09 to 1.41) for the time to the first cardiovascular event. Within-city estimates tended to be larger than between-city estimates, but the differences were not significant (P=0.07); the city-adjusted approach and the estimate of within-city effects yielded similar results (1.69 and 1.64, respectively). A similar pattern emerged for coronary heart disease and cerebrovascular events.

The magnitude of effects observed was largest for mortality end points (Table 3). The strongest overall association was with death definitely associated with coronary heart disease (hazard ratio, 2.21; 95% CI, 1.17 to 4.16), the fatal event characterized by greatest diagnostic certainty. The effect size increased across the range of exposure concentrations that were measured (Fig. 1). We did not observe associations between other pollutants and cardiovascular disease in single-pollutant models, and adjustment for other mea-

Table 1. (Continued.)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>3.4–10.9 μg/m$^3$ (N = 12,906)</th>
<th>11.0–12.4 μg/m$^3$ (N = 13,139)</th>
<th>12.5–14.2 μg/m$^3$ (N = 13,568)</th>
<th>14.3–16.4 μg/m$^3$ (N = 13,035)</th>
<th>16.5–28.3 μg/m$^3$ (N = 13,245)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical activity — MET/wk</td>
<td>14.3±14.9</td>
<td>14.3±14.6</td>
<td>14.6±14.5</td>
<td>14.0±14.6</td>
<td>13.5±14.4</td>
</tr>
<tr>
<td>Hypertension — %</td>
<td>29.8</td>
<td>29.3</td>
<td>28.1</td>
<td>30.0</td>
<td>30.8</td>
</tr>
<tr>
<td>Diabetes mellitus — %</td>
<td>5.0</td>
<td>3.8</td>
<td>3.7</td>
<td>4.9</td>
<td>4.8</td>
</tr>
<tr>
<td>Hypercholesterolemia — %</td>
<td>13.3</td>
<td>12.7</td>
<td>12.1</td>
<td>12.4</td>
<td>13.1</td>
</tr>
<tr>
<td>Waist circumference — cm</td>
<td>83.9±13.3</td>
<td>84.5±13.1</td>
<td>84.0±13.2</td>
<td>84.2±13.5</td>
<td>84.5±13.7</td>
</tr>
<tr>
<td>Hormone-replacement therapy — %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past use</td>
<td>19.8</td>
<td>20.5</td>
<td>20.0</td>
<td>19.4</td>
<td>20.1</td>
</tr>
<tr>
<td>Current use</td>
<td>53.2</td>
<td>49.6</td>
<td>43.4</td>
<td>48.8</td>
<td>51.1</td>
</tr>
<tr>
<td>Alcohol consumption — no. of drinks/wk</td>
<td>2.5±5.0</td>
<td>2.9±5.7</td>
<td>2.7±5.2</td>
<td>2.5±5.0</td>
<td>2.4±5.1</td>
</tr>
<tr>
<td>Time lived in current state — %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤9 yr</td>
<td>9.6</td>
<td>7.4</td>
<td>4.1</td>
<td>6.8</td>
<td>5.3</td>
</tr>
<tr>
<td>10–19 yr</td>
<td>10.0</td>
<td>7.6</td>
<td>5.7</td>
<td>7.9</td>
<td>7.3</td>
</tr>
<tr>
<td>≥20 yr</td>
<td>80.4</td>
<td>85.0</td>
<td>90.2</td>
<td>85.3</td>
<td>87.3</td>
</tr>
<tr>
<td>Time spent outdoors (summer) — %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30 min/day</td>
<td>31.7</td>
<td>30.8</td>
<td>30.2</td>
<td>31.5</td>
<td>33.7</td>
</tr>
<tr>
<td>30 min to 2 hr/day</td>
<td>49.0</td>
<td>49.8</td>
<td>49.6</td>
<td>49.6</td>
<td>49.6</td>
</tr>
<tr>
<td>&gt;2 hr/day</td>
<td>19.4</td>
<td>19.5</td>
<td>20.2</td>
<td>18.9</td>
<td>16.7</td>
</tr>
<tr>
<td>Time spent outdoors (nonsummer) — %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30 min/day</td>
<td>37.2</td>
<td>37.8</td>
<td>38.2</td>
<td>37.9</td>
<td>38.2</td>
</tr>
<tr>
<td>30 min to 2 hr/day</td>
<td>50.4</td>
<td>50.0</td>
<td>50.8</td>
<td>51.0</td>
<td>50.7</td>
</tr>
<tr>
<td>≥2 hr/day</td>
<td>12.5</td>
<td>12.2</td>
<td>11.0</td>
<td>11.1</td>
<td>11.0</td>
</tr>
</tbody>
</table>

* A total of 2113 events are listed for the 65,893 women in the study, even though 7283 of the women had missing data for at least one covariate. Therefore, the main analyses were conducted on data for only 1816 events among 58,610 women. Plus–minus values are means ±SD. Percentages may not total 100 because of rounding. GED denotes general equivalency diploma, BMI body-mass index (the weight in kilograms divided by the square of the height in meters), and MET metabolic equivalent.
† Events include myocardial infarction, revascularization, stroke, and death from coronary heart disease or cerebrovascular disease.
‡ Race or ethnic group was reported by the subjects.
Sensitivity Analyses

A sensitivity analysis incorporating a random-effect term for each city allowed for the possibility that effects might vary from city to city in the estimation of the main effect and its variance for each of the overall, between-city, and within-city effects. Results from this sensitivity analysis were consistent with the primary analysis. The effect estimates were not diminished, and in all cases, the lower confidence limits above 1 did not decrease (see the Supplementary Appendix). Adjustment for additional covariates (i.e., presence or absence of environmental tobacco smoke, occu-
LONG-TERM EXPOSURE TO AIR POLLUTION AND CARDIOVASCULAR EVENTS IN WOMEN

Susceptibility to Effects of Air Pollution

Differences in the relationship between PM$_{2.5}$ and cardiovascular disease according to the characteristics of the subjects are summarized in Table 4. The association between cardiovascular events and the level of PM$_{2.5}$ increased with increasing categories of BMI and waist-to-hip ratio and with a shorter duration of residence in the current state.

DISCUSSION

In a large, prospective cohort of postmenopausal women, long-term (annual average) exposure to increased concentrations of fine particulate air pollution was associated with an increased risk of first cardiovascular events. The increased risk applied to nonfatal and fatal cardiovascular events, including both coronary and cerebrovascular events. We found that estimates of effects within cities were often larger than those of effects between cities; the latter had been the primary measure in previous U.S. studies of long-term exposure to pollutants.

The risk of death associated with higher levels of PM$_{2.5}$ was generally larger than the risk of all first events; it was also larger than mortality estimates reported in previous U.S. cohort studies that used only death certificates. For death from cardiovascular causes (including coronary heart disease and cerebrovascular disease), we estimated an overall 76% increase in risk with each increase of 10 μg per cubic meter in long-term PM$_{2.5}$ exposure — accounting for subjects in approximately the 10th to 90th percentiles for exposure.

Our measurement of the between-city effect is similar to that in the American Cancer Society's
Table 4. Estimated Hazard Ratios for Cardiovascular Events Associated with an Increase of 10 μg per Cubic Meter in the Level of Fine Particulate Matter (PM$_{2.5}$), According to Selected Characteristics.*  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. of Subjects with Event</th>
<th>Hazard Ratio (95% CI)</th>
<th>P Value</th>
<th>Hazard Ratio Adjusted for City (95% CI)†</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects with any cardiovascular event</td>
<td>1816</td>
<td>1.24 (1.09–1.41)</td>
<td>0.64</td>
<td>1.69 (1.26–2.27)</td>
<td>0.81</td>
</tr>
<tr>
<td>Household income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>388</td>
<td>1.30 (1.10–1.53)</td>
<td>0.34</td>
<td>1.75 (1.28–2.40)</td>
<td>0.45</td>
</tr>
<tr>
<td>$20,000–49,999</td>
<td>886</td>
<td>1.23 (1.08–1.41)</td>
<td></td>
<td>1.69 (1.25–2.27)</td>
<td></td>
</tr>
<tr>
<td>≥$50,000</td>
<td>542</td>
<td>1.20 (1.02–1.40)</td>
<td></td>
<td>1.66 (1.22–2.26)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.34</td>
<td></td>
<td>0.54</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td>0.22</td>
<td></td>
<td>0.31</td>
</tr>
<tr>
<td>Not high-school graduate</td>
<td>112</td>
<td>1.40 (1.11–1.75)</td>
<td>0.34</td>
<td>1.88 (1.32–2.67)</td>
<td>0.42</td>
</tr>
<tr>
<td>Graduate of high school or trade school or GED</td>
<td>575</td>
<td>1.33 (1.14–1.55)</td>
<td></td>
<td>1.79 (1.32–2.44)</td>
<td></td>
</tr>
<tr>
<td>Some college or associate degree</td>
<td>514</td>
<td>1.26 (1.09–1.44)</td>
<td></td>
<td>1.74 (1.29–2.34)</td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>615</td>
<td>1.11 (0.94–1.31)</td>
<td></td>
<td>1.54 (1.13–2.10)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.07</td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td>0.42</td>
</tr>
<tr>
<td>&lt;60 yr</td>
<td>234</td>
<td>1.21 (0.84–1.73)</td>
<td>0.20</td>
<td>1.66 (1.05–2.61)</td>
<td>0.20</td>
</tr>
<tr>
<td>60–69 yr</td>
<td>785</td>
<td>1.14 (0.93–1.39)</td>
<td></td>
<td>1.53 (1.09–2.14)</td>
<td></td>
</tr>
<tr>
<td>≥70 yr</td>
<td>797</td>
<td>1.34 (1.11–1.63)</td>
<td></td>
<td>1.85 (1.34–2.56)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.36</td>
<td></td>
<td>0.38</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td>0.55</td>
<td></td>
<td>0.51</td>
</tr>
<tr>
<td>Current smoker</td>
<td>150</td>
<td>1.68 (1.06–2.66)</td>
<td>0.20</td>
<td>2.28 (1.33–3.92)</td>
<td>0.20</td>
</tr>
<tr>
<td>Former smoker</td>
<td>750</td>
<td>1.24 (1.01–1.52)</td>
<td></td>
<td>1.71 (1.23–2.39)</td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>916</td>
<td>1.18 (0.99–1.40)</td>
<td></td>
<td>1.60 (1.16–2.21)</td>
<td></td>
</tr>
<tr>
<td>Living with smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently</td>
<td>158</td>
<td>1.28 (0.84–1.97)</td>
<td>0.02</td>
<td>1.65 (0.99–2.76)</td>
<td>0.02</td>
</tr>
<tr>
<td>Formerly</td>
<td>1206</td>
<td>1.18 (1.00–1.38)</td>
<td></td>
<td>1.59 (1.16–2.16)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>436</td>
<td>1.39 (1.07–1.80)</td>
<td></td>
<td>1.90 (1.31–2.78)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.02</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Body-mass index</td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>&lt;22.5</td>
<td>227</td>
<td>0.99 (0.80–1.21)</td>
<td>0.003</td>
<td>1.35 (0.96–1.88)</td>
<td>0.004</td>
</tr>
<tr>
<td>22.5–24.7</td>
<td>337</td>
<td>1.16 (0.96–1.40)</td>
<td></td>
<td>1.58 (1.14–2.19)</td>
<td></td>
</tr>
<tr>
<td>24.8–27.2</td>
<td>359</td>
<td>1.24 (1.05–1.45)</td>
<td></td>
<td>1.69 (1.24–2.30)</td>
<td></td>
</tr>
<tr>
<td>27.3–30.9</td>
<td>439</td>
<td>1.38 (1.18–1.61)</td>
<td></td>
<td>1.88 (1.38–2.56)</td>
<td></td>
</tr>
<tr>
<td>&gt;30.9</td>
<td>454</td>
<td>1.35 (1.12–1.64)</td>
<td></td>
<td>1.84 (1.33–2.55)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.02</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Waist-to-hip ratio</td>
<td></td>
<td></td>
<td>0.02</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>&lt;0.74</td>
<td>199</td>
<td>1.07 (0.90–1.29)</td>
<td>0.003</td>
<td>1.45 (1.05–2.00)</td>
<td>0.004</td>
</tr>
<tr>
<td>0.74–0.77</td>
<td>272</td>
<td>1.12 (0.95–1.31)</td>
<td></td>
<td>1.51 (1.11–2.06)</td>
<td></td>
</tr>
<tr>
<td>0.78–0.80</td>
<td>305</td>
<td>1.24 (1.07–1.44)</td>
<td></td>
<td>1.68 (1.23–2.27)</td>
<td></td>
</tr>
<tr>
<td>0.81–0.86</td>
<td>482</td>
<td>1.30 (1.13–1.50)</td>
<td></td>
<td>1.76 (1.30–2.38)</td>
<td></td>
</tr>
<tr>
<td>&gt;0.86</td>
<td>558</td>
<td>1.29 (1.11–1.50)</td>
<td></td>
<td>1.75 (1.29–2.37)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td></td>
<td></td>
<td>0.02</td>
<td></td>
<td>0.02</td>
</tr>
</tbody>
</table>

*From the analyses conducted using 10 μg/m$^3$ as the baseline level of PM$_{2.5}$.
†Adjusted for city of residence.

[From the NEJM website]
Table 4. (Continued.)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. of Subjects with Event</th>
<th>Hazard Ratio (95% CI)</th>
<th>P Value</th>
<th>Hazard Ratio Adjusted for City (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waist circumference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;73 cm</td>
<td>203</td>
<td>1.05 (0.86–1.27)</td>
<td>0.09</td>
<td>1.43 (1.02–1.99)</td>
<td>0.11</td>
</tr>
<tr>
<td>73–78 cm</td>
<td>295</td>
<td>1.20 (1.02–1.41)</td>
<td></td>
<td>1.63 (1.19–2.23)</td>
<td></td>
</tr>
<tr>
<td>79–85 cm</td>
<td>373</td>
<td>1.22 (1.05–1.41)</td>
<td></td>
<td>1.66 (1.22–2.24)</td>
<td></td>
</tr>
<tr>
<td>86–95 cm</td>
<td>452</td>
<td>1.33 (1.15–1.53)</td>
<td></td>
<td>1.80 (1.33–2.43)</td>
<td></td>
</tr>
<tr>
<td>&gt;95 cm</td>
<td>493</td>
<td>1.27 (1.07–1.51)</td>
<td></td>
<td>1.73 (1.26–2.36)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td>0.06</td>
<td></td>
<td></td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Hormone-replacement therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current use</td>
<td>704</td>
<td>1.33 (1.09–1.61)</td>
<td>0.13</td>
<td>1.85 (1.32–2.58)</td>
<td>0.08</td>
</tr>
<tr>
<td>No current use</td>
<td>1077</td>
<td>1.16 (0.98–1.39)</td>
<td></td>
<td>1.57 (1.14–2.17)</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>219</td>
<td>0.96 (0.67–1.37)</td>
<td></td>
<td>1.24 (0.78–1.96)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1597</td>
<td>1.28 (1.12–1.47)</td>
<td></td>
<td>1.75 (1.30–2.36)</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>935</td>
<td>1.22 (1.02–1.45)</td>
<td>0.79</td>
<td>1.65 (1.09–2.27)</td>
<td>0.69</td>
</tr>
<tr>
<td>No</td>
<td>881</td>
<td>1.26 (1.05–1.51)</td>
<td></td>
<td>1.74 (1.25–2.40)</td>
<td></td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td></td>
<td></td>
<td>0.92</td>
<td></td>
<td>0.94</td>
</tr>
<tr>
<td>Yes</td>
<td>333</td>
<td>1.25 (0.94–1.67)</td>
<td></td>
<td>1.71 (1.15–2.54)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1483</td>
<td>1.23 (1.07–1.42)</td>
<td></td>
<td>1.69 (1.25–2.28)</td>
<td></td>
</tr>
<tr>
<td>Family history of cardiovascular disease</td>
<td></td>
<td></td>
<td>0.19</td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>Yes</td>
<td>1359</td>
<td>1.30 (1.12–1.51)</td>
<td></td>
<td>1.80 (1.32–2.44)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>436</td>
<td>1.07 (0.83–1.37)</td>
<td></td>
<td>1.46 (1.00–2.12)</td>
<td></td>
</tr>
<tr>
<td>Time lived in current state</td>
<td></td>
<td></td>
<td>0.39</td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td>≥20 yr</td>
<td>1585</td>
<td>1.21 (1.06–1.39)</td>
<td></td>
<td>1.66 (1.23–2.23)</td>
<td></td>
</tr>
<tr>
<td>10–19 yr</td>
<td>114</td>
<td>1.39 (1.12–1.72)</td>
<td></td>
<td>1.97 (1.40–2.79)</td>
<td></td>
</tr>
<tr>
<td>≤9 yr</td>
<td>108</td>
<td>1.54 (1.06–2.26)</td>
<td></td>
<td>2.24 (1.39–3.59)</td>
<td></td>
</tr>
<tr>
<td>P for trend</td>
<td>0.05</td>
<td></td>
<td></td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Health insurance coverage</td>
<td></td>
<td></td>
<td>0.33</td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Yes</td>
<td>1763</td>
<td>1.22 (1.07–1.39)</td>
<td></td>
<td>1.71 (1.27–2.30)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>42</td>
<td>1.82 (0.81–4.10)</td>
<td></td>
<td>2.65 (1.12–6.28)</td>
<td></td>
</tr>
<tr>
<td>Time spent outdoors</td>
<td></td>
<td></td>
<td>0.35</td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td>&lt;30 min</td>
<td>510</td>
<td>1.09 (0.86–1.39)</td>
<td></td>
<td>1.56 (1.05–2.31)</td>
<td></td>
</tr>
<tr>
<td>≥30 min</td>
<td>945</td>
<td>1.26 (1.05–1.50)</td>
<td></td>
<td>1.82 (1.29–2.57)</td>
<td></td>
</tr>
</tbody>
</table>

* All estimates were adjusted for age, race or ethnic group, educational level, household income, smoking status, systolic blood pressure, body-mass index, and presence or absence of a history of diabetes, hypertension, or hypercholesterolemia. Data were missing for some subjects, so the number of subjects in each category may not total 1816.
† The city-adjusted models included an indicator variable for each metropolitan area.

Cancer Prevention Study II and the Harvard Six Cities Study. The between-city hazard ratio was 1.63 for death from cardiovascular causes as defined in our study and 1.42 for the broader definition similar to that used in the other studies. In the independent reanalysis, the estimated hazard ratio for death from all cardiovascular causes associated with an increase of 10 μg per cubic meter in long-term PM$_{2.5}$ exposure was 1.19 (95% CI, 1.05 to 1.34) in the Six Cities Study and
1.13 (95% CI, 1.08 to 1.18) in the American Cancer Society’s study. However, subjects in those cohorts differed substantially from ours, especially by the inclusion of men and persons with previous cardiovascular disease. The larger effect sizes observed for levels of PM$_{2.5}$ in our study may be due to these factors or to our efforts to reduce misclassification of outcomes and exposures. However, other studies have suggested greater effects of particulate air pollution in women than in men.$^{16,21}$ The increased association we observed between the PM$_{2.5}$ level and death from cardiovascular causes, as compared with all cardiovascular events, could be related to methodologic considerations, such as a reduced misclassification of fatal events. Alternatively, fine particulate exposures could exert effects that disproportionately result in fatal events (such as arrhythmic events or hemorrhagic stroke), as compared with nonfatal cardiovascular events.

In addition to the increased risk of coronary heart disease, we identified an association between long-term exposure to air pollution and the incidence of cerebrovascular disease. For each increase of 10 μg per cubic meter of exposure, there was a 35% increase in the risk of cerebrovascular events and an 83% increase in the risk of death from cerebrovascular causes. Previous evidence in this area included ecologic studies suggesting that the rate of death from stroke may be elevated in areas near main roads or with increased pollution,$^{22,23}$ and short-term exposure has been linked to stroke, for example.$^{24}$

We observed a stronger association between the PM$_{2.5}$ level and cardiovascular disease with increasing obesity, as measured by either the BMI or the waist-to-hip ratio. These findings require replication. In contrast to the findings in the American Cancer Society’s study,$^{5,6}$ we observed a uniform pollution-related risk of cardiovascular events across age groups, possibly because of the greater homogeneity of subjects in our study.

Our study benefited from well-defined outcomes, extensive data regarding risk factors for cardiovascular disease, and long-term geographic stability of a cohort without previous cardiovascular disease. Critics of earlier studies have suggested that poorly measured or unmeasured confounding factors may vary from city to city and account, at least in part, for the observed city-to-city differences in death rates associated with air pollution.$^{25}$ We ascertained key characteristics of subjects that might confound the relationship with exposure, and study results were not sensitive to adjustment for these characteristics, although some residual confounding cannot be excluded. Aspects of our analytic approach also reduce the concern over confounding, such as our examination of the between-city and within-city components of exposure. We controlled for the factors that vary from city to city (e.g., imperfectly measured subject characteristics, the composition or toxicity of particulate matter, and particle infiltration) in the analysis, which included a city indicator variable. By investigating many potential covariates, and by including both within-city and between-city exposures, we provided confirmation of the observed association between long-term exposure to air pollution and cardiovascular disease.

The role of socioeconomic status has received attention in air-pollution epidemiology. Beyond controlling for educational level and household income, our results were not sensitive to further adjustment for occupation or Census-derived measures of income, wealth, or poverty on the basis of ZIP Code. Neither educational level nor household income significantly modified the relationship between air pollution and cardiovascular disease, although there was a trend toward greater effects among those with less education.

Since our study included a large number of women who lived in many locales, regional or smaller-scale differences in medical practice might have influenced our findings, particularly regarding coronary revascularization. However, the results did not change when revascularization procedures were not included in the analysis. Furthermore, it is possible that women living close to one another may be more similar than those living farther apart, which could affect the variance estimation. However, we found no evidence that such a bias influenced our results.

Our assessment of exposure remains necessarily limited, since exposure levels were assigned from one monitor with the use of the subjects’ primary residential ZIP Codes, which could potentially introduce some inaccuracy. The degree to which ambient pollution monitors represent the exposures of the subjects is imperfect, though we took measures to exclude nonrepresentative monitors. We were unable to assess microclimate differences in exposure, whether some participants may have moved, or details regarding the
LONG-TERM EXPOSURE TO AIR POLLUTION AND CARDIOVASCULAR EVENTS IN WOMEN

subjects' activity and location, such as time spent in traffic and indoors. These factors contribute to errors in measurement and misclassification of exposure but are unlikely to have introduced a bias that would explain the study's findings.

We used data regarding PM$_{2.5}$ levels from a single year at the midpoint of follow-up, rather than a baseline or multiyear average, because of the presence of a much greater number of measuring stations in the year 2000 than at other times. Concentrations of particulate pollution were stable during the period under study; year-to-year PM$_{2.5}$ values were very highly correlated during years available (Pearson's correlation, $r=0.92$). Analysis in the American Cancer Society's study showed a strong correlation between the sites at a 20-year interval and no time dependence of the hazard function, indicating that fine particulate pollution measured at any time during follow-up is a reasonable surrogate for the relevant exposure (long-term exposure to particulate matter).$^{26}$

The mechanism by which long-term exposure to fine particulate air pollution may increase the risk of cardiovascular disease remains uncertain and the subject of intensive speculation and investigation.$^{1,14,27}$ Accelerated atherosclerosis and vulnerability to plaque rupture have been documented in an experimental model,$^{28}$ and ambient pollution has been correlated with carotid intima-media thickness in humans.$^{16}$

Our results were specific to fine particulate pollution; we did not observe robust effects with other sizes of particulate matter or with other measured air pollutants (see the Supplementary Appendix). Finally, since we investigated long-term exposure and first cardiovascular events, the results in this cohort cannot be ascribed primarily to the short-term effects of increases in levels of pollution occurring on a day-to-day basis or to effects of pollution limited to those who were already ill.

Our study provides evidence of the association between long-term exposure to air pollution and the incidence of cardiovascular disease. Our study confirms previous reports and indicates that the magnitude of health effects may be larger than previously recognized. These results suggest that efforts to limit long-term exposure to fine particulate pollution are warranted.

Supported by grants from the Environmental Protection Agency (R827355, to the Northwest Center for Particulate Matter and Health) and the National Institute of Environmental Health Sciences (T32ES07262, K24ES013195, K23ES011139, and P30ES07043, to the University of Washington). The WHI program is funded by the National Heart, Lung, and Blood Institute.

No potential conflict of interest relevant to this article was reported.

The views expressed in this article do not necessarily reflect the views or policies of the Environmental Protection Agency.

We thank the WHI investigators (listed in the Supplementary Appendix); the women who participated in this study; and Nancy Bonnimming, Jill Shupe, and Jonathan Schildcrout for their assistance with data management.

REFERENCES


Copyright © 2007 Massachusetts Medical Society.
JAK2 Exon 12 Mutations in Polycythemia Vera and Idiopathic Erythrocytosis

Linda M. Scott, Ph.D., Wei Tong, Ph.D., Ross L. Levine, M.D.,
Mike A. Scott, Ph.D., Philip A. Beer, M.R.C.P., M.R.C.Path.,
Michael R. Stratton, M.D., Ph.D., P. Andrew Futreal, Ph.D.,
Wendy N. Erber, M.D., Mary Frances McMullin, F.R.C.P., F.R.C.Path.,
D. Gary Gilliland, M.D., Ph.D., Harvey F. Lodish, Ph.D.,
and Anthony R. Green, F.R.C.Path., F.Med.Sci.

BACKGROUND

The V617F mutation, which causes the substitution of phenylalanine for valine at position 617 of the Janus kinase (JAK) 2 gene \(\text{(JAK2)}\), is often present in patients with polycythemia vera, essential thrombocythemia, and idiopathic myelofibrosis. However, the molecular basis of these myeloproliferative disorders in patients without the V617F mutation is unclear.

METHODS

We searched for new mutations in members of the JAK and signal transducer and activator of transcription (STAT) gene families in patients with V617F-negative polycythemia vera or idiopathic erythrocytosis. The mutations were characterized biochemically and in a murine model of bone marrow transplantation.

RESULTS

We identified four somatic gain-of-function mutations affecting JAK2 exon 12 in 10 V617F-negative patients. Those with a JAK2 exon 12 mutation presented with an isolated erythrocytosis and distinctive bone marrow morphology, and several also had reduced serum erythropoietin levels. Erythroid colonies could be grown from their blood samples in the absence of exogenous erythropoietin. All such erythroid colonies were heterozygous for the mutation, whereas colonies homozygous for the mutation occur in most patients with V617F-positive polycythemia vera. BaF3 cells expressing the murine erythropoietin receptor and also carrying exon 12 mutations could proliferate without added interleukin-3. They also exhibited increased phosphorylation of JAK2 and extracellular regulated kinase 1 and 2, as compared with cells transduced by wild-type JAK2 or V617F JAK2. Three of the exon 12 mutations included a substitution of leucine for lysine at position 539 of JAK2. This mutation resulted in a myeloproliferative phenotype, including erythrocytosis, in a murine model of retroviral bone marrow transplantation.

CONCLUSIONS

JAK2 exon 12 mutations define a distinctive myeloproliferative syndrome that affects patients who currently receive a diagnosis of polycythemia vera or idiopathic erythrocytosis.
The myeloproliferative disorders comprise a spectrum of chronic hematologic diseases that are likely to arise from a mutan multipotent hematopoietic stem cell. The V617F somatic mutation in the Janus kinase (JAK) 2 gene (JAK2), which causes the substitution of phenylalanine for valine at position 617, has recently been found in the majority of patients with polycythemia vera and in many with essential thrombocytopenia or idiopathic myelofibrosis. This gene encodes a cytoplasmic tyrosine kinase. The mutation, which occurs in the JAK homology 2 (JH2) negative regulatory domain, increases JAK2 kinase activity and causes cytokine-independent growth of cell lines and cultured bone marrow cells. Mutant JAK2 transfected into murine bone marrow cells produces erythrocytosis and subsequent myelofibrosis in recipient animals, suggesting a causal role for the mutation. 

Allele-specific polymerase chain reaction (PCR) can be used to detect the V617F mutation in approximately 95% of patients with polycythemia vera and in 50 to 60% of patients with essential thrombocytopenia or idiopathic myelofibrosis. The mutation is also present in hematopoietic progenitors committed to granulocytic or erythroid differentiation and in purified hematopoietic stem cells from patients with polycythemia vera. 

Many patients with polycythemia vera or idiopathic myelofibrosis are homozygous for the V617F mutation, as a result of mitotic recombination affecting chromosome 9p, but homozygosity is rare in patients with essential thrombocytopenia. The mutation occurs infrequently in patients with myelodysplasia or acute myeloid leukemia but does not occur in those with lymphoid tumors, epithelial cancers, or sarcomas. 

The JAK2 mutation allows for a distinction between two subtypes of idiopathic myelofibrosis and essential thrombocytopenia. The phenotype of V617F-positive, but not V617F-negative, essential thrombocytopenia resembles that of polycythemia vera. However, patients with V617F-negative essential thrombocytopenia do have cytogenetic abnormalities, dysplastic megakaryocytes, and a risk of transformation to myelofibrosis or acute myeloid leukemia, all of which are features of a myeloproliferative disorder. Activating mutations in the thrombopoietin receptor have been reported in 10% of patients with V617F-negative idiopathic myelofibrosis and in a few patients with essential thrombocytopenia. However, the molecular basis of V617F-negative polycythemia vera is unknown.

**METHODS**

**PATIENTS**

We recruited patients from Addenbrooke’s Hospital in Cambridge, St. Thomas’ Hospital in London, and Belfast City Hospital in Belfast (all in the United Kingdom) and from those enrolled in the Myeloproliferative Disorders Study of Harvard University in Boston. Diagnoses assigned by local physicians were reviewed centrally and revised according to established criteria for polycythemia vera, essential thrombocytopenia, and idiopathic myelofibrosis. The Addenbrooke’s National Health Service Trust Research Ethics Committee approved this study. Written informed consent was obtained from each patient.

**MUTATION SCREENING**

The isolation of granulocytes and T lymphocytes and hematopoietic colony assays were performed as previously described. Individual burst-forming units and erythropoietin-independent erythroid colonies were harvested into water and boiled. Primers for the coding exons of JAK1, JAK2, JAK3, the tyrosine kinase 2 gene (TYK2), and of two signal transducer and activator of transcription genes (STAT5A and STAT5B) are listed at www.sanger.ac.uk/genetics/CGP; all additional primers used are listed in Table 1 in the Supplementary Appendix (available with the full text of this article at www.nejm.org). We performed allele-specific PCR using DNA from granulocytes or from total peripheral blood, an annealing temperature of 62°C, JAK2 exon 12 control primers, and primers specific for the alleles containing the K539L mutation (leading to the replacement of lysine at position 539 with a leucine), the N542-E543del mutation (causing the deletion of asparagine at position 542 and glutamic acid at position 543), the F537-K539delinsL mutation (leading to the replacement of phenylalanine at position 537 through lysine at position 539 by a single leucine), or the H538QK539L mutation (causing a substitution of glutamine for histidine at position 538 and leucine for lysine at position 539). We amplified DNA from in vitro colonies using exon 12 primers and sequenced or genotyped the PCR products using digestion with AspI.
Bone marrow biopsy specimens from the iliac crest were fixed in neutral buffered formalin. Some were processed in paraffin and others in methylmethacrylate after decalcification in 5.5% EDTA. Sections (1 to 3 μm thick) were cut and visualized using hematoxylin and eosin or Wright–Giemsa stain. All stained sections were viewed under a light microscope (Olympus-BX51) equipped with a 10×-H26.5 ocular lens. Low-power (20×) and high-power (40×) images were obtained with a digital camera (Pixera Pro150ES) and Studio 3.0.1 software (Adobe Systems).

Site-directed mutagenesis and production of retrovirus

We introduced the mutations V617F, H538QK539L, K539L, N542-E543del, and F537-K539delinsL into murine Jak2 complementary DNA in a bicistronic retroviral vector encoding green fluorescent protein (MSCViresGFP), using QuikChange site-directed mutagenesis (Stratagene). The complete nucleotide sequence of each retroviral vector was confirmed before use. For the production of each retrovirus, equal amounts of Jak2 retroviral vector and packaging plasmids (Ecopak) were combined, incubated with FuGene (Roche) for 15 minutes, and then added to the human embryonic kidney-cell line, 293T. The supernatants were harvested 48 hours later and were used to transduce BaF3 cells expressing the murine erythropoietin receptor (BaF3/EpoR cells) or murine bone marrow cells.

BaF3-cell proliferation assays and Western blotting

BaF3/EpoR cells were maintained in RPMI-1640 medium containing 10% fetal-calf serum and 10% medium conditioned with WEHI-3B cells, as a source of interleukin-3, and infected with retroviral supernatants containing MSCViresGFP vectors encoding mutant or wild-type Jak2. The green fluorescent protein–positive population from each transduction was confirmed before use. For the production of each retrovirus, equal amounts of Jak2 retroviral vector and packaging plasmids (Ecopak) were combined, incubated with FuGene (Roche) for 15 minutes, and then added to the human embryonic kidney-cell line, 293T. The supernatants were harvested 48 hours later and were used to transduce BaF3 cells expressing the murine erythropoietin receptor (BaF3/EpoR cells) or murine bone marrow cells.

Bone marrow transplantation assay in mice

Bone marrow transplantation was performed as previously described. Briefly, retroviral supernatants were titrated by determining the percentage of BaF3 cells that were positive for green fluorescent protein 48 hours after the introduction of the retroviral vector. Supernatants containing equal titers of wild-type Jak2 or V617F or K539L Jak2 were used to transfect bone marrow cells. BALB/c donor mice were treated with 150 mg of 5-fluorouracil per kilogram of body weight, and cells harvested from femurs and tibias 7 days later were cultured for 24 hours in transplantation medium (RPMI-1640 medium, 10% fetal-calf serum, 6 ng of murine interleukin-3 per milliliter, 10 ng of human interleukin-6 per milliliter, and 10 ng of murine stem-cell factor per milliliter). Bone marrow cells were centrifuged at 2500 rpm for 90 minutes in the presence of 1 ml of retroviral supernatant and 10 μg of polybrene per 4×10^6 cells. Exposure to retroviral supernatant and centrifugation...
were repeated 1 day later. Aliquots of $1 \times 10^6$ bone marrow cells were resuspended in 0.7 ml of Hank’s balanced salt solution and then injected into lethally irradiated BALB/c mice. Peripheral-blood counts and cell morphology were evaluated for each recipient 38 days after transplantation.

### Statistical Analysis

We used an unpaired Student’s t-test to compare demographic and laboratory features at the time of diagnosis between patients with a V617F JAK2 mutation and those with a JAK2 exon 12 mutation and to compare peripheral-blood counts among

---

#### Figure 1. Somatic Mutations of JAK2 Exon 12 in Patients with Polycythemia Vera or Idiopathic Erythrocytosis.

Panel A shows DNA-sequence traces from peripheral-blood granulocytes and T lymphocytes and from erythropoietin-independent erythroid colonies. Nucleotides are indicated by capital letters, with N representing sites at which wild-type and mutant nucleotides are apparent at the same position. The traces reveal four acquired mutations within JAK2 exon 12 (indicated by arrowheads), often with low-level involvement in granulocytes. Panel B (top) shows the alignment of wild-type and mutant exon 12 JAK2 alleles (shown in red) (nucleotides are indicated by capital letters and amino acids by bold capital letters; dashes indicate the positions of deleted nucleotides). The amino acid alignment across multiple species (Panel B, bottom) shows conservation of the mutated amino acids, indicated in red.
mouse recipients of bone marrow cells expressing wild-type \( \text{V}617F \), or \( \text{K}539L \) Jak. Fisher’s exact test was used to compare frequencies of mutations in the wild-type, \( \text{V}617F \), or \( \text{K}539L \) Jak.

### Results

#### Somatic Mutations Affecting JAK2 Exon 12

Of the 73 patients with polycythemia vera in our original cohort, 2 did not have the \( \text{V}617F \) mutation and were studied further. In these two patients, mutations were not found in the coding exons of \( \text{JAK1, JAK3, TYK2, STAT5A, or STAT5B} \). However, both patients had alterations in \( \text{JAK2} \) exon 12 that affected residues laying approximately 80 amino acids before \( \text{V}617F \).

- One patient had a 6-bp in-frame deletion affecting positions 1611 to 1616, resulting in an \( \text{F}537-\text{K}539\text{delinsL} \) mutation.
- The second patient had a \( \text{CAA} \rightarrow \text{ATT} \) mutation at positions 1614 through 1616, resulting in an \( \text{H}538\text{QK}539\text{L} \) mutation (Fig. 1A).

These mutations were acquired, since they could be detected in peripheral-blood granulocytes but not in T lymphocytes.

### JAK2 Exon 12 Mutations

JAK2 exon 12 mutations were identified in eight of an additional nine patients who received a diagnosis of \( \text{V}617F \)-negative polycythemia vera from their local physicians. The mutations were frequently present at low levels in granulocyte DNA but were readily identifiable in clonally derived erythropoietin-independent erythroid colonies (Fig. 1A).

In total, four exon 12 alleles were identified, all of which had changes affecting conserved residues between \( \text{K}537 \) and \( \text{E}543 \) (Fig. 1); three of the alleles (in Patients 1 through 6) contained a \( \text{K}539\text{L} \) substitution.

### JAK2 Exon 12 Mutations

JAK2 exon 12 mutations were not detected by sequencing granulocyte DNA from 55 patients with \( \text{V}617F \)-positive polycythemia vera, 25 patients with \( \text{V}617F \)-negative essential thrombocythemia, and 12 patients with \( \text{V}617F \)-negative cases of idiopathic myelofibrosis (and data not shown). Since mutation-bearing granulocytes may represent only a minority of peripheral blood granulocytes, DNA from an additional 90 patients with \( \text{V}617F \)-negative essential thrombocythemia was screened using sensitive allele-specific PCR assays for each exon 12 mutation, but no mutations were detected (data not shown). These results indicate that JAK2 exon 12 mutations occur only in patients with myelo...
proliferative syndrome who present with erythrocytosis.

**CLINICAL PHENOTYPE ASSOCIATED WITH JAK2 EXON 12 MUTATIONS**

Table 1 shows the clinical and laboratory features of the patients with exon 12 mutations. All had platelet counts of $450 \times 10^3$ or less per cubic millimeter and neutrophil counts that were within the normal range or were insufficiently raised to fulfill the criteria for a diagnosis of polycythemia vera. A low serum erythropoietin level was found in four of eight tested patients, and in six of six tested patients, erythropoietin-independent erythroid colonies could be grown from peripheral-blood cells, a key feature of the myeloproliferative disorders. Central review of clinical and laboratory features revealed that six patients fulfilled the criteria of the Polycythemia Vera Study Group for polycythemia vera, and four patients fulfilled criteria for idiopathic erythrocytosis. Patients with exon 12 mutations were significantly younger at diagnosis than 86 patients from Addenbrooke’s Hospital who had V617F-positive polycythemia vera (median age, 52 years vs. 58 years; P = 0.003) and had significantly higher hemoglobin levels (mean, 202 g per liter vs. 180 g per liter; P = 0.002), lower white-cell counts (mean, $8.4 \times 10^3$ per cubic millimeter vs. $14.1 \times 10^3$ per cubic millimeter; P = 0.008), and lower platelet counts (mean, $311 \times 10^3$ per cubic millimeter vs. $605 \times 10^3$ per cubic millimeter; P < 0.001) (Table 2 in the Supplementary Appendix). Bone marrow trephine biopsy was performed in five patients at diagnosis; the biopsy specimens were examined in a blinded manner. All showed a characteristic pattern of erythroid hyperplasia without morphologic abnormalities of the megakaryocyte or granulocyte lineages (Fig. 2, and Fig. 1A in the Supplementary Appendix).

Hematopoietic progenitors that are homozygous for the V617F mutation are detectable in most patients with polycythemia vera. To seek such homozygosity in patients with exon 12 mutations, individual hematopoietic progenitors from Patients 3, 4, 5, and 7 were genotyped with the use of Asel digestion (Fig. 2B in the Supplementary Appendix), sequence analysis, or both. Homozygosity was not observed in any of the 151 erythroid colonies carrying an exon 12 mutation, whether they were grown in the presence or absence of erythropoietin (Fig. 2C in the Supplementary Appendix). In one patient, granulocyte–macrophage colonies were also heterozygous for the exon 12 mutation, demonstrating that this genetic change occurred at the level of the common myeloid progenitor or the hematopoietic stem cell.

**PROLIFERATION AND SIGNALING IN CELLS BEARING EXON 12 MUTATIONS**

The expression of each Jak2 exon 12 mutant in interleukin-3–dependent BaF3/EpoR cells caused
the cells to proliferate in the absence of added exogenous cytokine, with kinetics indistinguishable from those observed for cells with the V617F mutation (Fig. 3A). This proliferation required expression of the erythropoietin receptor; it was not observed in parental BaF3 cells (data not shown). In the absence of stimulation with erythropoietin, all mutants were consistently associated with increased levels of tyrosine-phosphorylated Jak2 and Stat5, as compared with wild-type Jak2 (Fig. 3B). Moreover, the three alleles containing a K539L substitution all generated consistently higher levels of phosphorylated Erk1 and Erk2 that were markedly higher than those obtained with wild-type Jak2 and higher than those obtained with V617F Jak2 (Fig. 3C). In summary, when transduced into BaF3/EpoR cells, all four Jak2 exon 12 mutations caused growth-factor hypersensitivity and activated biochemical pathways associated with erythropoietin signaling.

**Figure 3. Proliferation and Increased Signaling in the Absence of Exogenous Cytokine from Jak2 Exon 12 Mutations.**

BaF3/EpoR cells (10⁵ per cubic millimeter) — transduced with an empty retroviral vector or stably expressing wild-type murine Jak2 or Jak2 with V617F, F537-K539delinsL, H538Q/K539L, K539L, or N542-E543del mutations — were cultured in the absence of interleukin-3 for 4 days (Panel A). On days 2 and 4, we assessed cell numbers and viability in quadruplicate using trypan-blue exclusion. Results reflect four independent experiments; mean (±SD) counts for each cell line at both time points are shown. BaF3/EpoR cells transduced with an empty MSCViresGFP retroviral vector (Panel B), or BaF3/EpoR cells containing wild-type Jak2 or Jak2 with V617F, F537-K539delinsL, H538Q/K539L, N542-E543del, or K539L mutations were depleted of cytokines for 4 hours. Cells were lysed and underwent immunoprecipitation (IP) with antibody specific for Jak2 or Stat5. Western blot (WB) was then performed with antibodies against phosphotyrosine (4G10), total Jak2, phosphotyrosine-694 Stat5, or total Stat5 (Panel B). BaF3/EpoR cells expressing the Jak2 alleles were analyzed by Western blot with antibodies specific for phosphorylated or total extracellular regulated kinase 1 (Erk1) and 2 (Erk2) (Panel C). BaF3/EpoR cells stimulated with 10 U per milliliter of erythropoietin for 10 minutes were used as positive controls in Panels B and C. Plus signs indicate presence and minus signs absence of exogenous Jak2 or erythropoietin.

**RETROVIRAL TRANSFER OF JAK2 MUTATIONS INTO MICE**

To assess the effects of exon 12 mutations in vivo, murine bone marrow cells were transduced with retroviral vectors encoding wild-type, V617F, or K539L Jak2 and then were transplanted into lethally irradiated BALB/c mice, which are especially susceptible to the development of myeloid disorders after transfer of the V617F mutant.⁸ Five weeks after transplantation, animals that received V617F-transduced bone marrow cells had erythrocytosis and leukocytosis (Fig. 4A), results that are consistent with previous observations.⁸
as well as a modest thrombocytosis. Recipients of K539L-transduced cells also had an elevated hematocrit, reticulocytosis, and leukocytosis and a modest thrombocytosis (Fig. 4). Consistent with the human phenotypes associated with exon 12 and V617F mutations, the mean white-cell and platelet counts were lower in recipients of K539L-transduced cells than in recipients of V617F-transduced cells (P=0.005 and P=0.07, respectively). Fluorescence-activated cell-sorting analysis of bone marrow cells from these mice showed that, as compared with wild-type Jak2, K539L-transduced cells resulted in expansion of the erythroid and granulocytic lineages but not those of T lymphocytes, B lymphocytes, or megakaryocytes (data not shown).

**DISCUSSION**

We have identified a distinctive myeloproliferative syndrome, associated with gain-of-function JAK2 exon 12 mutations, that includes patients who are currently given a diagnosis of polycythemia vera or idiopathic erythrocytosis. Patients with JAK2 exon 12 mutations present with erythrocytosis, low serum erythropoietin levels, and a distinctive histologic appearance of the bone marrow. As in other myeloproliferative diseases, erythropoietin-independent erythroid progenitors can be cultured from peripheral-blood cells, and cytogenetic abnormalities, splenomegaly, or transformation to myelofibrosis has been observed in some patients. Unlike erythroid colonies in patients with V617F-positive polycythemia vera, those in patients with exon 12 mutations are not homozygous for the JAK2 mutation.

The diagnosis of individual patients with a myeloproliferative disorder can be difficult. Different centers use different diagnostic criteria, and several diagnostic tests are not widely used. A patient may therefore be given a diagnosis of polycythemia vera by one clinician and a diagnosis of idiopathic erythrocytosis by another. Our results emphasize the importance of molecular classifi-

---

*Figure 4. A Myeloproliferative Phenotype, Resulting from Retroviral Expression of K539L Jak2, in a Murine Model of Bone Marrow Transplantation.*

Panel A shows the mean (±SD) hematocrit, white-cell count, and platelet count in the peripheral blood of BALB/c mouse recipients of bone marrow expressing wild-type, V617F, or K539L Jak2. Mice (five in each group) were evaluated 38 days after transplantation. P values are shown for the comparison with recipients of wild-type Jak2. Panels B, C, and D (hematoxylin and eosin) show representative peripheral-blood smears from mice 38 days after transplantation.
cation of these diseases. Exon 12 mutations may have previously been missed when peripheral-blood leukocyte DNA was analyzed, since granulocyte involvement in patients with these mutations is often low. For the molecular diagnosis of this syndrome, it is therefore important to sequence DNA from bone marrow cells or, preferably, from individual clonogenic hematopoietic colonies.

It is not clear how mutations that affect residues 537 through 543 result in unregulated JAK2 activity. To date, only the structure of the JAK2 kinase domain has been elucidated, and for this reason the details of interdomain interactions in JAK2 are unknown. However, homology-based molecular modeling suggests that residues 537 through 543 lie within a region linking the predicted SRC homology 2 (SH2) and JH2 domains of JAK2. These residues are near the predicted loop carrying V617 in a theoretical model of the full-length JAK2 protein (Fig. 3 in the Supplementary Appendix). Verification of this model awaits detailed structural and biochemical analysis.

Our results also shed light on the various clinical phenotypes associated with exon 12 and V617F mutations. Compared with the V617F mutation, exon 12 mutations result in stronger ligand-independent signaling through JAK2; exon 12 mutations generate higher levels of JAK2 and ERK1 and ERK2 phosphorylation than does the V617F mutation. Moreover, the absence of exon 12 mutations in patients with essential thrombocytemia accords with the proposal that low levels of JAK2 signaling favor thrombocytemia, whereas more-active signaling favors erythrocytosis.

Supported by grants from the U.K. Leukaemia Research Fund and the Wellcome Trust (to Dr. Green), the Leukemia and Lymphoma Society, the Doris Duke Charitable Foundation, and the Howard Hughes Medical Institute (to Dr. Gilliland), Amgen (to Dr. Lodish), the National Cancer Institute (K01 CA115679, to Dr. Tong), the National Institutes of Health (P01 HL32262, to Dr. Lodish, and DK50654 and CA66996, to Dr. Gilliland), and the American Society of Hematology and the Doris Duke Charitable Foundation (to Dr. Levine).

No potential conflict of interest relevant to this article was reported.

We thank Romano Kroemer for the coordinates of the JAK2 model; Melanie Percy, Betty Cheung, Anthony Bence, and the staff of the Addenbrooke’s Haematological Disorders Sample Bank for the processing of clinical samples; Sara Zarnegar for technical assistance; Brian Huntly for comments on the manuscript; Yana Pikman for assistance with transplant experiments; and Martha Wadleigh for providing clinical details.

REFERENCES

8. Yana Pikman for assistance with transplant experiments; and Martha Wadleigh for providing clinical details.


Copyright © 2007 Massachusetts Medical Society.
A Communication Strategy and Brochure for Relatives of Patients Dying in the ICU

Alexandre Lautrette, M.D., Michael Darmon, M.D., Bruno Megarbane, M.D., Ph.D., Luc Marie Joly, M.D., Sylvie Chevret, M.D., Ph.D., Christophe Adrie, M.D., Ph.D., Didier Barnoud, M.D., Gérard Bleichner, M.D., Cédric Bruel, M.D., Gérald Choukroun, M.D., J. Randall Curtis, M.D., M.P.H., Fabienne Fieux, M.D., Richard Galliot, M.D., Maité Garroutte-Orgeas, M.D., Hugues Georges, M.D., Dany Goldgran-Toledano, M.D., Mercé Jourdain, M.D., Ph.D., Georges Loubert, M.D., Jean Reignier, M.D., Fayçal Saidi, M.D., Bertrand Souweine, M.D., Ph.D., François Vincent, M.D., Nancy Kentish Barnes, Ph.D., Frédéric Pochard, M.D., Ph.D., Benoit Schlemmer, M.D., and Elie Azoulay, M.D., Ph.D.

BACKGROUND

There is a need for close communication with relatives of patients dying in the intensive care unit (ICU). We evaluated a format that included a proactive end-of-life conference and a brochure to see whether it could lessen the effects of bereavement.

METHODS

Family members of 126 patients dying in 22 ICUs in France were randomly assigned to the intervention format or to the customary end-of-life conference. Participants were interviewed by telephone 90 days after the death with the use of the Impact of Event Scale (IES; scores range from 0, indicating no symptoms, to 75, indicating severe symptoms related to post-traumatic stress disorder [PTSD]) and the Hospital Anxiety and Depression Scale (HADS; subscale scores range from 0, indicating no distress, to 21, indicating maximum distress).

RESULTS

Participants in the intervention group had longer conferences than those in the control group (median, 30 minutes [interquartile range, 19 to 45] vs. 20 minutes [interquartile range, 15 to 30]; P<0.001) and spent more of the time talking (median, 14 minutes [interquartile range, 8 to 20] vs. 5 minutes [interquartile range, 5 to 10]). On day 90, the 56 participants in the intervention group who responded to the telephone interview had a significantly lower median IES score than the 52 participants in the control group (27 vs. 39, P=0.02) and a lower prevalence of PTSD-related symptoms (45% vs. 69%, P=0.01). The median HADS score was also lower in the intervention group (11, vs. 17 in the control group; P=0.004), and symptoms of both anxiety and depression were less prevalent (anxiety, 45% vs. 67%; P=0.02; depression, 29% vs. 56%; P=0.003).

CONCLUSIONS

Providing relatives of patients who are dying in the ICU with a brochure on bereavement and using a proactive communication strategy that includes longer conferences and more time for family members to talk may lessen the burden of bereavement. (ClinicalTrials.gov number, NCT00331877.)
HAVING A LOVED ONE DIE IN THE INTENSIVE CARE UNIT (ICU) is an extraordinarily stressful event. The patient is usually unable to communicate with the family or with ICU staff. Qualitative and quantitative studies of families in this situation have identified effective communication between caregivers and families and support from caregivers throughout the decision-making process as important to family members.

In many ICUs, an end-of-life family conference, which is rooted in findings from epidemiologic and interventional studies on communicating with families of dying patients, is an important part of ICU practice. In these conferences, family members and ICU staff discuss the patient’s situation in a quiet room. Ideally, family members are given opportunities to ask questions, express concerns, and confront painful emotions with the help of caring, compassionate professionals.

Although the conference is important, the effect of its structure on bereaved family members has not been evaluated in a randomized trial. We conducted a multicenter, randomized, controlled study to evaluate the effect of a proactive communication strategy that consisted of an end-of-life family conference conducted according to specific guidelines and that concluded with the provision of a brochure on bereavement. We hypothesized that this intervention, as compared with the customary end-of-life conference, would decrease stress-related symptoms and symptoms of anxiety and depression in family members 90 days after the patient’s death.

**METHODS**

We conducted a prospective, randomized, controlled trial in 22 ICUs (Table 1) in France from May 2005 to October 2005. The study was approved by the institutional review board of the French Society for Critical Care, and oral informed consent was obtained from the participating families. At each ICU, one investigator was responsible for the study, which included six consecutive patients and their surrogates. On day 90, one mem-

| Table 1. Characteristics of the 22 ICUs in the Study. |
|---------------------------------|-----------------|
| Characteristic                  | Value           |
| Teaching hospital — no. (%)     | 15 (68)         |
| Type of ICU — no. (%)           |                 |
| Medical                         | 10 (45)         |
| Surgical                        | 3 (14)          |
| Medical and surgical            | 9 (41)          |
| No. of attending physicians — median (interquartile range) | 6 (5–6) |
| No. of residents — median (interquartile range) | 3 (3–4) |
| No. of patients per nurse — median (interquartile range) | 3 (3–3) |
| No. of beds — median (interquartile range) | 16 (12–21) |
| Rooms with more than two beds — no. (%) | 10 (45) |
| Regular (at least weekly) nurse–physician meetings — no. (%) | 19 (86) |
| Availability of bereavement brochure before study began — no. | 0 |
| Research group on end-of-life family care — no. (%) | 8 (36) |
| End-of-life family conferences held before study began — no. | 0 |
| Routine involvement of family members in daily care — no. (%) | 8 (36) |
| Routine involvement of family members in decisions — no. (%) | 8 (36) |
| No. of family–staff conflicts in 2004 — median (interquartile range) | 25 (12–41) |
| No. of visiting hours per day — median (interquartile range) | 4 (2–8) |
| Unrestricted visiting hours — no. (%) | 5 (23) |
| Psychologist present in ICU — no. (%) | 5 (23) |

* The research groups consisted of nurses and doctors who met weekly to discuss how to improve the quality of care.
### Table 2. Characteristics of Patients and Enrolled Family Members at Time of ICU Admission.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control Group (N = 63)</th>
<th>Intervention Group (N = 63)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age — yr</td>
<td></td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Median</td>
<td>68</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Interquartile range</td>
<td>56–76</td>
<td>56–80</td>
<td></td>
</tr>
<tr>
<td>Male sex — no. (%)</td>
<td>37 (59)</td>
<td>33 (52)</td>
<td>0.47</td>
</tr>
<tr>
<td>French descent — no. (%)</td>
<td>56 (89)</td>
<td>58 (92)</td>
<td>0.60</td>
</tr>
<tr>
<td>Unmarried — no. (%)</td>
<td>15 (24)</td>
<td>21 (33)</td>
<td>0.23</td>
</tr>
<tr>
<td>Direct admission to ICU — no. (%)</td>
<td>34 (54)</td>
<td>37 (59)</td>
<td>0.77</td>
</tr>
<tr>
<td>Coexisting conditions — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>13 (21)</td>
<td>13 (21)</td>
<td>0.99</td>
</tr>
<tr>
<td>Chronic heart failure</td>
<td>10 (16)</td>
<td>14 (22)</td>
<td>0.36</td>
</tr>
<tr>
<td>Cancer</td>
<td>21 (33)</td>
<td>12 (19)</td>
<td>0.10</td>
</tr>
<tr>
<td>Cirrhosis</td>
<td>2 (3)</td>
<td>5 (8)</td>
<td>0.24</td>
</tr>
<tr>
<td>Poor performance status — no. (%)</td>
<td>28 (44)</td>
<td>27 (43)</td>
<td>0.61</td>
</tr>
<tr>
<td>Reason for ICU admission — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute respiratory failure</td>
<td>28 (44)</td>
<td>27 (43)</td>
<td>0.85</td>
</tr>
<tr>
<td>Coma</td>
<td>27 (43)</td>
<td>25 (40)</td>
<td>0.71</td>
</tr>
<tr>
<td>Shock</td>
<td>21 (33)</td>
<td>24 (38)</td>
<td>0.57</td>
</tr>
<tr>
<td>Acute renal failure</td>
<td>11 (18)</td>
<td>14 (22)</td>
<td>0.50</td>
</tr>
<tr>
<td>Cardiac arrest</td>
<td>14 (22)</td>
<td>16 (25)</td>
<td>0.67</td>
</tr>
<tr>
<td>Simplified Acute Physiology Score — median</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(interquartile range)†</td>
<td>64 (52–76)</td>
<td>59 (52–81)</td>
<td>0.85</td>
</tr>
<tr>
<td>Treatment needed at end of life — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td>56 (89)</td>
<td>58 (92)</td>
<td>0.54</td>
</tr>
<tr>
<td>Vasopressors</td>
<td>42 (67)</td>
<td>49 (78)</td>
<td>0.23</td>
</tr>
<tr>
<td>Dialysis</td>
<td>16 (25)</td>
<td>14 (22)</td>
<td>0.67</td>
</tr>
<tr>
<td>Sedation</td>
<td>47 (75)</td>
<td>49 (78)</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>Family members‡</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male sex — no. (%)</td>
<td>12 (23)</td>
<td>17 (30)</td>
<td>0.39</td>
</tr>
<tr>
<td>Age — yr</td>
<td></td>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td>Median</td>
<td>54</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Interquartile range</td>
<td>46–64</td>
<td>47–58</td>
<td></td>
</tr>
<tr>
<td>French descent — no. (%)</td>
<td>46 (88)</td>
<td>48 (86)</td>
<td>0.35</td>
</tr>
<tr>
<td>Catholic — no. (%)</td>
<td>35 (67)</td>
<td>35 (63)</td>
<td>0.78</td>
</tr>
<tr>
<td>Married — no. (%)</td>
<td>24 (46)</td>
<td>22 (39)</td>
<td>0.57</td>
</tr>
<tr>
<td>Relationship to patient — no. (%)</td>
<td></td>
<td></td>
<td>0.45</td>
</tr>
<tr>
<td>Spouse</td>
<td>22 (42)</td>
<td>20 (36)</td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td>22 (42)</td>
<td>30 (54)</td>
<td></td>
</tr>
<tr>
<td>Parent</td>
<td>5 (10)</td>
<td>2 (4)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (6)</td>
<td>4 (7)</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates that the participants and their parents were born in France.
† Scores range from 0 to 163, with higher scores indicating more severe illness.
‡ Data are for the 52 family members in the control group and the 56 family members in the intervention group who were interviewed at 90 days.
number of each family — either the patient’s designated surrogate or the person who ranked highest in the hierarchy for surrogate decision making — was interviewed. Additional methodologic details are presented in the Supplementary Appendix, available with the full text of this article at www.nejm.org.

**PARTICIPANT SELECTION AND STUDY PROCEDURES**
The only criterion for inclusion in the study was the belief by the physician in charge that the patient would die within a few days. Patients younger than 18 years of age were excluded from the study, as were family members who had insufficient knowledge of French for a telephone interview. Table 2 lists characteristics of the patients and family members. Surrogates were assigned at random to the intervention or control group. In the control group, interactions between the family and the ICU staff, including the end-of-life conference, occurred according to the usual practice at each center. In the intervention group, the end-of-life family conference was held in accordance with detailed guidelines developed by

<table>
<thead>
<tr>
<th>Table 3. Implementation of the Intervention and End-of-Life Care, Including Decisions to Forgo Life-Sustaining Treatments.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Implementation of intervention</td>
</tr>
<tr>
<td>Family informed of decision to forgo life-sustaining treatment — no. (%)</td>
</tr>
<tr>
<td>More than one family member informed of decision — no. (%)</td>
</tr>
<tr>
<td>Involvement of family in decision — no. (%)</td>
</tr>
<tr>
<td>No involvement</td>
</tr>
<tr>
<td>Family members expressed patient’s wishes</td>
</tr>
<tr>
<td>Family members expressed their own wishes</td>
</tr>
<tr>
<td>End-of-life conference</td>
</tr>
<tr>
<td>No. of family members present</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Nurse present — no. (%)</td>
</tr>
<tr>
<td>No. of ICU physicians present</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Duration of conference — min</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Total time that family members spoke — min</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Total time that nurse spoke — min</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Clinicians’ observations — no. (%)</td>
</tr>
<tr>
<td>Family expressed guilt</td>
</tr>
<tr>
<td>Family reported successful expression of emotions</td>
</tr>
<tr>
<td>Family believed that patient’s symptoms were controlled</td>
</tr>
<tr>
<td>Family reported conflicts with ICU staff</td>
</tr>
</tbody>
</table>
one of the authors at the University of Washington.\textsuperscript{10,14,15} Families were given a brochure on bereavement (see the Supplementary Appendix for the original French version and a version translated into English by the authors). The end-of-life conference used in the intervention group had five objectives for the caregivers, summarized by the mnemonic VALUE\textsuperscript{10,14,15}: to value and appreciate what the family members said, to acknowledge the family members’ emotions, to listen, to ask questions that would allow the caregiver to understand who the patient was as a person, and to elicit questions from the family members. Each investigator received a detailed description of the conference procedure.\textsuperscript{10} Randomization was performed centrally in blocks of six, stratified according to the ICU, with group assignments sent in sealed envelopes to the study centers (for details see the Supplementary Appendix).

**OUTCOME MEASURES**

One family member per patient was interviewed over the telephone 90 days after the patient’s death; the interviews took place between August 2005 and January 2006. The primary outcome measure was the score on the Impact of Event Scale

<table>
<thead>
<tr>
<th>Table 3. (Continued.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td><strong>End-of-life care</strong></td>
</tr>
<tr>
<td>Decision to forgo life-sustaining treatments — no. (%)</td>
</tr>
<tr>
<td>No. of days from ICU admission to decision</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Nonbeneficial interventions after end-of-life conference — no. (%)</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
</tr>
<tr>
<td>Vasopressors</td>
</tr>
<tr>
<td>Dialysis</td>
</tr>
<tr>
<td>Other†</td>
</tr>
<tr>
<td>No. of nonbeneficial interventions provided after decision to forgo life-sustaining treatments</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Life-sustaining treatments withdrawn — no. (%)</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
</tr>
<tr>
<td>Vasopressors</td>
</tr>
<tr>
<td>Dialysis</td>
</tr>
<tr>
<td><strong>Other data</strong></td>
</tr>
<tr>
<td>No. of days from decision to forgo life-sustaining treatments to death</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>No. of days in ICU</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Interquartile range</td>
</tr>
<tr>
<td>Conflicts with family members reported by ICU staff — no. (%)</td>
</tr>
<tr>
<td>Patients who survived and were discharged — no. (%)</td>
</tr>
</tbody>
</table>

* The intervention began on the day that the end-of-life family conference was held.
† Other treatments were blood transfusions, antibiotics, and vitamins.
(IES), which assesses symptoms related to post-traumatic stress disorder (PTSD); scores range from 0 (no PTSD-related symptoms) to 75 (severe PTSD-related symptoms).\textsuperscript{5,16-18} We classified patients as having low or high IES scores, using 30 as the cutoff, in agreement with previous reports.\textsuperscript{5,18} Secondary outcome measures were symptoms of anxiety and depression, which we assessed using the Hospital Anxiety and Depression Scale (HADS); subscale scores range from 0 (no distress) to 21 (severe distress).\textsuperscript{19,20} HADS subscale scores above 8 were considered to indicate clinically significant symptoms of anxiety or depression.\textsuperscript{19}

**DATA COLLECTION**

Investigators recorded ICU and patient characteristics on standardized forms. The data elements included in Table 3 were gathered in a prospective fashion. In addition, a specific form was used to collect data describing the end-of-life family conference, and investigators were asked to clock family conference times. Primary-outcome data were collected by the interviewer 90 days after the patient’s death.

**STATISTICAL ANALYSIS**

On the basis of data from our previous study,\textsuperscript{2} we hypothesized that the intervention would decrease the risk of PTSD-related symptoms by 30%. To detect a significant difference between the two groups with a type I error of 0.05 and a power of 0.90, 100 families had to be recruited, 50 in each group. We decided to include 132 family members (66 in each group) to allow for families lost to follow-up on day 90 (up to 25%).\textsuperscript{5} Continuous variables were reported as medians and interquartile ranges, and categorical variables as proportions. Comparisons of continuous variables between the two randomized groups were performed with the Wilcoxon rank-sum test, whereas comparisons of categorical variables were performed with the Pearson chi-square test or Fisher’s exact test, as appropriate. All tests were two-sided, and \( P \) values of less than 0.05 were considered to indicate statistical significance. Statistical tests were performed with the SAS software package, version 9.1 (SAS Institute).

**RESULTS**

Of the 132 eligible family members, 126 were randomly assigned to a study group, and 108 (86%) were interviewed 3 months after the patient’s death (range, 90 to 104 days) (Fig. 1). Of the 22 ICUs in the study, 15 were in teaching hospitals, and 7 in general hospitals. In all the ICUs, nurses and physicians held regular meetings about end-of-life issues; however, only three ICUs had written procedures for delivering information to families of dying patients, and only five ICUs had unrestricted visiting hours. Before the study, none of the ICUs provided family members with written information about bereavement, and none were aware of the VALUE-based guidelines for end-of-life conferences. The characteristics of the patients at enrollment did not differ significantly between the two study groups. A decision to forgo life-sustaining treatment was made for all the study patients; at the time that the decision was implemented, 114 patients (90%) were receiving mechanical ventilation and 96 (76%) were deeply sedated, precluding meaningful communication between the patient and family.

A comparison of the characteristics of the end-of-life conferences in the two study groups provides a measure of the implementation of the intervention. The significant differences in the conduct of the conferences, shown in Table 3, suggest that the guidelines for the intervention conferences were followed.\textsuperscript{2,21}

Regarding the prespecified process-of-care measures listed in Table 3, although the length of stay in the ICU and in the hospital did not differ significantly between the intervention and control groups, there were fewer nonbeneficial interventions (continued life support after a decision to withhold or withdraw life-sustaining treatments) in the intervention group (see Fig. 1 of the Supplementary Appendix), and withdrawal of mechanical ventilation and vasopressors was more common in this group than in the control group. Among the relatives who initially disagreed with the ICU clinicians regarding decisions to forgo life-sustaining treatments, those in the intervention group were more likely to agree with the decisions eventually (six relatives in the intervention group vs. none in the control group, \( P = 0.02 \)). Among the family members in both groups, 96 (89%) reported that the amount of time spent providing information was sufficient, and 97 (90%) felt that the information was clear; 41 (38%) reported a desire for additional information that was not provided (Table 4). The proportions of family members who reported a desire...
for additional information, who received newly prescribed psychotropic drugs, and who expressed feelings of guilt were lower in the intervention group than in the control group. In addition, 95% of family members in the intervention group said they were able to express their emotions to the ICU team, as compared with only 75% of family members in the control group.

Regarding the prespecified main outcome variables recorded 90 days after the death of the patient (Table 4), the IES scores in the intervention group were lower than those in the control group (median score, 27 [interquartile range, 18 to 42] vs. 39 [interquartile range, 25 to 48]; P=0.02), indicating that 25 family members in the intervention group (45%) were at risk for PTSD as compared with 36 (69%) in the control group. Similarly, family members in the intervention group had significantly lower HADS scores than those in the control group (median score, 11 [interquartile range, 8 to 18] vs. 17 [interquartile range, 11 to 25]; P=0.004), with 25 family members (45%) reporting clinically significant symptoms of anxiety and 16 (29%) reporting clinically significant symptoms of depression, as compared with 35 (67%) and 29 (56%) in the control group, respectively (P=0.02 and P=0.003, respectively) (Fig. 2).

**DISCUSSION**

Over the past decade, epidemiologic studies have identified the specific needs of family members of dying patients, thereby allowing the development of proactive interventions that have improved communication with family members. End-of-life family conferences are rooted in the evidence provided by this literature, their main goals being to improve communication between ICU staff and family members and to assist families when difficult decisions need to be made.

In our multicenter, randomized study, we compared two end-of-life conference formats, one reflecting a proactive approach to communication and ending with the provision of a brochure on bereavement, and the other reflecting the typical approach used by each center. The proactive communication strategy decreased PTSD-related symptoms and symptoms of anxiety and depression among family members.

In the intervention group, ICU clinicians were...
asked to follow detailed published guidelines\textsuperscript{14,15} to ensure a uniform and effective change in their approach to communication. As compared with the control conferences, the intervention conferences were attended by a larger number of relatives and were associated with longer times spent delivering information and listening to relatives. The intervention conferences also provided family members with more opportunities to discuss the patient’s wishes, to express emotions, to alleviate feelings of guilt, and to understand the goals of care. Our finding that patients in the intervention group received fewer nonbeneficial treatments concurs with evidence of the efficacy of proactive strategies such as ethics consultation\textsuperscript{24} and early palliative-care consultation for dying patients in the ICU.\textsuperscript{25}

A bereavement brochure was given to the family at the end of the intervention conference. Previous studies by our research group showed that comprehension was markedly improved by simply delivering standardized written information for families.\textsuperscript{23} This experience prompted us to include a brochure in our proactive communication strategy. Furthermore, prior research suggests that multifaceted interventions are necessary to effect changes in clinicians’ behavior.\textsuperscript{26}

Our study has several limitations. First, it was performed in France, where the patient–physician relationship is perceived as more paternalistic than elsewhere,\textsuperscript{27} with physicians having final authority in decisions to forgo life-sustaining treatments.\textsuperscript{5} Nonetheless, the intervention used in our study was rooted in the international literature and is relevant to other countries.\textsuperscript{2} It might be argued that the gap between the intervention and the control groups was larger as a result of paternalistic attitudes in the control group, since this group replicated usual practice; if this view is correct, the magnitude of the beneficial effect of the intervention in France would be greater than could be expected in countries where shared decision making with family members is more firmly established. A strong argument against this view, however, is the fact that interactions with family members in the control group were similar to those reported in other European countries and in North America.\textsuperscript{14,15} Furthermore, the results of our intervention were consistent with those in earlier studies of proactive interventions.\textsuperscript{22,24,25} In addition, 22 centers participated in our study, further enhancing the generalizability of our findings.

Second, our only criterion for inclusion in the study was the belief on the part of the physician in charge that death was inevitable and that a decision to forgo life-sustaining treatment was in order. In some cases, however, patients in such

---

**Table 4.** Outcomes Assessed on Day 90.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control Group (N=52)</th>
<th>Intervention Group (N=56)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IES score</td>
<td></td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Median</td>
<td>39</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Interquartile range</td>
<td>25–48</td>
<td>18–42</td>
<td></td>
</tr>
<tr>
<td>Presence of PTSD-related symptoms (IES score &gt;30) — no. (%)</td>
<td>36 (69)</td>
<td>25 (45)</td>
<td>0.01</td>
</tr>
<tr>
<td>HADS score</td>
<td></td>
<td></td>
<td>0.004</td>
</tr>
<tr>
<td>Median</td>
<td>17</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Interquartile range</td>
<td>11–25</td>
<td>8–18</td>
<td></td>
</tr>
<tr>
<td>Symptoms of anxiety — no. (%)</td>
<td>35 (67)</td>
<td>25 (45)</td>
<td>0.02</td>
</tr>
<tr>
<td>Symptoms of depression — no. (%)</td>
<td>29 (56)</td>
<td>16 (29)</td>
<td>0.003</td>
</tr>
<tr>
<td>Saw a psychologist after death of patient — no. (%)</td>
<td>6 (12)</td>
<td>4 (7)</td>
<td>0.41</td>
</tr>
<tr>
<td>Received newly prescribed psychotropic drugs after death of patient — no. (%)</td>
<td>12 (23)</td>
<td>6 (11)</td>
<td>0.05</td>
</tr>
<tr>
<td>Effectiveness of overall information provided — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time allotted to provide information was sufficient</td>
<td>45 (87)</td>
<td>51 (91)</td>
<td>0.45</td>
</tr>
<tr>
<td>Information was clear</td>
<td>45 (87)</td>
<td>52 (93)</td>
<td>0.34</td>
</tr>
<tr>
<td>Additional information requested</td>
<td>24 (46)</td>
<td>17 (30)</td>
<td>0.05</td>
</tr>
</tbody>
</table>
circumstances survive. Conceivably, our intervention might have a negative effect on the family members of patients who survive, a situation that transpired only once in this study. Nevertheless, the possible negative effects of such an event must be compared with the negative effects of suboptimal communication on the much larger number of families whose relatives die.

Third, we did not determine how many families read the bereavement brochure or how those who did reacted to it. The multicenter design of the study and the fact that each ICU physician held only three intervention conferences did not allow us to evaluate the physicians' learning curve. Previous work has shown that even a brief course of training may improve communication skills. A study over time would be useful to determine whether benefits to the families increase as ICU physicians improve their communication skills.

Fourth, because we did not assess the HADS score before the critical illness or at the time of the patient’s death, we cannot be sure that the two groups of family members were not different at baseline. However, in a recent noninterventional study, we recorded the HADS score for family members 90 days after the patient’s discharge or death. The median score was 17 (interquartile range, 10 to 22), suggesting not only that symptoms of anxiety and depression were common and lasting but also that the proactive communication strategy we tested in the current study had positive effects.

Fifth, although the interviewer and the analyst were unaware of the group assignments, blinding of family members and ICU clinicians was not feasible. Consequently, we cannot exclude the possibility that the investigators believed strongly in the effectiveness of the intervention and that this may have influenced other interactions with family members.

Finally, the positive results of the current study might in theory indicate that in the control group, communication was less personalized and interactive than the norm. However, we believe that the characteristics of the control conferences (reported in Table 3) — notably, their longer duration, as compared with that in earlier work by our group (20 minutes vs. 10 minutes) — show that communication with families was as good as, or better than, the norm. In addition, the proportion of relatives who were satisfied with the information they received and the proportion who requested additional information indicate that the standard of care for providing information was met. The fact that the IES and HADS scores in the control group were similar to those in our previous studies argues against the possibility that the control conferences were substandard, as does the extensive experience acquired over the years by the ICU physicians in our study group.

In summary, a proactive strategy for routine end-of-life family conferences that included provision of a brochure on bereavement, as compared with customary practice, resulted in longer meetings in which families had more opportunities to speak and to express emotions, felt more supported in making difficult decisions, experienced more relief from guilt, and were more likely to accept realistic goals of care. The result of this strategy was a decrease in PTSD-related symptoms and symptoms of anxiety and depression 3 months after the patient’s death.
Supported by grants from Assistance Publique–Hôpitaux de Paris and the French Society for Critical Care Medicine (AOR01004). The study was performed on behalf of the Famirea Study Group. Dr. Curtis was supported by a grant from the National Institute of Nursing Research (RO1NR005226).

Dr. Azoulay reports receiving grant support from Pfizer. No other potential conflict of interest relevant to this article was reported.

We thank A. Wolfe, M.D., for help with the manuscript.

REFERENCES


(C) 2007 Massachusetts Medical Society.)
Brief Report

Prepubertal Gynecomastia Linked to Lavender and Tea Tree Oils

Derek V. Henley, Ph.D., Natasha Lipson, M.D., Kenneth S. Korach, Ph.D., and Clifford A. Bloch, M.D.

Summary

Most cases of male prepubertal gynecomastia are classified as idiopathic. We investigated possible causes of gynecomastia in three prepubertal boys who were otherwise healthy and had normal serum concentrations of endogenous steroids. In all three boys, gynecomastia coincided with the topical application of products that contained lavender and tea tree oils. Gynecomastia resolved in each patient shortly after the use of products containing these oils was discontinued. Furthermore, studies in human cell lines indicated that the two oils had estrogenic and antiandrogenic activities. We conclude that repeated topical exposure to lavender and tea tree oils probably caused prepubertal gynecomastia in these boys.

Gynecomastia is generally attributed to conditions that disrupt sex-steroid signaling pathways, resulting in increased or unopposed estrogen action on breast tissue. In contrast to gynecomastia in adolescent boys and men, prepubertal gynecomastia is rare and should always be considered pathological, prompting a search for a source of estrogen. Although hyperestrogenemia may be endogenous or exogenous in origin, most persons with prepubertal gynecomastia have normal serum concentrations of sex steroids, and an underlying cause is not identified. In such cases, possible exposure to exogenous sources of estrogen should be considered. We investigated the cause of prepubertal gynecomastia in three otherwise healthy boys with normal serum concentrations of endogenous steroids.

Case Reports

Patient 1
A boy who was 4 years 5 months old presented with gynecomastia of apparently 2 to 3 weeks' duration. He had no exposure to any known exogenous form of estrogens (ingestants, salves, or ointments). His height and weight were at the 97th percentile and between the 75th and 90th percentiles, respectively. He had bilateral gynecomastia with firm, nontender breast tissue measuring 2 cm by 2 cm in diameter. His testes were 3 ml in volume and of normal consistency. His genitalia were prepubertal (Tanner stage 1). Laboratory investigation showed normal thyroid function; the follicle-stimulating hormone (FSH) concentration was 1.04 IU per liter (reference range, 0.25 to 1.92), luteinizing hormone 0.47 IU per liter (reference range, 0.02 to 1.03), testosterone 0.08 ng per milliliter (0.27 nmol per liter) (reference range, 0.02 to 0.25 ng per milliliter), estradiol less than 20 pg per milliliter (73 pmol per liter)
per liter) (normal value, <20), dehydroepiandrosterone (DHEA) sulfate less than 5.0 μg per deciliter (0.14 μmol per liter) (reference range, 1 to 40), 17-alpha-hydroxyprogesterone 0.32 μg per liter (0.97 nmol per liter) (reference range, 0.2 to 0.8), and prolactin 8.0 μg per liter (reference range, 2 to 29); the serum biochemistry values, including liver-function tests, were normal. On evaluation 3 months later, the breast buds were tender to palpation and had increased to 2.5 cm by 2.5 cm in diameter with an increased breast mound. The patient’s mother reported applying a compounded “healing balm” containing lavender oil to his skin starting shortly before the initial presentation. The gynecomastia partially resolved within 4 months after application of the healing balm was discontinued, at which time the breast buds measured 1.5 cm by 1.5 cm in diameter and were soft in consistency. Several months later, his pediatrician stated that the gynecomastia had resolved completely.

**Patient 2**

A boy who was 10 years 1 month old presented with a 5-month history of gynecomastia. He and his mother reported that the condition seemed more prominent in the evening and a little less so in the morning. His medical history and family history were unremarkable. His height and weight were above the 97th percentile, and his body-mass index (the weight in kilograms divided by the square of the height in meters) was 21.1. He had firm, tender breast buds, measuring 3.5 cm by 4.0 cm in length and width and approximately 3.5 cm in depth, with stretching of the areolae. His testes were 3 ml in volume and of normal consistency. His pubic hair was Tanner stage 2 (a small amount of long hair at the base of the scrotum), and his genitalia were Tanner stage 1. Laboratory testing showed a testosterone concentration of 0.36 ng per milliliter (1.25 nmol per liter) (normal value, <0.25), free testosterone 0.0066 ng per milliliter (0.0229 nmol per liter) (reference range, 0.0006 to 0.0057), and DHEA sulfate 278 μg per deciliter (7.6 μmol per liter) (normal value, <75). On questioning, it was determined that the patient was not currently using drugs, herbal supplements, or herbal lotions but was applying a styling gel to his hair and scalp every morning and regularly using a shampoo. The labels of both the gel and the shampoo listed Lavandula angustifolia (lavender) oil and Melaleuca alternifolia (tea tree) oil as ingredients. Re-evaluation 9 months after use of these products was discontinued showed that his areolar mounds had decreased in depth to approximately 1 cm with almost no palpable glandular tissue.

**Patient 3**

A boy who was 7 years 10 months old presented with a 1-month history of gynecomastia that had appeared gradually. His height was between the 75th and 90th percentiles, and his weight was at the 50th percentile. He had bilateral gynecomastia with firm, nontender breast tissue that corresponded to Tanner stage 2. His testes were 3 ml in volume and of normal consistency. His genitalia were Tanner stage 1, and there was no pubic hair present. Laboratory testing showed normal thyroid function, FSH 0.49 IU per liter (reference range, 0.25 to 1.92), luteinizing hormone 0.16 IU per liter (reference range, 0.02 to 1.03), estradiol 5 pg per milliliter (18 pmol per liter) (normal value, <10), estril less than 0.1 ng per milliliter (0.3 nmol per liter) (normal value, <0.1), estrone less than 13 pg per milliliter (48 pmol per liter) (normal value, <13), total estrogens 61 pg per milliliter (225 pmol per liter) (normal value, <130 in adult men), DHEA sulfate 22 μg per deciliter (0.6 μmol per liter) (normal value, <130 in adult men), 17-alpha-hydroxyprogesterone 0.13 μg per liter (0.39 nmol per liter) (reference range, 15 to 65), and free beta subunit of human chorionic gonadotropin less than 2 mIU per milliliter (normal value, <5); the serum biochemistry values, including liver-function tests, were normal. His history was positive for the use of lavender-scented soap and intermittent use of lavender-scented commercial skin lotions. The gynecomastia resolved completely a few months after use of scented soap and skin lotions was discontinued (personal communication from the patient’s family). His fraternal twin used the same skin lotions, but not the lavender-scented soap, and did not have any gynecomastia.

**METHODS**

**MAMMALIAN CELL CULTURE**

Human breast-cancer (MCF-7) cells that express estrogen receptors were grown in phenol red-free Dulbecco’s modified Eagle’s medium containing 10% fetal-calf serum (Atlanta Biologicals), penicillin (100 U per milliliter), and streptomycin (100 μg
Figure 1. Estrogenic Activity of Lavender and Tea Tree Oils in Human Breast-Cancer (MCF-7) Cells.

MCF-7 cells were transiently transfected with both the estrogen-inducible 3X-ERE-TATA-luciferase (firefly) plasmid and the constitutively active renilla luciferase reporter plasmid (Promega) and treated for 18 hours with increasing concentrations of lavender oil (Panel A) and tea tree oil (Panel B) in the presence or absence of 1 μM fulvestrant. Treatment with 1 nM 17β-estradiol served as a positive control for activation of the reporter plasmid. The firefly luciferase activity was normalized to that of renilla luciferase activity and the total protein content for each sample. The results are expressed as the average (±SE) fold increase relative to the control solvent of the values obtained from independent experiments (five experiments in Panel A and four in Panel B), each conducted in duplicate. The dashed line at the top represents treatment of the cells with estradiol alone, and the dashed line at the bottom represents treatment of the cells with estradiol in the presence of fulvestrant. In Panels A and B, for the comparison between treatment with estradiol, lavender oil (at 0.005%, 0.01%, and 0.025%), or tea tree oil (at 0.005%, 0.01%, and 0.025%) and treatment with ethanol (the solvent control) alone, P<0.001. For the comparison between treatment with tea tree oil alone (at 0.001%) and treatment with ethanol, P<0.01. For the comparison between treatment with estradiol and treatment with estradiol plus fulvestrant, P<0.01 in Panel A and P<0.001 in Panel B. For the comparison between treatment with lavender oil or tea tree oil and treatment with either of the two oils plus fulvestrant, P<0.001. MCF-7 cells were treated for 2, 6, 12, or 18 hours with dimethylsulfoxide, 0.025% (vol/vol) lavender oil, 0.025% (vol/vol) tea tree oil, or 1 nM 17β-estradiol in the presence or absence of 1 μM fulvestrant (Panel C). Real-time PCR was performed to measure the steady-state mRNA levels of MYC, CTSD, and IGFBP3. The data shown represent a single time point corresponding to the maximum 17β-estradiol–induced expression of each gene (MYC, 2 hr; CTSD, 18 hr; and IGFBP3, 6 hr). All values were normalized to glyceraldehyde-3-phosphate dehydrogenase, and each data point represents the average increase relative to the vehicle control of the values obtained from four independent experiments. In Panel C, for the comparison between treatment with estradiol, lavender oil, or tea tree oil and treatment at the same point in time with ethanol, P<0.05. For CTSD, for the comparison between treatment with estradiol and treatment at the same point in time with ethanol, P<0.05. For IGFBP3, for the comparison between treatment with lavender oil or tea tree oil and the identical treatment plus fulvestrant at the same point in time, P<0.05. For IGFBP3, for the comparison between treatment with lavender oil or tea tree oil and the identical treatment plus fulvestrant at the same point in time, P=0.056.
per milliliter). Human breast-cancer (MDA-kb2) cells that express the androgen receptor were maintained as previously described.\(^4\) Cell-culture reagents were obtained from Invitrogen Life Technologies, unless otherwise indicated. For all experiments, the lavender oil (\textit{L. officinalis}, which is a synonym for \textit{L. angustifolia}) and tea tree oil (\textit{M. alternifolia}) (both from Sigma Chemical) were diluted in dimethylsulfoxide before they were added to culture media.

**Luciferase Assays and Reverse-Transcriptase and Real-Time Polymerase-Chain-Reaction Analysis**

MCF-7 and MDA-kb2 cells were assayed for luciferase activity with the use of the Dual–Luciferase reporter assay system (Promega) and an LMAX II\(^{384}\) luminometer (Molecular Devices). Total RNA was isolated from MCF-7 cells and MDA-kb2 cells with the use of the RNeasy Mini Kit (Qiagen), according to the manufacturer’s protocol. Synthesis of complementary DNA (cDNA) and analyses of gene-specific cDNA concentrations were performed by real-time polymerase chain reaction (PCR), as previously described.\(^5\) The PCR primers were designed with the use of Primer Express software, version 2.0 (Applied Biosystems) (see the Supplementary Appendix, available with the full text of this article at www.nejm.org).

**Statistical Analysis**

The data were analyzed for statistical significance by the Mann–Whitney nonparametric test.

**RESULTS**

**Estrogen-Receptor–Dependent Estrogenic Activity In Vitro**

To determine whether lavender oil and tea tree oil are estrogenic, we performed dose–response experiments in MCF-7 cells that were positive for estrogen receptors and were transiently transfected with an estrogen-inducible luciferase reporter plasmid containing three copies of an estrogen-response element (3X-ERE-TATA-luciferase). Both oils stimulate ERE-dependent luciferase activity in a dose-dependent manner, with the maximum activity observed at 0.025% volume per volume (vol/vol) for each oil, corresponding to approximately 50% of the activity elicited by 1 nM 17β-estradiol (Fig. 1A and 1B). Treatment with higher doses of the oils was cytotoxic. The pure estrogen-receptor antagonist fulvestrant inhibited transactivation of the 3X-ERE-TATA-luciferase reporter plasmid by both oils, indicating that their activity is estrogen-receptor–dependent (Fig. 1A and 1B). Additional experiments indicated that lavender oil was able to transactivate the estrogen-inducible reporter plasmid in estrogen-receptor–negative SK-BR-3 human breast-cancer cells only after simultaneous transfection with an estrogen-receptor–expression vector (data not shown).

Further experiments in MCF-7 cells indicated that the two oils modulated the expression of the estrogen-regulated endogenous genes MYC (also called \textit{C-MYC}),\(^6\) CTSD,\(^7\) and IGFBP3.\(^8\) Lavender oil and tea tree oil increased the expression of messenger RNA (mRNA) for MYC and CTSD and decreased the expression of mRNA for IGFBP3, as compared with the dimethylsulfoxide controls, in a manner that was similar to the effect of 1 nM 17β-estradiol on the magnitude and timing of the
responses (Fig. 1C). These responses were attenuated in the presence of 1 μM fulvestrant (Fig. 1C).

**In Vitro Antiandrogenic Activity**
To evaluate the potential androgenic properties of lavender oil and tea tree oil, we performed dose–response experiments in MDA-kb2 cells, a line of human breast-cancer cells that are positive for the androgen receptor and were stably transfected with an androgen-inducible and glucocorticoid-inducible mouse mammary-tumor virus (MMTV)-luciferase reporter plasmid. Treatment
of MDA-kb2 cells with the androgen-receptor agonist dihydrotestosterone (DHT) at 0.1 nM, the lowest observed effective dose in this cell line,\(^4\) resulted in an increase in luciferase activity that was almost four times higher than that in the dimethylsulfoxide controls (Fig. 2A and 2B). In contrast, neither lavender oil nor tea tree oil trans-activated the MMTV-luciferase reporter plasmid at any concentration tested (Fig. 2A and 2B).

The antiandrogenic properties of the two oils were assessed by simultaneously treating the MDA-kb2 cells with DHT and increasing the concentration of lavender oil or tea tree oil. The androgen-receptor antagonist flutamide was also included in these assays, as a positive control for androgen-receptor antagonism. Transactivation of the MMTV-luciferase reporter plasmid by 0.1 nM DHT was inhibited in a concentration-dependent manner by both lavender oil and tea tree oil, as well as by flutamide (Fig. 2A and 2B). Maximum inhibition occurred at 0.005% vol/vol for both lavender oil and tea tree oil, corresponding to a decrease in luciferase activity of 52% and 41%, respectively, in the presence of 0.1 nM DHT. The observed inhibitory effects appear to be specific to the androgen receptor, since neither of the two oils attenuated the glucocorticoid-receptor–mediated transactivation of the MMTV-luciferase reporter plasmid in the presence of 5 nM dexamethasone, the lowest observed effective dose in this cell line\(^9\) (data not shown). Further experiments in MDA-kb2 cells indicated that the antiandrogenic properties of lavender oil and tea tree oil extended to inhibition of DHT-stimulated expression of the androgen-inducible endogenous genes CYP1B1, CYP1B1, and CYP1B1, and UGT2B28, and SEC14L2\(^10\) (Fig. 2C). The antiandrogenic effects of the two oils are not caused by down-regulation of the expression of the androgen receptor, since neither of the oils altered the amount of androgen-receptor mRNA or protein in these experiments (data not shown).

**DISCUSSION**

In contrast to gynecomastia, which occurs in more than 60% of boys during puberty, prepubertal gynecomastia is extremely uncommon. Since there is no known physiologic cause of prepubertal gynecomastia, pathologic causes should be considered. However, a specific cause is rarely identified, and in 90% of patients, prepubertal gynecomastia is labeled idiopathic.\(^2,3\) In such patients, the condition may be caused by exposure to an environmental chemical that disrupts the endocrine system and leads to disproportionate estrogen and androgen pathway signaling, a finding reported in a limited number of adults with gynecomastia.\(^11,12\)

In this report, we describe three otherwise healthy boys with prepubertal gynecomastia, all of whom had normal serum concentrations of endogenous steroids and none of whom had been exposed to any known exogenous endocrine disruptor such as medications, oral contraceptives, marijuana, or soy products.\(^1\) The repeated topical application of one or more over-the-counter personal care products that contained lavender oil or lavender oil and tea tree oil was documented for all three patients. Case 1 provided the clinical clue to lavender oil as a potential source, because it was the only topically applied agent used by that child. Use of lavender oil was considered trivial by the child’s mother, who acknowledged its use only after repeated questioning. In Case 2, the boy had biochemical evidence of physiologic adrenarche, but the evidence was unrelated to his gynecomastia, which resolved after discontinuation of the use of products containing lavender oil and tea tree oil, despite the persistence of adrenarche. The daily temporal fluctuation in the severity of the gynecomastia reported by the patient’s mother might have been caused by the transdermal absorption kinetics of the oils after application each morning. In Case 3, the patient was exposed intermittently to various over-the-counter personal-care products containing lavender oil. His twin brother used the same lotions but not the scented soap, and gynecomastia did not develop in him.

The common use of products containing lavender oil, tea tree oil, or both by the three boys and the resolution of their gynecomastia within months after ceasing use of those products suggest that these oils may possess endocrine-disrupting activity that causes an imbalance in estrogen and androgen pathway signaling. Other components in these products may also possess endocrine-disrupting activity that contributed to the gynecomastia, but those components were not tested because we chose to evaluate only the component that was found in all the products used by the patients (lavender oil) and a chemically similar component that was found in some of the products (tea tree oil).
Our in vitro studies confirm that lavender oil and tea tree oil possess weak estrogenic and antiandrogenic activities that may contribute to an imbalance in estrogen and androgen pathway signaling. Estrogenic or antiandrogenic activities have been reported for other essential oils and some of their monoterpene constituents.13-18 On the basis of the three case reports and the in vitro studies, we suspect that repeated topical application of over-the-counter products containing lavender oil or tea tree oil was the cause of gynecomastia in the three patients.

This report raises an issue of concern, since lavender oil and tea tree oil are sold over the counter in their “pure” form and are present in an increasing number of commercial products, including shampoos, hair gels, soaps, and body lotions. Whether the oils elicit similar endocrine-disrupting effects in prepubertal girls, adolescent girls, or women is unknown. Since gynecomastia is labeled idiopathic in approximately 10% of men, one might speculate that unidentified exogenous sources of endocrine-disrupting chemicals may contribute to the onset or progression of the condition, or both, in such patients.1 The results of our in vitro studies indicate a dose–response relationship in the estrogenic and antiandrogenic activities of lavender oil and tea tree oil, suggesting that susceptibility to gynecomastia or other manifestations of endocrine disruption may require exposure to a threshold dose of these oils. The threshold might depend on several undefined factors, including the concentration of the oil in a product; the duration, frequency, and quantity of use of the product; and the genetic characteristics of persons exposed. Until epidemiologic studies are performed to determine the prevalence of gynecomastia associated with exposure to lavender oil and tea tree oil, we suggest that the medical community should be aware of the possibility of endocrine disruption and should caution patients about repeated exposure to any products containing these oils.

Supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences (NIEHS).

Dr. Bloch reports receiving grant support and lecture fees from Eli Lilly, Genentech, Novo Nordisk, Pfizer, Tercica, and Serono. No other potential conflict of interest relevant to this article was reported.

We thank the laboratory of Dr. Donald P. McDonnell for the 3X-ERE-TATA-luciferase plasmid; the laboratory of Dr. L.E. Gray, Jr., for the MDA-kb2 cell line; Dr. Grace Kissling (NIEHS) for performing statistical analyses; Drs. Donna Baird, John F. Couse, and Walter J. Rogan for critical review of the manuscript; and Ms. Sue Edelstein for graphical assistance.

REFERENCES

Public Reporting and Pay for Performance in Hospital Quality Improvement

Peter K. Lindenauer, M.D., M.Sc., Denise Remus, Ph.D., R.N., Sheila Roman, M.D., M.P.H., Michael B. Rothberg, M.D., M.P.H., Evan M. Benjamin, M.D., Allen Ma, Ph.D., and Dale W. Bratzler, D.O., M.P.H.

ABSTRACT

BACKGROUND
Public reporting and pay for performance are intended to accelerate improvements in hospital care, yet little is known about the benefits of these methods of providing incentives for improving care.

METHODS
We measured changes in adherence to 10 individual and 4 composite measures of quality over a period of 2 years at 613 hospitals that voluntarily reported information about the quality of care through a national public-reporting initiative, including 207 facilities that simultaneously participated in a pay-for-performance demonstration project funded by the Centers for Medicare and Medicaid Services; we then compared the pay-for-performance hospitals with the 406 hospitals with public reporting only (control hospitals). We used multivariable modeling to estimate the improvement attributable to financial incentives after adjusting for baseline performance and other hospital characteristics.

RESULTS
As compared with the control group, pay-for-performance hospitals showed greater improvement in all composite measures of quality, including measures of care for heart failure, acute myocardial infarction, and pneumonia and a composite of 10 measures. Baseline performance was inversely associated with improvement; in pay-for-performance hospitals, the improvement in the composite of all 10 measures was 16.1% for hospitals in the lowest quintile of baseline performance and 1.9% for those in the highest quintile (P<0.001). After adjustments were made for differences in baseline performance and other hospital characteristics, pay for performance was associated with improvements ranging from 2.6 to 4.1% over the 2-year period.

CONCLUSIONS
Hospitals engaged in both public reporting and pay for performance achieved modestly greater improvements in quality than did hospitals engaged only in public reporting. Additional research is required to determine whether different incentives would stimulate more improvement and whether the benefits of these programs outweigh their costs.
The need to improve both the quality and the safety of health care in the United States is well documented.\textsuperscript{1-5} Traditional strategies to stimulate improvement include regulation, measurement of performance and subsequent feedback, and marketplace competition.\textsuperscript{6} Despite limited evidence, public reporting of hospital quality data and pay for performance have emerged as two of the most widely advocated strategies for accelerating quality improvement.\textsuperscript{7-11} Public reporting stimulates interest in quality on the part of physicians and hospital leaders, perhaps by appealing to their professional ethos.\textsuperscript{12} Pay-for-performance programs are intended to strengthen the business case for quality improvement by rewarding excellence and reversing what have been described as perverse financial incentives that can deter hospitals from investing in quality-improvement efforts.\textsuperscript{9,13,14} In enacting the Deficit Reduction Act of 2005, Congress demonstrated its support for financial incentives by calling on the Centers for Medicare and Medicaid Services (CMS) to develop a plan for hospital “value based purchasing” by 2009.\textsuperscript{15}

Despite the instinctive appeal of pay for performance and public reporting, little is known about the individual or combined benefits of such programs,\textsuperscript{12,16,17} and both are the subject of ongoing debate.\textsuperscript{18-23} In order to determine the incremental effect of pay for performance, we measured improvements in hospital quality that occurred when financial incentives were combined with public reporting and compared these improvements with gains associated with public reporting alone.

\begin{center}
\textbf{METHODS}
\end{center}

\textbf{HOSPITAL QUALITY ALLIANCE}

In December 2002, the American Hospital Association, the Federation of American Hospitals, and the Association of American Medical Colleges launched the Hospital Quality Alliance (HQA), a national public–private collaboration to encourage hospitals to collect and report data regarding the quality of care on a voluntary basis.\textsuperscript{24} The HQA was designed to provide information about the quality of hospital care to the public and to “invigorate efforts to improve quality.” All acute care hospitals in the United States were invited to participate, and by linking participation in the program to the annual Medicare payment update, the CMS was able to achieve participation rates of more than 98%. Participating hospitals were required to collect and report data on a minimum of 10 quality measures regarding three clinical conditions: heart failure, acute myocardial infarction, and pneumonia (Table 1). Hospitals began submitting data in the fourth quarter of 2003, and this information was made available on the Hospital Compare Web site.\textsuperscript{25} In order to provide stable rate estimates, data from hospitals that submitted information on fewer than 25 cases were not included in the analysis.

\begin{table}[h!]
\centering
\begin{tabular}{|l|}
\hline
\textbf{Acute myocardial infarction}  \\
Percentage of patients who were given aspirin on arrival  \\
Percentage of patients who were given an ACE inhibitor or ARB for left ventricular systolic dysfunction  \\
Percentage of patients for whom aspirin was prescribed at discharge  \\
Percentage of patients who were given a beta-blocker on arrival  \\
Percentage of patients for whom a beta-blocker was prescribed at discharge  \\
\hline
\textbf{Heart failure}  \\
Percentage of patients who were assessed for left ventricular function  \\
Percentage of patients who were given an ACE inhibitor or ARB for left ventricular systolic dysfunction  \\
\hline
\textbf{Pneumonia}  \\
Percentage of patients who were assessed for oxygenation  \\
Percentage of patients who were given initial antibiotics within 4 hours after arrival  \\
Percentage of patients who were assessed and given pneumococcal vaccination  \\
\hline
\end{tabular}
\caption{Quality Measures Shared by the Hospital Quality Alliance and Hospital Quality Incentive Demonstration.*}
\end{table}

\* ACE denotes angiotensin-converting enzyme, and ARB angiotensin-receptor blocker.
HOSPITAL QUALITY INCENTIVE DEMONSTRATION

In March 2003, hospitals subscribing to a quality-benchmarking database, known as Perspective, which is maintained by Premier Healthcare Informatics, were invited to participate in the CMS–Premier Hospital Quality Incentive Demonstration (HQID), a multiyear collaborative whose goal was “to determine if economic incentives are effective at improving the quality of inpatient care.” Hospitals that accepted the invitation collected and submitted data on 33 quality measures regarding five clinical conditions: heart failure, acute myocardial infarction, community-acquired pneumonia, coronary-artery bypass grafting, and hip and knee replacement. This set of conditions included the previously described 10 measures reported on the Hospital Compare Web site. The remaining 23 measures are described elsewhere.

Hospitals needed to have a minimum of 30 cases per condition annually to be eligible for the demonstration. For each of the clinical conditions, hospitals performing in the top decile on a composite measure of quality for a given year received a 2% bonus payment in addition to the usual Medicare reimbursement rate. Hospitals in the second decile received a 1% bonus. Bonuses averaged $71,960 per year and ranged from $914 to $847,227. These additional payments are anticipated to be partially offset by financial penalties ranging from 1 to 2% of Medicare payments for hospitals that by the end of the third year of the program had failed to exceed the performance of hospitals in the lowest two deciles, as established during the program’s first year.

Of 421 hospitals that were invited to participate, 266 (63%) initially accepted, 155 declined, and 11 later withdrew. In several instances, multiple hospitals shared the same Medicare provider number. These multihospital organizations submit billing claims and clinical quality data to the CMS as a single entity and were treated as a single hospital for the purpose of our analysis. The demonstration project began in the fourth quarter of 2003 and continued through the third quarter of 2006. Informed consent and institutional review board approval were not required because the data were collected for administration of the Medicare program, not for research, and access to these data is provided to the program by law. All the authors assume full responsibility for the accuracy and completeness of the data presented.

STATISTICAL ANALYSIS

The overlapping reporting requirements between the HQA and HQID allowed us to compare improvements in quality associated with public reporting with those achieved when financial incentives are combined with public reporting. Hospitals were eligible for our analysis if they participated in the HQA program and submitted data on a minimum of 30 cases for a single condition annually, including at least 8 cases in both the fourth quarter of 2003 and the third quarter of 2005. In our primary analyses, we matched each HQID participant with as many as two HQA hospitals on the basis of the number of beds (matched to within five beds), teaching status (teaching or nonteaching), region (Northeast, Midwest, South, or West), location (urban or rural), and ownership status (not-for-profit or for-profit). These analyses focused on HQID hospitals that had participated in the program throughout the entire 2-year study period. From the pool of HQA hospitals available for matching, we excluded those that had either declined participation in the HQID or started the demonstration and then withdrew, since these decisions may have reflected doubts about whether the hospitals would be successful or other confounding factors.

We treated the matched sets as the primary units of analysis. We calculated the change in adherence to each of the HQA quality measures over a period of eight quarters for each hospital. We then calculated the difference in the improvement for each HQA quality measure for pay-for-performance hospitals, as compared with the control group. In sets with two control hospitals, the improvements at the two facilities were averaged. A paired t-test was performed to evaluate the difference in improvement between pay-for-performance and control hospitals. In addition, we calculated the percentage change in adherence to two sets of compound measures for each of the clinical conditions. First, we calculated a “composite process score” by adding up the total number of opportunities for each condition for which correct care was provided and dividing this result by the sum of the number of correct care opportunities. Using this same approach, we calculated a summary composite score that com-
bined all 10 individual measures. Second, we created an “appropriate care measure” by calculating the percentage of patients who received all recommended interventions for a given clinical condition. As compared with composite process measures, appropriate care measures may better represent the interests and likely desires of patients, are more sensitive to subtle improvements, and can help foster a system perspective in quality measurement.

We performed a series of stratified analyses to evaluate the effects of baseline performance, teaching status, and number of beds on the response to these incentives and compared the improvement of pay-for-performance hospitals with that of control hospitals. To estimate the incremental effect of financial incentives, multiple linear regression was applied to the matched sample. The dependent variable in these analyses was the difference in improvement between pay-for-performance hospitals and control hospitals for each matched set. We controlled for baseline hospital performance and diagnosis-specific hospital volume. The four composite process measures were used for all stratified and multivariable analyses.

To provide additional validation of our results, we conducted another multiple linear regression using the entire set of HQA participants, not only those identified through matching, with the individual hospital as the unit of analysis. In this regression, the dependent variable was the improvement over the 2-year study period, and we adjusted for baseline hospital performance, diagnosis-specific hospital volume, and all other available hospital characteristics.

To evaluate the potential contribution of a “volunteer bias” among the pay-for-performance group, we repeated our multivariable analysis by grouping hospitals that had either declined participation in the HQID or had withdrawn, together with those that had accepted and had completed the 2 years. Finally, we repeated our multiple linear regression using the entire set of HQA participants, with the hospital as the unit of analysis, and added an interaction term to explore whether the effect of pay for performance varied across quintiles of baseline performance. All analyses were carried out with the use of SAS software, version 9.1 (SAS Institute). P values of less than 0.05 were considered to indicate statistical significance.

**RESULTS**

Of the 4691 hospitals that submitted data for the HQA between the fourth quarter of 2003 and the third quarter of 2005, 2490 met our enrollment criteria, including 266 participants in the HQID. Eleven hospitals withdrew from the HQID during the first 2 years, leaving 255 pay-for-performance hospitals eligible for our primary analysis. We successfully matched 207 of these 255 HQID hospitals with 406 HQA controls, including 199 with two matches and 8 with one match. The typical hospital included in the study was a small-to-mid-size, nonteaching, not-for-profit facility serving an urban population in the South (Table 2). As compared with all hospitals participating in the HQA, study hospitals were larger, less likely to have for-profit ownership, more likely to be urban, and more likely to have house staff. Hospitals that declined participation in the HQID were on average smaller, more rural, and less engaged in housestaff training (see Table 2A of the Supplementary Appendix, available with the full text of this article at www.nejm.org).

**IMPROVEMENTS IN QUALITY**

Over the 2-year study period, both pay-for-performance hospitals and control hospitals showed evidence of improvement in each of the individual and compound measures of performance (Table 3 and Fig. 1). Pay-for-performance hospitals showed significantly greater improvement than did control hospitals in 7 of the 10 individual measures of performance, with absolute differences in improvement ranging from 0.6% for oxygen assessment among patients with pneumonia (P<0.001) to 10.9% for vaccination among patients with pneumonia (P<0.001) (Table 3). Pay-for-performance hospitals also achieved greater improvement in all the composite process measures, with differences ranging from 4.1% for pneumonia (P<0.001) to 5.2% for heart failure (P<0.001). For each of the conditions, differences in the composite measures of performance between the two hospital groups increased throughout the 2-year study period (Fig. 1). A similar pattern was observed for the appropriate care measures (i.e., percentages of patients who received all recommended treatments for the condition), with absolute differences in changes ranging from 6.0% for heart failure (P<0.001) to 7.5% for acute myocardial infarction (P<0.001) (Table 3). Hospitals that declined to participate in the pay-for-per-
formance demonstration improved less than did participating hospitals (Table 3A of the Supplementary Appendix).

Stratified analyses showed an inverse relationship between baseline performance and improvement in both groups of hospitals (Table 4). This factor influenced comparisons on the basis of hospital size and teaching status (Table 4A of the Supplementary Appendix). The difference in improvement between pay-for-performance hospitals and control hospitals varied with baseline performance, ranging from 1.2% for the composite measure of care for heart failure among hospitals with the highest baseline performance to 9.6% for the same measure among hospitals with the poorest baseline performance (Table 4).

After adjustment for the effects of baseline performance and for differences in baseline performance and condition-specific volumes between pay-for-performance hospitals and control hospitals, the incremental effect of financial incentives decreased to 2.6% for the composite process measure of acute myocardial infarction (P<0.001) and 4.1% for heart failure (P<0.001) (Table 5). A second multivariable analysis, which included the entire pool of HQA participants (with adjustment for baseline performance, condition-specific volume, and all hospital characteristics), yielded similar findings. A third analysis, intended to account for a volunteer effect by including hospitals that declined to participate in pay for performance, showed a persistent, albeit smaller, effect of financial incentives (Table 5). A final multivariable analysis, which accounted for the effects of baseline performance and other hospital characteristics, showed that the effect of financial incentives varied according to baseline performance for the composite measure of care for heart failure, with the largest improvements observed among hospitals with the poorest baseline performance. In contrast, for the composite measures of acute myocardial infarction and pneumonia

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Hospitals with Pay for Performance plus Public Reporting (N=207)</th>
<th>Hospitals with Public Reporting Only (Control Group) (N=406)</th>
<th>All Hospitals with Public Reporting (N=2490)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of beds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200</td>
<td>63 (30.4)</td>
<td>128 (31.5)</td>
<td>1355 (54.4)</td>
</tr>
<tr>
<td>200–400</td>
<td>77 (37.2)</td>
<td>146 (36.0)</td>
<td>697 (28.0)</td>
</tr>
<tr>
<td>&gt;400</td>
<td>67 (32.4)</td>
<td>132 (32.5)</td>
<td>438 (17.6)</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>36 (17.4)</td>
<td>69 (17.0)</td>
<td>873 (35.1)</td>
</tr>
<tr>
<td>Urban</td>
<td>171 (82.6)</td>
<td>337 (83.0)</td>
<td>1617 (64.9)</td>
</tr>
<tr>
<td>Teaching status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td>97 (46.9)</td>
<td>191 (47.0)</td>
<td>783 (31.4)</td>
</tr>
<tr>
<td>Nonteaching</td>
<td>110 (53.1)</td>
<td>215 (53.0)</td>
<td>1707 (68.6)</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>45 (21.7)</td>
<td>90 (22.2)</td>
<td>662 (26.6)</td>
</tr>
<tr>
<td>Northeast</td>
<td>28 (13.5)</td>
<td>54 (13.3)</td>
<td>417 (16.7)</td>
</tr>
<tr>
<td>South</td>
<td>109 (52.7)</td>
<td>212 (52.2)</td>
<td>1032 (41.4)</td>
</tr>
<tr>
<td>West</td>
<td>25 (12.1)</td>
<td>50 (12.3)</td>
<td>379 (15.2)</td>
</tr>
<tr>
<td>Ownership</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-profit</td>
<td>4 (1.9)</td>
<td>7 (1.7)</td>
<td>461 (18.5)</td>
</tr>
<tr>
<td>Not-for-profit</td>
<td>203 (98.1)</td>
<td>399 (98.3)</td>
<td>2029 (81.5)</td>
</tr>
</tbody>
</table>

* Listed are all hospitals that submitted performance data on a minimum of 30 cases per year and 8 cases in the fourth quarter of 2003 and the third quarter of 2005. Percentages may not total 100 because of rounding.
### Table 3. Improvements in Quality over a 2-Year Period among Hospitals Engaged in Pay for Performance and Public Reporting.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hospitals with Pay for Performance plus Public Reporting</th>
<th>Hospitals with Public Reporting Only (Control Group)</th>
<th>Absolute Difference between Hospital Groups</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4th Quarter, 2003</td>
<td>3rd Quarter, 2005</td>
<td>Absolute Change</td>
<td>4th Quarter, 2003</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>no. of patients (%)</td>
<td>%</td>
<td></td>
<td>no. of patients (%)</td>
</tr>
<tr>
<td>Aspirin on arrival</td>
<td>8,293 (92.2)</td>
<td>6,776 (96.3)</td>
<td>4.1</td>
<td>12,780 (94.9)</td>
</tr>
<tr>
<td>Aspirin at discharge</td>
<td>10,204 (93.9)</td>
<td>8,358 (96.1)</td>
<td>2.2</td>
<td>16,660 (92.7)</td>
</tr>
<tr>
<td>ACE inhibitor for LVSD</td>
<td>2,016 (78.2)</td>
<td>1,809 (88.9)</td>
<td>10.7</td>
<td>3,129 (82.9)</td>
</tr>
<tr>
<td>Beta-blocker on arrival</td>
<td>10,122 (89.1)</td>
<td>8,607 (95.9)</td>
<td>6.8</td>
<td>16,524 (90.4)</td>
</tr>
<tr>
<td>Beta-blocker at discharge</td>
<td>7,224 (87.6)</td>
<td>5,345 (93.5)</td>
<td>5.9</td>
<td>11,187 (89.6)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>no. of patients (%)</td>
<td>%</td>
<td></td>
<td>no. of patients (%)</td>
</tr>
<tr>
<td>LV assessment</td>
<td>17,301 (83.2)</td>
<td>14,933 (93.3)</td>
<td>10.1</td>
<td>25,431 (85.1)</td>
</tr>
<tr>
<td>ACE inhibitor for LVSD</td>
<td>5,728 (77.7)</td>
<td>5,970 (86.3)</td>
<td>8.6</td>
<td>8,606 (77.6)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>no. of patients (%)</td>
<td>%</td>
<td></td>
<td>no. of patients (%)</td>
</tr>
<tr>
<td>Antibiotic timing</td>
<td>17,280 (68.2)</td>
<td>7,877 (79.9)</td>
<td>11.7</td>
<td>25,944 (68.9)</td>
</tr>
<tr>
<td>Vaccination</td>
<td>9,357 (42.7)</td>
<td>5,515 (72.7)</td>
<td>30.0</td>
<td>13,599 (44.6)</td>
</tr>
<tr>
<td>Oxygen assessment</td>
<td>18,523 (97.9)</td>
<td>9,757 (99.6)</td>
<td>1.7</td>
<td>28,077 (98.3)</td>
</tr>
<tr>
<td>Appropriate care measures</td>
<td>no. of patients (%)</td>
<td>%</td>
<td></td>
<td>no. of patients (%)</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>13,187 (77.0)</td>
<td>11,193 (88.6)</td>
<td>11.6</td>
<td>21,665 (81.6)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>17,569 (76.1)</td>
<td>15,264 (88.3)</td>
<td>12.2</td>
<td>25,578 (77.8)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>18,531 (48.5)</td>
<td>9,758 (70.6)</td>
<td>22.1</td>
<td>28,077 (50.0)</td>
</tr>
<tr>
<td>Composite process scores</td>
<td>no. of patients (%)</td>
<td>%</td>
<td></td>
<td>no. of patients (%)</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>39,543 (88.7)</td>
<td>32,247 (94.8)</td>
<td>6.1</td>
<td>62,944 (91.3)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>23,656 (81.2)</td>
<td>21,588 (91.5)</td>
<td>10.3</td>
<td>34,829 (82.9)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>46,339 (75.2)</td>
<td>23,724 (86.4)</td>
<td>11.2</td>
<td>70,363 (76.2)</td>
</tr>
<tr>
<td>All 10 measures</td>
<td>116,613 (81.0)</td>
<td>96,695 (90.5)</td>
<td>9.5</td>
<td>192,381 (82.9)</td>
</tr>
</tbody>
</table>

* ACE denotes angiotensin-converting enzyme, LVSD left ventricular systolic dysfunction, and LV left ventricular.
† Student’s t-test was performed to compare the absolute difference in the change between hospitals with pay for performance plus public reporting and those with public reporting only (control hospitals).
and all 10 measures combined, estimates of the improvement attributable to financial incentives were similar, regardless of baseline performance.

**Discussion**

Public reporting and pay for performance are two of the most important methods that have been proposed to close persistent gaps in the quality and safety of health care. To evaluate whether combining pay for performance with public reporting results in more improvement than public reporting alone, we took advantage of a natural experiment involving several thousand hospitals engaged in a national public-reporting initiative, with more than 200 simultaneously participating in a pay-for-performance demonstration. We found that hospitals that were offered a 1 to 2% bonus for achieving high levels of performance relative to their peers had greater improvements in quality over a 2-year period than did those receiving no financial incentives. After adjustment for differences in baseline performance and other characteristics between the two groups of hospitals,
<table>
<thead>
<tr>
<th>Composite Process Score</th>
<th>Hospitals with Pay for Performance plus Public Reporting</th>
<th>Hospitals with Public Reporting Only (Control Group)</th>
<th>Absolute Difference between Hospital Groups</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. of patients (%)</td>
<td>%</td>
<td>no. of patients (%)</td>
<td>%</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 5</td>
<td>5,678 (72.9)</td>
<td>17.8</td>
<td>9,728 (79.5)</td>
<td>10.3</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>6,658 (86.5)</td>
<td>6.7</td>
<td>13,212 (89.5)</td>
<td>2.6</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>9,596 (91.1)</td>
<td>3.3</td>
<td>17,055 (93.2)</td>
<td>1.5</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>9,218 (94.0)</td>
<td>2.9</td>
<td>19,805 (95.8)</td>
<td>1.1</td>
</tr>
<tr>
<td>Quintile 1</td>
<td>9,934 (97.9)</td>
<td>-1.1</td>
<td>12,800 (98.7)</td>
<td>-3.4</td>
</tr>
<tr>
<td>Heart failure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 5</td>
<td>2,412 (62.6)</td>
<td>25.2</td>
<td>6,038 (67.0)</td>
<td>15.6</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>5,202 (79.0)</td>
<td>12.3</td>
<td>7,902 (79.7)</td>
<td>8.6</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>5,564 (83.6)</td>
<td>9.6</td>
<td>9,012 (85.2)</td>
<td>3.4</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>5,395 (88.1)</td>
<td>3.4</td>
<td>8,019 (89.3)</td>
<td>0.5</td>
</tr>
<tr>
<td>Quintile 1</td>
<td>6,003 (93.7)</td>
<td>-0.1</td>
<td>9,632 (95.2)</td>
<td>-1.3</td>
</tr>
<tr>
<td>Pneumonia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 5</td>
<td>11,532 (62.8)</td>
<td>19.6</td>
<td>16,292 (64.4)</td>
<td>13.1</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>10,856 (70.3)</td>
<td>13.7</td>
<td>17,449 (70.7)</td>
<td>9.4</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>10,742 (74.9)</td>
<td>10.2</td>
<td>17,566 (75.5)</td>
<td>7.8</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>9,113 (79.8)</td>
<td>9.3</td>
<td>15,257 (80.3)</td>
<td>4.9</td>
</tr>
<tr>
<td>Quintile 1</td>
<td>8,536 (88.4)</td>
<td>3.4</td>
<td>12,281 (87.9)</td>
<td>1.1</td>
</tr>
<tr>
<td>Composite of 10 measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quintile 5</td>
<td>19,625 (69.7)</td>
<td>16.1</td>
<td>26,297 (70.5)</td>
<td>13.2</td>
</tr>
<tr>
<td>Quintile 4</td>
<td>21,969 (77.5)</td>
<td>13.1</td>
<td>39,711 (79.3)</td>
<td>7.8</td>
</tr>
<tr>
<td>Quintile 3</td>
<td>23,513 (81.3)</td>
<td>10.6</td>
<td>45,420 (82.9)</td>
<td>5.3</td>
</tr>
<tr>
<td>Quintile 2</td>
<td>26,984 (85.4)</td>
<td>6.1</td>
<td>45,170 (87.5)</td>
<td>3.1</td>
</tr>
<tr>
<td>Quintile 1</td>
<td>24,522 (91.2)</td>
<td>1.9</td>
<td>35,783 (94.2)</td>
<td>-3.0</td>
</tr>
</tbody>
</table>

* Student’s t-test was performed to compare the absolute difference in the change between hospitals with pay for performance plus public reporting and those with public reporting only (control hospitals) for each quintile.

† Analysis of variance was performed in each category to compare the change from the fourth quarter of 2003 to the third quarter of 2005 for hospitals with pay for performance plus public reporting and those with public reporting only (control hospitals). P<0.001 for all comparisons.
the incremental effect of financial incentives was reduced, amounting to 2.6 to 4.1% over a period of 2 years.

Why are these findings important? Although the effect of the incentives was modest, our results suggest that financial incentives are capable of catalyzing quality-improvement efforts among hospitals already engaged in public reporting. And although the lion’s share of bonus payments were made to hospitals with the highest baseline performance, participants across the entire spectrum responded similarly, perhaps equally motivated by the desire to avoid financial penalties.

However, before widespread application of financial incentives is considered, it should be acknowledged that pay for performance is more complex than public reporting in several ways. First, unless new money is infused into the payment system or savings are identified from improvements in quality, the size of any bonuses will need to be balanced by reductions in reimbursements across the entire system or to underperforming hospitals, creating significant concern about the possibility of harm to safety-net institutions. Second, complex and politically charged judgments need to be made about fundamental system design. For example, should bonuses be paid to top-performing hospitals, to those with the greatest improvements, or to all those that meet a performance threshold? Third, the costs of administering pay-for-performance programs are likely to be higher than those for public-reporting programs. With these issues in mind, it will be important to determine not simply whether the addition of pay for performance results in more improvement than public reporting alone, but whether the benefits of such a program are worth the added cost and complexity.

Little is known about the effect of public reporting on the quality of care. In one well-documented case, rates of death after coronary bypass surgery in New York State were observed to fall after hospital-specific rates became public, but the mechanisms through which this occurred have been debated. Similarly, according to a report on the QualityCounts program, run by the Employer Health Care Alliance Cooperative, hospitals with public-reporting programs engaged in more quality-improvement activities and were more likely to have improved outcomes than were controls.

---

Table 5. Estimates of Incremental Effect of Pay for Performance.*

<table>
<thead>
<tr>
<th>Analytic Approach</th>
<th>Incremental Effect of Pay for Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acute Myocardial Infarction</td>
</tr>
<tr>
<td>Matched for hospital characteristics</td>
<td>4.3 (2.5–6.1)</td>
</tr>
<tr>
<td>Matched for hospital characteristics and adjusted for baseline performance and condition-specific volume†</td>
<td>2.6 (1.3–3.9)</td>
</tr>
<tr>
<td>Unmatched and adjusted for baseline performance, condition-specific volume, and all hospital characteristics‡</td>
<td>1.9 (0.8–3.1)§</td>
</tr>
<tr>
<td>Unmatched and adjusted for baseline performance, condition-specific volume, and all hospital characteristics; hospitals that declined participation in pay for performance included and grouped with those that agreed to participate¶</td>
<td>1.8 (0.9–2.8)</td>
</tr>
</tbody>
</table>

* P<0.001 for all categories, unless otherwise noted.
† Multiple linear regression of matched pairs was adjusted for baseline performance and condition-specific hospital volume.
‡ Multiple linear regression of data for 2490 hospitals that engaged in pay for performance and public reporting was adjusted for hospital size, teaching status, region, location, ownership status, baseline performance, and condition-specific volume.
§ P=0.002.
¶ Multiple linear regression of data for 2490 hospitals that engaged in pay for performance and public reporting was adjusted for hospital size, teaching status, region, location, ownership status, baseline performance, and condition-specific volume; hospitals that declined to participate or withdrew from the Hospital Quality Incentive Demonstration were added to the pay-for-performance group to attempt to account for a volunteer effect.
Even less is known about the effect of pay for performance on quality or outcomes. In a recent analysis, Rosenthal et al., showed that offering financial incentives to physician groups produced little gains in measures of the quality of ambulatory care and largely rewarded groups with high performance at baseline.

Our study has a number of limitations. First, it is unclear how either pay for performance or public reporting alone would have compared with no reporting, and previous studies have noted improvements over time associated with other quality-improvement efforts. Second, hospitals that were involved in the pay-for-performance demonstration differed from the entire pool of HQA applicants, and our findings should be generalized with caution. Third, baseline performance for 5 of the 10 measures approached or exceeded 90%, thereby limiting our power to detect differences between the two groups. Fourth, although the HQID involved 33 measures spread across 5 conditions and procedures, we assessed the effect of financial incentives only on the 10 conditions that were shared by the HQA. Fifth, our attempt to adjust for volunteer bias may have underestimated the true effect of pay for performance. Sixth, the financial incentives offered to hospitals were modest, and larger bonuses might have led to more sizable improvements in quality. Finally, only the top 20% of participants were eligible for financial rewards, and interest in the program might decline over time, especially among hospitals that consistently fail to garner bonus payments. A choice to use alternative strategies, such as incentives for threshold achievement or absolute or relative improvement, might have led to different outcomes.

In conclusion, financial incentives can modestly increase improvements in quality among hospitals already engaged in public reporting. Additional research is required to determine whether larger incentives or the restructuring of payment models can stimulate more meaningful improvements and to evaluate whether the benefits of these programs outweigh their costs.

Supported by a contract (500-02-OK-03) from the CMS.

No potential conflict of interest relevant to this article was reported.

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the Department of Health and Human Services.

We thank Robert Wachter, M.D., for his thoughtful comments on an earlier version of the manuscript, and Penelope Pekow, Ph.D., for her guidance with biostatistics.

REFERENCES

Public Reporting and Pay for Performance in Hospital Quality Improvement


Nolan T, Berwick DM. All-or-none measurement raises the bar on performance. JAMA 2006;295:1168-70.


Copyright © 2007 Massachusetts Medical Society.
Inhaled Insulin for Diabetes Mellitus

Graham T. McMahon, M.D., M.M.Sc., and Ronald A. Arky, M.D.

This Journal feature begins with a case vignette that includes a therapeutic recommendation. A discussion of the clinical problem and the mechanism of benefit of this form of therapy follows. Major clinical studies, the clinical use of this therapy, and potential adverse effects are reviewed. Relevant formal guidelines, if they exist, are presented. The article ends with the authors’ clinical recommendations.

A 52-year-old man with an 8-year history of type 2 diabetes mellitus visits his primary care provider for advice. His glucometer readings at home have been high despite treatment with a sulfonylurea, a thiazolidinedione, and metformin at maximal doses. He has never smoked. His glycated hemoglobin value is 8.6% and his fasting blood glucose concentration ranges between 170 and 220 mg per deciliter (9.4 and 12.2 mmol per liter). His blood pressure, weight, and lipid profile are within recommended target ranges. The patient and his physician discuss therapeutic options and agree that insulin treatment should be initiated. The physician wonders whether the patient might benefit from inhaled insulin and refers him to an endocrinologist for evaluation.

The Clinical Problem

Diabetes mellitus, a major cause of illness and death across the globe, is responsible for a growing proportion of national health care expenditures. Insulin treatment is necessary for a substantial minority of patients with diabetes; more than 5 million Americans take insulin injections every day. A wide range of subcutaneous insulins are available, many administered with penlike delivery devices and ultrafine needles that enhance the comfort and convenience of insulin treatment. However, surveys indicate substantial resistance to insulin therapy on the part of both patients with type 2 diabetes who are not taking insulin and clinicians who care for such patients; the reasons for this resistance include anticipated pain, inconvenience, fear of hypoglycemia, and concern about weight gain. True insulin and needle phobias are uncommon, although many patients appear to avoid insulin injections and blood glucose testing because of anxiety. The youngest and oldest patients are least likely to accept injectable therapy and thus pose the greatest challenge for physicians who want to initiate insulin treatment. Although resistance can be mitigated through education, efforts to develop oral, nasal, and inhaled formulations of insulin have been driven by the preference of patients to avoid subcutaneous injections.

Pathophysiology and Effect of Therapy

Insulin is lifesaving for patients with type 1 diabetes, a disease characterized by beta-cell failure and insulin deficiency. Type 2 diabetes, by contrast, is characterized by defects in both insulin secretion and insulin action, with insulin deficiency usually emerging later in the course of the disease. Insulin supplementation is often required to attain good glycemic control in type 2 diabetes and is typically initiated if the glycated hemoglobin level is not in the target range despite treatment with a combination of oral hypoglycemic agents.
Most proteins and peptides used for systemic therapeutic purposes, including insulin, have high molecular weights and are hydrophilic; as a result, the only suitable means of administration has been injection.\textsuperscript{15} However, inhalation devices can now facilitate delivery of drugs to the lungs. Since the lung is a large microvascular organ, molecules that are formulated to reach the alveoli can gain access to the systemic circulation.\textsuperscript{15,16} Effective distribution in the lung requires particles that have an aerodynamic diameter between 1 and 5 \( \mu \text{m} \).\textsuperscript{15,16}

Many inhaled medications do not require a high degree of precision in dosing, and portable devices for inhaled drug delivery may be characterized by considerable dose-to-dose variation because of differences in inhalational flow rates. These devices are unsuitable for the administration of drugs such as insulin, for which dose consistency is critical.\textsuperscript{17} The development of suitable inhalation devices has therefore been a limiting factor in the production of a reliable, clinically useful form of inhaled insulin.

So far, the only device for insulin inhalation that has been approved by the Food and Drug Administration (FDA) is an inhaler that delivers a dry-powder formulation of human insulin produced by means of recombinant DNA technology (Exubera, Pfizer). After oral inhalation of a single dose of human insulin by means of this device, approximately 40\% of the dose reaches the deep lung, and 10\% of the total dose is bioavailable.\textsuperscript{18-20} The amount of drug that is delivered to the oropharynx or swallowed is unlikely to have a clinical effect.\textsuperscript{20}

The interval between the administration of insulin and the onset of glucose-lowering activity is shorter with inhaled insulin (10 to 20 minutes) than with subcutaneously administered soluble (regular) human insulin and is similar to the interval with subcutaneously administered rapid-acting insulin analogues such as aspart, glulisine, and lispro. These pharmacokinetic features make inhaled insulin a suitable agent for preprandial administration. Its duration of action is between that of the rapidly acting insulin analogues and that of regular human insulin.\textsuperscript{20-22}

\textbf{Clinical Evidence}

Inhaled insulin has been compared with subcutaneous insulin regimens in patients with type 1 diabetes and in those with type 2 disease and has been compared with oral hypoglycemic agents in patients with type 2 diabetes.\textsuperscript{23} All these trials were open label; most lasted for less than 6 months, and more than 90\% of the participants were white.\textsuperscript{23,24}

Among patients with type 1 or type 2 diabetes who received either a combination of neutral protamine Hagedorn (NPH) and regular insulin two to three times daily or a combination of ultralente each night and inhaled insulin before each meal, the glycated hemoglobin level at 6 months did not differ significantly between the two treatment groups. Patients who received ultralente and inhaled insulin had slightly lower rates of hypoglycemia.\textsuperscript{25,26}

Adding thrice-daily inhaled insulin to existing oral therapy is generally more effective over a 12-to-24-week period than adding a second oral hypoglycemic drug taken once or twice a day.\textsuperscript{27-29} However, as compared with oral agents for diabetes, inhaled insulin is consistently associated with a significantly higher incidence of hypoglycemic events.\textsuperscript{23,27-30}

In clinical trials, patients have been generally more satisfied with inhaled insulin than with subcutaneous insulin.\textsuperscript{25,26,31,32} Whether this outcome will be borne out in clinical practice remains to be determined.

\textbf{Clinical Use}

The FDA and the European Medicines Agency have both approved the Exubera inhalation delivery system for the preprandial treatment of patients with type 1 or 2 diabetes.\textsuperscript{18,33} Therefore, most of the available information regarding the use of inhaled insulin is based on studies of this agent. Several other manufacturers have preparations of inhaled insulin that are being evaluated in clinical trials but have not yet been approved.

Because of its rapid onset of activity, inhaled insulin is suitable for preprandial but not for basal use. Patients with diabetes that is suboptimally controlled with the use of oral agents alone can usually be successfully treated at the outset by adding a single subcutaneous dose of either NPH or glargine insulin that is given before bedtime and titrated to a target fasting glucose level of approximately 100 mg per deciliter (5.5 mmol per liter).\textsuperscript{34} Patients who comply with such an ap-
proach and whose glycated hemoglobin levels remain above target levels while they are receiving a basal insulin benefit from additional preprandial insulin therapy. Preprandial insulins such as inhaled insulin are therefore most suitable for patients with glycated hemoglobin levels that remain elevated after fasting glucose levels have been controlled with a basal insulin.

Inhaled insulin therapy may be especially useful for patients with a true needle phobia and those with extensive cutaneous lipodystrophy at injection sites, although the incidence of the latter problem is declining. Inhaled insulin is not approved for use in pregnant women, children, or adolescents.

Smoking is a contraindication to the use of inhaled insulin; active smoking significantly increases the rate and extent of insulin absorption. In contrast, passive exposure to tobacco smoke in nonsmokers decreases the rate and extent of insulin absorption. Clinicians should therefore exercise caution if they are prescribing inhaled insulin for patients who work or live in a smoky environment.

The use of inhaled insulin in patients with underlying lung disease such as asthma or chronic obstructive pulmonary disease is not recommended, since the absorption of insulin in these patients can be unpredictable, particularly when they are also using an inhaled bronchodilator. A simple upper respiratory tract infection may be less problematic: according to the manufacturer of Exubera, an experimental rhinovirus infection did not change the absorption of inhaled insulin. There are no data regarding the effect of more severe respiratory tract infections, such as pneumonia, on the absorption of inhaled insulin. Nevertheless, it is prudent for patients initiating treatment with inhaled insulin to be trained in the use and receive a supply of subcutaneous insulin for situations in which pulmonary absorption might not be reliable.

All candidates for inhaled insulin therapy should be taught how to check their glucose level before meals. They should also undergo spirometry, and the drug should not be used if the forced expiratory volume in 1 second (FEV₁) is below 70% of the predicted value. Measurement of the diffusing capacity for carbon monoxide is not mandatory but can provide a useful baseline for monitoring changes in pulmonary function over time.

With the Exubera inhalational device, the administration of the dose and the inhalation are separated into two steps (see the video in the Supplementary Appendix, available with the full text of this article at www.nejm.org). When a dose of insulin is required, the patient extends the chamber and places a single blister of powdered insulin into a slot in the front of the device (Fig. 1). The patient creates a compressed volume of air by squeezing the pneumatic handle. Once the device is activated, the powder is released into a visible cloud, where it is suspended in a small volume of air that can be inhaled. A 5-second breath-hold allows the drug to settle in the lungs.

The dose of inhaled insulin is measured in milligrams rather than in units. The manufacturer's guidelines suggest that the initial estimate of the appropriate premeal dose should be 0.05 mg per kilogram of body weight. Thus, a person who weighs 100 kg should take 5 mg of inhaled insulin before each meal. However, unlike subcutaneous insulins, inhaled insulin is currently available in only two fixed doses (1 mg and 3 mg, approximately equivalent to 3 units and 8 units of insulin, respectively). Since only one blister can be used at each inhalation, multiple inhalations before each meal are necessary if the required dose of insulin is not exactly 1 mg or 3 mg. Furthermore, the received dose varies depending on the combination of blisters used. Consecutive inhalation of insulin from three blisters containing 1 mg of insulin apiece causes a 30 to 40% higher insulin exposure than inhalation of insulin from one blister containing 3 mg of insulin. Therefore, patients should not replace a single 3-mg dose with three consecutive 1-mg doses.

Patient education regarding the use of inhaled insulin is critical to maximize the consistency of technique and dose delivery. Maintenance of the inhaler is also essential. The device must be cleaned weekly and allowed to air dry, since moisture in the chamber absorbs the insulin powder. In addition, an internal valve (included with each box of insulin blister packs) must be replaced every 2 weeks; this step requires manual dexterity.

Follow-up should include spirometry at 6 months and then every year because of the potential effect of inhaled insulin on pulmonary function. If the FEV₁ is confirmed to have declined by more than 20% or by more than 500 ml from the baseline value, inhaled insulin should be discontinued indefinitely.
Inhaled insulin is more expensive than other mealtime insulin. The average monthly cost of inhaled insulin in the amount recommended for a 100-kg patient is approximately $112. In comparison, the average monthly wholesale cost for a similar dose of injectable insulin is $33 for regular insulin, $76 for a rapid-acting insulin analogue, and $102 for a rapid-acting insulin analogue in a penlike delivery device. Many managed-care organizations offer limited coverage for inhaled insulin, placing it in a tier of medications that require preapproval, higher patient copayments, or both.

### Adverse Effects

Two studies involving patients with type 1 diabetes and one study involving patients with type 2 diabetes showed a lower overall incidence of hypoglycemia among patients who received inhaled insulin than among those who received injected regular insulin. However, two of these trials showed an increased incidence of severe hypoglycemia among the patients who received inhaled insulin. The rate of hypoglycemia after the use of the Exubera device has not been compared with that associated with the alternative preprandial insulins (aspart, glulisine, or lispro) in head-to-head trials.

Diabetes is associated with abnormal lung function. Inhaled insulin has small additional effects on both the diffusing capacity for carbon monoxide and the FEV₁, suggesting effects on the alveolar-capillary membrane and lung elastic recoil, respectively; it is not clear whether these effects are correlated. However, the FEV₁ declined by more than 15% from the baseline value in 1.3% of patients with type 1 diabetes who received inhaled insulin and in 5% of patients with type 2 diabetes who received inhaled insulin. This loss of lung function appeared to resolve within 6 weeks of discontinuation of inhaled insulin after up to 2 years of treatment. It is not known whether these changes in pulmonary function can be predicted on the basis of cough or dyspnea; cough has frequently been reported in clinical trials of inhaled insulin.

### Areas of Uncertainty

Insulin acts as a weak growth factor when it binds to the type 1 insulin-like growth factor receptor. Short-term studies in animals have not shown a substantial effect on cell-proliferation indexes in the alveolar or bronchiolar areas of the lung. The long-term effects of supraphysiologic doses of insulin in the human lung or on neoplastic lung tissue are unknown.

Insulin antibody levels rise progressively with the increased duration of exposure to inhaled insulin in patients with type 1 or type 2 diabetes. These levels stabilize within 9 to 12 months after the start of treatment and decline but do not normalize after cessation of treatment. Antibody levels are especially elevated among patients with type 1 diabetes, increasing by
more than a factor of 8 after 6 months of the use of inhaled insulin. The frequency of severe hypoglycemia and the onset or duration of insulin activity have not been shown to be altered in the presence of insulin antibodies, but further study is required to confirm that these antibodies do not act as a reservoir for delayed insulin release.

Studies have suggested that patients with diabetes are likely to prefer inhaled insulin over insulin injection, in some cases by a ratio of 8:1. It is not clear whether any increases in patient preference, acceptability, or satisfaction will be translated into increased compliance and improved glucose control. Managed-care companies and patients will need to decide whether they are willing to pay the additional price for this alternative insulin delivery system. Other inhaled insulin systems are in various stages of development and will need to be compared with the Exubera inhalation device. Finally, the longer-term safety and efficacy of this form of therapy have not yet been established.

GUIDELINES

In the United Kingdom, the National Institute for Health and Clinical Excellence recommends that inhaled insulin be prescribed only by diabetes specialists and for patients with needle phobia or severe problems at injection sites. The German Institute for Quality and Efficiency in Health Care has concluded that inhaled insulin offered no additional benefit over subcutaneously administered insulin.

No guidelines for the use of inhaled insulin have been developed by expert groups or societies in the United States.

RECOMMENDATIONS

The patient described in the vignette presents with circumstances that are typical of many persons for whom insulin therapy is recommended. Although the concept of inhaled insulin is likely to be attractive to many such patients, we would first target the fasting glucose before introducing a preprandial insulin. After appropriate education and with the necessary support in place, we would begin treatment with a basal insulin given before sleep, adjusting the dose to achieve a mean fasting glucose level of approximately 100 mg per deciliter. Thus, we do not recommend the use of inhaled insulin in this patient. Should the patient later require preprandial insulin, the freedom from subcutaneous injection offered by inhaled insulin should be weighed against the necessity for multiple inhalations (sometimes at each dose), added cost, limited portability, risk of hypoglycemia, and unknown long-term adverse effects of this form of therapy.

No potential conflict of interest relevant to this article was reported.

We thank Christopher H. Fanta, M.D., for helpful comments.

A video showing the use of inhaled insulin is available with the full text of this article at www.nejm.org.

Copyright © 2007 Massachusetts Medical Society.
A Medical Mystery: Dilated Bowel

A 70-year-old man presented with a history of increasing abdominal distention. On physical examination, the patient was afebrile and his abdomen was soft and nontender. An abdominal radiograph raised concern about a sigmoid volvulus; thus, a computed tomographic (CT) scan of the abdomen was obtained. What disorder does this man have on the basis of his scout CT image (Panel A) and electrocardiogram (Panel B)?

Editor's note: We invite our readers to submit their answers at www.nejm.org/mystery. We will publish the diagnosis in the Correspondence section of the March 29 issue and e-mail it to everyone who submits an answer. All answers must be received by February 14.

Copyright © 2007 Massachusetts Medical Society.
A 50-year-old Asian woman presented with a papulonodular, erythematous rash on her legs below the knees. The skin lesions were nontender and nonpruritic and were accompanied by paresthesias. She had no fever, arthralgias, or other systemic symptoms.

The papulonodular rash on the legs coupled with paresthesias raises the possibility of a diagnosis of erythema nodosum. Erythema nodosum can develop as a delayed hypersensitivity response to a number of diseases, including streptococcal infection, tuberculosis, fungal infection, leprosy, inflammatory bowel disease, and sarcoidosis, as well as to drugs. It usually occurs on the front of the legs and consists of painful lesions that may resolve spontaneously; the lack of tenderness in this case decreases the likelihood of erythema nodosum. Rashes that can be confused with erythema nodosum include other forms of panniculitis, such as Behçet’s disease, subcutaneous infections caused by bacteria or fungi, superficial thrombophlebitis, and cutaneous vasculitides.

A biopsy of one of the skin lesions showed granulomatous dermatitis. Stains and cultures for acid-fast bacilli and fungal organisms were negative. On the basis of the clinical and histopathological findings, a diagnosis of sarcoidosis was made by the patient’s physician. No treatment was administered at that time.

Cutaneous lesions of sarcoidosis, which occur in about 25% of patients with sarcoidosis, may be specific, such as lupus pernio and plaques histologically showing noncaseating granulomas, or nonspecific, such as erythema nodosum histologically showing a nondiagnostic inflammatory reaction. Other skin lesions in patients with chronic sarcoidosis include maculopapular eruptions, subcutaneous nodules, changes in old scars, alopecia, and hypopigmented and hyperpigmented areas.

Skin lesions are occasionally the mode of presentation of sarcoidosis; however, they rarely appear as an isolated finding. Most patients have pulmonary or upper respiratory tract involvement; peripheral lymphadenopathy and ocular involvement are also common. When a granulomatous dermatitis is not accompanied by involvement of other organ systems, other causes must be considered, such as tuberculosis, berylliosis, leprosy, leishmaniasis, syphilis, or a deep fungal infection. Further evaluation is needed before concluding that this patient has sarcoidosis.

A chest radiograph reportedly revealed scattered nodules and increased interstitial markings that were considered to be consistent with sarcoidosis, and the patient was referred to a pulmonologist. Three months later, shortly before her appointment with the pulmonologist, dyspnea on minimal exertion, hoarseness, and dysphagia...
with both solids and liquids developed. She did not have fever, chills, weight loss, night sweats, arthralgias, chest pain, palpitations, or ocular symptoms. She was admitted to the hospital for further evaluation.

The new dyspnea, hoarseness, and dysphagia require immediate medical attention. Dysphagia with both solids and liquids is often due to a motility disorder of the esophagus; however, hoarseness indicates a lesion in the larynx or vocal cords. Immediate nasopharyngeal laryngoscopy may be required. The present symptoms may be explained by laryngeal sarcoidosis, which can occur in isolation or as a component of systemic sarcoidosis. Laryngeal tuberculosis and fungal disease of the larynx, although uncommon, should be considered, as should collagen vascular disorders that cause dysphagia (such as scleroderma, dermatomyositis, or systemic lupus erythematosus). Malignant conditions should also be considered. Although the findings on the chest radiograph may be consistent with sarcoidosis, high-resolution computed tomography (CT) of the chest is necessary to better evaluate diffuse lung abnormalities or lymphadenopathy.

The patient’s medical history included endometriosis. She was taking no medications and had no history of recent travel or use of tobacco, alcohol, or illicit drugs. She was employed as an administrator. On examination, she appeared cachectic and had a hoarse voice and difficulty completing full sentences. Her temperature was 37.2°C, and she had a blood pressure of 167/94 mm Hg, a heart rate of 97 beats per minute, and a respiratory rate of 20 breaths per minute. Her oxygen saturation was 99% while she was breathing ambient air. The head and neck examination revealed asymmetric elevation of the soft palate, with deviation of the uvula toward the right. She had no palpable cervical lymphadenopathy. Her lungs were clear to auscultation, and the cardiac examination was normal. Skin examination revealed a papulonodular, erythematous rash over the dorsal and ventral surfaces of the legs below the knees (Fig. 1). The lesions were nontender and less than 1 cm in diameter. The remainder of the neurologic examination was normal.

The persistent papulonodular, erythematous rash over the legs may still represent cutaneous nodular sarcoidosis, but the size of the lesions and their persistence during a 3-month period are not typical of erythema nodosum. Although the patient’s cachexia, hoarseness, and severe dyspnea and the cranial-nerve findings could still represent manifestations of sarcoidosis, I am increasingly worried about the possibility of a malignant condition.

The findings on examination of the cranial nerves involve the glossopharyngeal and vagus nerves and should be evaluated promptly. Could these findings be explained by neurosarcoidosis? Perhaps. Neurosarcoidosis is rare, but it can appear in an acute, explosive fashion or as a slow, long-term illness. It may affect any part of the nervous system, but the cranial nerves, hypothalamus, and pituitary gland are involved most commonly. However, other explanations for this patient’s cranial neuropathy, such as cancer with central nervous system metastases and fungal or mycobacterial infection causing a basilar meningitis, must still be considered.

The complete blood count, the differential count, and levels of serum electrolytes and creatinine were within normal limits. The serum calcium level was 9.5 mg per deciliter (2.4 mmol per liter), and the 24-hour urinary calcium excretion was 371 mg (9.3 mmol) (normal value, <300 mg [7.5 mmol]).
A chest radiograph showed increased interstitial markings and bilateral nodular opacities (Fig. 2). Results of pulmonary-function tests were normal, except for a decreased carbon monoxide diffusing capacity, which was 66% of normal. At this point, the physicians caring for the patient thought that sarcoidosis was the most likely diagnosis; however, no treatment was administered pending pulmonary consultation.

Her chest radiograph shows abnormalities that are consistent with stage III sarcoidosis (interstitial disease without clinically significant lymphadenopathy), but other diseases, including metastatic cancer or other granulomatous processes, could also explain these findings. The finding of an isolated reduction in the carbon monoxide diffusing capacity is not surprising, but it does not help in making the diagnosis of sarcoidosis, nor does it suggest another process. Hypercalcemia is present in almost a third of patients with sarcoidosis (hypercalcemia is present in approximately 10%), but its presence cannot be used to make the diagnosis. High-resolution CT of the chest is needed to better demonstrate the type, extent, and distribution of abnormalities.

High-resolution CT of the chest (Fig. 3) showed multiple nodules, 3 mm to 3 cm in diameter, located predominantly in the middle and lower lung zones. The nodules were randomly distributed, and several abutted the pleural surfaces and fissures. Prominent subcentimeter lymph nodes were noted in the aortopulmonary window, but the mediastinal and hilar lymph nodes were normal.

Although the findings on this CT scan are consistent with sarcoidosis, they are not classic for the disease. The role of CT in the diagnosis of sarcoidosis remains poorly defined. Several findings, especially in combination, support the diagnosis of sarcoidosis — namely, mediastinal and hilar lymphadenopathy, lung involvement with upper-lobe predominance, small nodules distributed along bronchovascular bundles or subpleurally, and interlobular septal thickening and architectural distortion. None of these findings, however, are specific enough to be considered diagnostic of sarcoidosis.

I continue to be concerned about cancer — particularly metastatic disease or lymphoma — and would recommend that the patient undergo lung biopsy. Endobronchial and transbronchial lung biopsy have low rates of complications and a reasonable diagnostic yield for the diseases being considered. The diagnostic yield for sarcoidosis largely depends on the experience of the operator, ranging from 40% to more than 90% when four or five specimens are taken on transbronchial lung biopsy. In the case of cancer, the combination of cytologic examinations plus histopathological examinations of tissue sections has a greater than 75% diagnostic yield for lesions involving the lung diffusely. Before this approach is pursued, though, the patient’s abnormal neurologic findings require urgent evaluation, includ-
An ophthalmologic examination was normal. Nasopharyngolaryngoscopy revealed a partially paralyzed left vocal cord; a biopsy was not performed. MRI of the brain revealed a nodule, 10 mm in diameter, in the suprasellar cistern and enhancement of the third cranial nerve bilaterally. Analysis of cerebrospinal fluid obtained by means of lumbar puncture revealed 9 white cells per cubic millimeter, of which 95% were lymphocytes; glucose and protein levels were normal.

The absence of ocular findings is not helpful in this case; only 20% of patients with sarcoidosis have ophthalmologic disease, and the other diseases under consideration infrequently present with ocular pathologic features. The findings on MRI are consistent not only with neurosarcoidosis but also with several other diseases, including metastatic cancer, lymphoma, Langerhans’ cell histiocytosis, and fungal or mycobacterial infection. Further testing is still warranted to rule out causes other than sarcoidosis.

The patient underwent fiberoptic bronchoscopy. The airways were normal. Stains and cultures of bronchoalveolar-lavage specimens for bacteria, fungi, and acid-fast bacilli were negative. A cell count in the lavage fluid showed 708 white cells per cubic millimeter, 57% macrophages (normal, 80% or higher), 25% lymphocytes (normal, 15% or lower), 15% eosinophils (normal, 1% or lower), and 3% neutrophils (normal, 2% or lower). The ratio of total CD4 to CD8 T lymphocytes was 6.4 (normal value ±SE, 2.6±0.6). Lung biopsies yielded eight tissue specimens that showed scattered areas of mixed acute and chronic inflammation without granulomas.

Lymphocytosis in bronchoalveolar-lavage fluid is common in patients with sarcoidosis, tuberculosis, or hypersensitivity pneumonitis. In cases of sarcoidosis, bronchoalveolar-lavage fluid reveals a high percentage of lymphocytes in 90% of patients. A predominance of T lymphocytes with an elevated CD4:CD8 ratio is also characteristic; a CD4:CD8 ratio of 3.5 or greater has a specificity for sarcoidosis of 94% and a sensitivity of 52%. Although these findings alone would have provided further support for sarcoidosis as the cause of the patient’s presenting symptoms and signs, I am troubled by the percentages of neutrophils and eosinophils in the bronchoalveolar-lavage fluid in this patient, since these are unusual for cases of sarcoidosis. As a general rule, bronchoalveolar-lavage fluid with more than 2% neutrophils or more than 1% eosinophils should suggest an alternative diagnosis. Surgical lung biopsy is warranted at this point to clarify the diagnosis.

Given the absence of granulomas in the biopsy specimens and the unusually high eosinophil count in the bronchoalveolar-lavage fluid, the diagnosis of sarcoidosis was considered unlikely to explain the patient’s clinical findings, so the patient underwent a video-assisted thoracoscopic lung biopsy (Fig. 4). Pathological examination revealed sheets of lymphocytes clustering around and within the walls of blood vessels (Fig. 5). Immunohistochemical staining revealed a predominantly B-cell population. In addition, an in situ hybridization test for Epstein–Barr virus was positive, confirming the diagnosis of lymphomatoid granulomatosis.

The patient was enrolled in a trial of interferon alfa therapy. Her response to this therapy was poor and was complicated by the development of idiopathic thrombocytopenic purpura. She was therefore treated with a regimen of rituximab, etoposide, cyclophosphamide, vincristine, doxorubicin, and prednisone, followed by maintenance therapy with interferon alfa, and she had a complete response to this regimen. She remains disease-free 23 months after diagnosis and has returned to part-time employment; her hoarseness is improving.

Figure 4. Cut Surface of a Specimen from a Video-Assisted Thoracoscopic Lung Biopsy, Showing Several Pale, Solid Nodules.
First described in 1972, lymphomatoid granulomatosis is an angiocentric large-B-cell lymphoproliferative process marked pathologically by transmural infiltration of vessel walls with both B and T lymphocytes.\(^1\) The lung is the most commonly involved organ, but the skin and nervous system are also frequently affected. Pulmonary disease commonly presents with multiple bilateral nodules principally involving the middle and lower lung fields; the radiographic presentation may be difficult to distinguish from other, more common diseases, including lymphoma, lymphocytic interstitial pneumonia, metastatic disease, sarcoidosis, Wegener's granulomatosis, and cryptogenic organizing pneumonia.\(^2\) In contrast to other pulmonary lymphomas, mediastinal and hilar lymphadenopathy are atypical in patients with lymphomatoid granulomatosis. Cutaneous disease may take various forms, ranging from an erythematous maculopapular eruption to subcutaneous nodules; the rash is most commonly found on the arms and legs and is not confluent.\(^3\) Neurologic disease may present as an isolated peripheral or cranial neuropathy, as a central mass lesion, or with seizures.

Since the clinical presentation of lymphomatoid granulomatosis is nonspecific, diagnosis requires the histopathological demonstration of polymorphic lymphoid infiltrates with focal areas of necrosis, as well as transmural vascular infiltration by pathologic B cells and reactive T cells. Immunohistochemical staining with the B-cell marker CD-20 and in situ hybridization with Epstein–Barr virus DNA are frequently required to confirm the diagnosis.\(^2,4\) Although Epstein–Barr virus DNA and RNA are often found in B cells in diseased tissue, causality has not yet been proved; in some cases, particularly those with isolated cutaneous involvement, it may not be possible to demonstrate Epstein–Barr virus positivity.\(^3\) It is noteworthy, however, that lymphomatoid granulomatosis most commonly — but not always, as seen in this case — presents in immunocompromised patients.

Although some cases follow a benign clinical course, the majority progress rapidly to an aggressive high-grade lymphoma, and median survival after diagnosis is 14 months. Treatment commonly consists of corticosteroids and cyclophosphamide, although there are recent reports of successful treatment with rituximab or interferon alfa, as this patient received.\(^4,5\)

Why did the physicians initially involved in this case settle on a diagnosis of sarcoidosis? They most likely fell prey to two types of cognitive bias: the availability and anchoring heuristics. Heuristics are learned shortcuts that all clinicians use to improve the efficiency of the diagnostic process.\(^6\) Unfortunately, they can also lead to diagnostic errors. The availability heuristic states that new events are most easily grouped cognitively with memorable previous experiences that are similar in some detail or pattern to the present event.\(^7\) Several clinical manifestations of sarcoidosis mimic those of lymphomatoid granulomatosis. Many clinicians have cared for patients with sarcoidosis, whereas very few have seen patients with lymphomatoid granulomatosis. Thus, the combi-
nation of pulmonary, cutaneous, and neurologic involvement in this case may inspire recollection of a similar case of sarcoidosis instead of prompting consideration of all diseases that may present in this fashion.

The anchoring heuristic occurs when the estimation of the likelihood of a diagnosis is based heavily on initial test results or on a previous diagnosis. In this case, the primary clinicians relied heavily on the dermatologist’s interpretation of the skin biopsy as sarcoidosis. Although the skin biopsy was consistent with sarcoidosis, it was not diagnostic of sarcoidosis; as the discussant points out, no findings on skin biopsy are diagnostic of sarcoidosis. No frank granulomas were seen, and the prominent lymphohistiocytic infiltrate and mild lymphocytic atypia seen in the biopsy specimen have frequently been described in lymphomatoid granulomatosis. Nevertheless, the clinicians initially accepted the diagnosis and were falsely anchored to it from that point onward. Consequently, subsequent test results were interpreted by the clinical team as consistent with sarcoidosis, even when the results were atypical for that diagnosis. Fortunately, the missteps in clinical reasoning did not result in a substantial diagnostic or therapeutic delay.

The discussant recognized that many of the study results, such as the findings on the CT scan and eosinophilia in the bronchoalveolar-lavage fluid, were not typical for sarcoidosis. In consistently noting the need for diagnostic verification, the discussant averted the trap of “premature closure” — a common error in clinical reasoning that occurs when a diagnosis is accepted before it has been adequately validated, thus preventing consideration of alternative diagnoses that may fit the clinical scenario better.

In the clinical reasoning process, the initial diagnosis is often only a hypothesis and requires additional confirmatory testing, especially when some findings do not clearly support the diagnosis. Although heuristics can improve diagnostic efficiency, they may also result in pitfalls that lead the diagnostic process into troubled waters. It is worth remembering such biases to avoid being anchored to the wrong diagnosis.

Dr. Saint is supported by a Career Development Award from the Health Services Research & Development Program of the Department of Veterans Affairs and a Patient Safety Developmental Center Grant (P20-HS11540) from the Agency for Healthcare Research and Quality.

No potential conflict of interest relevant to this article was reported.

We thank Kirk D. Jones, M.D., of the Department of Pathology, University of California, San Francisco, for preparing the pathological images.

REFERENCES

Copyright © 2007 Massachusetts Medical Society.
Cardiovascular Risks from Fine Particulate Air Pollution
Douglas W. Dockery, Sc.D., and Peter H. Stone, M.D.

More than a decade ago, prospective epidemiologic studies showed that mortality was increased among people living in communities with elevated concentrations of fine particulate air pollution.\(^1,2\) Subsequent research has shown that particulate air pollution is statistically and mechanistically linked to increased cardiovascular disease.\(^3\) New data are beginning to shed light on which persons are at heightened risk.

In this issue of the Journal, Miller et al.\(^4\) report on data from the Women’s Health Initiative (WHI) observational study, which greatly expands our understanding of how fine particulate pollution affects health. Earlier long-term prospective cohort studies showed an association between levels of air pollution consisting of particulate matter of less than 2.5 μm in aerodynamic diameter (PM\(_{2.5}\)) and an elevated risk of death from all causes and from cardiovascular disease.\(^1,2,5\) The WHI study broadens the scope by finding that nonfatal cardiovascular events are also strongly associated with fine particulate concentrations in the community. Earlier work relied solely on death certificates to define the rate of death from cardiovascular disease. In the WHI study, cardiovascular events and mortality were defined by objective review of medical records. The earlier studies were designed to identify risk factors for respiratory disease\(^1\) and cancer\(^2\) and therefore had limited ability to adjust for cardiovascular risk factors. The WHI observational study was designed to assess the risk of cardiovascular events and therefore could exclude cardiovascular risk factors as explanations for the observed associations with air pollution.

Earlier studies did not include data on the full range of regulated community air pollutants — that is, PM\(_{2.5}\) (and the larger particle fraction, PM\(_{10}\)), sulfur dioxide, nitrogen dioxide, carbon monoxide, and ozone. The WHI study considered all of these community air pollutants and found cardiovascular risk associated only with PM\(_{2.5}\) concentrations. Whereas earlier work compared levels of air pollution and rates of death between various cities, the WHI investigators were also able to compare areas within individual cities. Their analysis demonstrated a relationship between increased levels of fine particulate pollution and higher rates of death and complications from cardiovascular and cerebrovascular disease, depending not only on which city a person lived in but also on where in that city she lived.

Perhaps most important, the WHI study established a stronger statistical association between fine particulate air pollution and death from coronary heart disease than that found in earlier studies. In the WHI study, Miller et al. found an increased relative risk of 1.76 for death from cardiovascular disease for every increase of 10 μg per cubic meter in the mean concentration of PM\(_{2.5}\).\(^4\) By comparison, a study by the American Cancer Society showed that each increase of 10 μg per cubic meter in the mean PM\(_{2.5}\) concentration was associated with an increased relative risk of 1.12 for death from cardiovascular disease, 1.18 for death from ischemic heart disease (the largest proportion of deaths), and 1.13 for death from arrhythmia, heart failure, or cardiac arrest.\(^5\)

Samples in previous studies consisted of subjects from the entire population of the cities being investigated. The WHI analysis was restricted to postmenopausal women with no history of cardiovascular health problems. A 22-year follow-up of a cohort of nonsmoking white adults in California showed an increased risk of death from coronary heart disease with rising levels of
fine particulate air pollution in women but not in men.\(^6\) Does this suggest that the WHI population, or women in general, are more sensitive to the cardiovascular effects of particulate air pollution?

Women have a distinctly different profile of coronary disease. In the Women’s Ischemia Syndrome Evaluation study, the cluster of conditions that increase the risk of vascular disease (e.g., hypertension, diabetes, obesity, and inactivity) was seen more frequently in postmenopausal women than in men.\(^7\) Women’s coronary arteries are smaller in size and tend to harbor more diffuse atherosclerosis than do men’s arteries, and women’s microvessels appear to be more frequently dysfunctional than those of men.\(^7\) Indeed, in the Euro Heart Survey, although women were less likely than men to have fixed atherosclerotic obstructive disease, among patients undergoing elective diagnostic angiography for angina, women with confirmed coronary disease had twice the risk of death or myocardial infarction as that of men.\(^8\) These findings suggest that sex may not define susceptibility to air pollution but, rather, may be an indicator of an underlying cardiac substrate that puts women at increased risk.

Characteristics that define increased cardiovascular susceptibility to particulate air pollution have also been identified in men. Stronger associations between fine particulate concentrations and abnormal variability in heart rate were reported in asymptomatic men with higher Framingham cardiovascular risk scores.\(^9\) PM\(_{2.5}\) was more strongly associated with impaired autonomic cardiovascular function in men with genotypic and phenotypic indicators of increased systemic inflammation and oxidative stress than in those without these markers.\(^10\) However, the increased susceptibility was not found among men taking statins, which both improve lipid profiles and reduce systemic inflammation.

The mechanisms by which fine particulate air pollution influence the risk of cardiovascular disease are still under investigation. There is evidence that inhalation of particulate air pollution creates and exacerbates both pulmonary and systemic inflammation and oxidative stress, leading to direct vascular injury, atherosclerosis, and autonomic dysfunction.\(^3\) Buildup of atherosclerotic plaque, measured by the carotid intima–media thickness, is higher in communities with higher mean PM\(_{2.5}\) concentrations.\(^11\) Particulate air pollution has been found to lead to rapid and significant increases in fibrinogen, plasma viscosity, platelet activation, and release of endothelins, a family of potent vasoconstrictor molecules.\(^3\)

Taken together, these studies suggest that the status of cardiovascular risk factors has a substantial effect on susceptibility to the adverse effects of particulate air pollution. A particularly appealing aspect of the design of the WHI study is the range of data collected on all subjects, including demographic and lifestyle characteristics, cardiovascular risk factors, medical history, diet, and medications. With this wealth of data, the next generation of analyses should be able to focus risk stratification even further to identify the characteristics of persons who are most susceptible to the adverse effects of air pollution.

A multifaceted approach that encompasses both public health and medical interventions is needed to reduce the burden of cardiovascular disease attributable to air pollution. Comprehensive management of the harmful effects of fine particles must start with intensive efforts to reduce this destructive form of air pollution. Fine particulate air pollution results not only from the combustion of carbonaceous fuels in our vehicles, power plants, and factories but also from secondary particles produced by oxidation of gaseous pollutants emitted by these same sources. The evidence that has accumulated thus far regarding the health threat from PM\(_{2.5}\) pollution is convincing enough to have prompted the Environmental Protection Agency (EPA) to lower the short-term (24-hour) standard for fine particulate concentration that communities must achieve. Unfortunately for public health, the EPA failed to follow the recommendation of its science advisers and reduce the long-term standard for fine particles.\(^12\) The findings of the WHI study strongly support the recommendation for tighter standards for long-term fine particulate air pollution.

Even with tighter standards, people will continue to be exposed to fine particulate air pollution. Although the public health burden of cardiovascular disease attributable to air pollution is large, the evidence suggests that individual risks are modest. If the WHI and other studies can identify intrinsic and acquired individual factors that lead to increased adverse cardiovascular responses to air pollution, then it should be possible to offer focused interventions to persons who...
are at greatest risk and thereby ameliorate at least some of the patient-specific damages of air pollution.

Dr. Dockery reports receiving a stipend from Industrial Economics for participation in an expert-opinion study. No other potential conflict of interest relevant to this article was reported.

From the Department of Environmental Health, Harvard School of Public Health (D.W.D.), and the Cardiovascular Division, Brigham and Women’s Hospital and Harvard Medical School (P.H.S.) — both in Boston.


The Healing Power of Listening in the ICU
Craig M. Lilly, M.D., and Barbara J. Daly, Ph.D, R.N.

Critical care services are highly valued because they can often restore function in patients with acute life-threatening illnesses. In this context, advances in medical science have led to increased expectations for favorable outcomes of episodes of critical illness, even when the patient has severe coexisting chronic disease. The growing demand for critical care has led both to increased numbers of patients who survived with desirable functional outcomes and to increased numbers of patients who die in the intensive care unit (ICU).

Today, many deaths in the ICU occur after a decision has been made to discontinue or forgo advanced supportive technology.1 Decisions to shift from apparently ineffective technology to a treatment plan that focuses primarily on the patient’s comfort are usually made in discussions between caregivers and family members.2 These discussions involve complex conversations and are important to families. Communication processes that have been shown to improve the well-being of patients and family members include proactive, multidisciplinary sessions that provide patients (when they are able to communicate) and family members with the opportunity to ask questions, articulate the patient’s values, express painful emotions, discuss concerns, and obtain help with managing feelings of guilt.3

A clinical course that runs counter to the family’s hopes and expectations is extraordinarily stressful and is an important contributor to ICU-related post-traumatic stress disorder (PTSD) among families.4 A better understanding of how intensive care clinicians can support families as they make the transition from a goal of cure to one of comfort and acceptance of death is clearly needed. Recognition of the relationship between satisfaction, on the one hand, and expectations, perceptions, and prognosis, on the other hand, can lead to communication processes that synchronize the perceptions of family members with those of providers and close gaps between reality and expectations. Curtis and colleagues have described some of the components of a system of communication that is being increasingly recognized as an effective means of promoting harmony between critical care providers and families.5 This five-part system, known by the mnemonic VALUE, includes the following elements: valuing and appreciating what the family mem-

Copyright © 2007 Massachusetts Medical Society.
bers communicate, acknowledging their emotions by using reflective summary statements, listening to family members, understanding who the patient is as a person by asking open-ended questions and listening carefully to the responses, and eliciting questions from the family more effectively than by simply asking, “Any questions?” A key skill is listening more and talking less.6 Structured, proactive, multidisciplinary communication processes7 that are supported by ethics consultation8 and palliative care teams9 and include bereavement conferences that encourage providers to use a structured approach (such as the VALUE system) for guiding effective communication during critical care10 are the foundations for improving end-of-life care for patients and interactions with their families.11

The importance of understanding how to use effective communication to improve end-of-life care is increasingly supported by randomized intervention studies — such as the study by Lautrette and colleagues reported in this issue of the Journal12 — that meet most of the accepted standards of good clinical science. Lautrette et al. found that formal bereavement meetings held at the time that the senior physician had concluded that death was inevitable improved the well-being of family members, as measured by validated instruments. This study is groundbreaking in its demonstration of a statistically and clinically significant improvement in symptoms of anxiety, depression, and PTSD among family members, and it shows that expanding the focus of critical care to include family-centered outcomes is appropriate and desirable. In reporting these advances in the peer-reviewed literature, it is often difficult for authors to fully explain the core of their interventions, in part because of the complex, diverse, and emotion-laden nature of these multidimensional conversations.

Although the amount of time spent listening in an individual case will be driven primarily by the medical facts and the needs of the persons facing loss, the study by Lautrette and colleagues12 suggests that spending an average of 30 minutes (or 10 minutes longer than typical practice) with the patient’s family members leads to a significant improvement in their well-being in the months after their loss. Since there is substantial variation in the frequency of deaths, depending on the size of the ICU and the mortality rate of the population served, the fraction of time caregivers spend to help families manage the critical illness and death of a loved one will vary. On the basis of our research, we propose that the time clinicians working in adult ICUs spend supporting family members in shared decision making should roughly correspond to the mortality rate of the ICU patient population. For example, a full-time clinician serving a surgical ICU with a mortality rate of 2% would spend about 2 hours a week supporting patients and families. The same clinician would spend about 2 hours per day supporting families when serving a medical ICU with a mortality rate of 20%.13 When used effectively, this time can translate into considerable savings in costs by reducing the number of days a patient spends in the ICU before death13; the time spent with families thus deserves support for compelling economic as well as humanistic reasons.

Recommendations to improve care for patients dying in ICUs are rooted in both observational and interventional studies. Observational studies confirm our own practical experience that nearly every American family will be affected by the loss of a loved one in an ICU and that the effect of this loss can be mitigated by high-quality care. The field has been advanced by interventional studies showing that proactive communication processes, including intensive communication13 as well as ethics8 and palliative care9 consultations, improve outcomes. Evidence that proactive multidisciplinary conferences in which care providers and family members address bereavement, with the provision of printed materials, is another important advance in the field of end-of-life care in the ICU. All providers of critical care should receive training that will allow them to offer the kind of support that they would want if they had a family member who was facing death in an ICU.

No potential conflict of interest relevant to this article was reported.

From the University of Massachusetts Medical School, Worcester (C.M.L.); and Case Western University, Cleveland (B.J.D.).

4. Azoulay E, Pochard F, Kentsh-Barnes N, et al. Risk of post-
Pay for Performance at the Tipping Point
Arnold M. Epstein, M.D.

It is hard to dispute the rationale behind realigning payment incentives in health care to encourage higher quality and more efficient care. Indeed, across the country and beyond, the number of “pay for performance” programs, as such realignment is called, has reached a tipping point. In the United States, more than half the health maintenance organizations (HMOs) in the private sector have now initiated such programs, covering more than 80% of the country’s HMO enrollees.1 Congress has mandated that the Center for Medicare and Medicaid Services (CMS) develop plans to introduce a pay-for-performance program into Medicare.2 The British have gone a league further, introducing their own version of pay for performance that puts 25 to 30% of the income of family practitioners at stake.3

Because the rationale behind pay for performance is so compelling, it may seem surprising that the evidence base linking such programs to a better quality of care is thin (at least, according to two recent review articles4,5). Most previous studies have looked at incentives to physicians and medical groups. The data showing efficacy are inconsistent, and some studies have revealed unintended effects, such as improvement in documentation without much change in the underlying quality of care.6 Only one previous study examined cost-effectiveness.7

Given this dearth of solid evidence, it seems apt to compare our adoption of pay for performance with our adoption of new surgical procedures or medical therapies. Many of my clinical colleagues would insist on hard evidence documenting efficacy before endorsing a new therapeutic approach. They cite sobering stories of what can happen when we introduce new approaches prematurely. Consider, for example, the numerous surgical procedures or medical therapies—including radical mastectomy for women with early-stage breast cancer and hormone-replacement therapy for postmenopausal women—that were diffused widely before solid evidence of their relative efficacy was available, only for us to learn later that they were, at best, no more effective than alternative therapies or, at worst, harmful.8-10 If pay for performance were a therapy, its rapid diffusion thus far would have to be considered premature.

The study by Lindenauer et al.11 in this issue of the Journal begins to address this information gap on pay for performance. The authors report the initial results of a 3-year program in which more than 200 hospitals participating in a quality-benchmarking database maintained by Premier volunteered for a Medicare demonstration in which payments would be allocated partially on the basis of quality performance. Hospitals performing in the top decile received a 2% increment in Medicare payments, whereas hospitals in the second decile received a 1% increment. Hospitals that underperformed by failing to exceed the performance of hospitals in the lowest two deciles (as established during the program’s first

12. Copyright © 2007 Massachusetts Medical Society.
year) were liable for a 1 to 2% financial penalty in the third year.

Lindenauer et al. matched these hospitals with 406 hospitals that were providing the CMS with a subgroup of the same quality-performance data. The latter data were intended for public reporting but not for additional payments. The pay-for-performance hospitals showed significantly greater improvement than the hospitals that engaged in public reporting alone. However, after adjustment for confounders, the overall differential was only 2.9%, and that number probably overestimates the effect of the program, since the participating hospitals were a self-selected group whose administrators probably thought their performance would exceed the payment threshold.

Besides gauging the effect of pay for performance, the study by Lindenauer et al. also challenges the leading rationale for providing financial incentives.12,13 For years we have assumed that rewarding higher quality with higher payments directly motivates physicians and hospitals to invest in personnel and systems to improve the quality of care. However, the data from this study suggest that the causal chain may be more complicated. If gaining financial reward were indeed the primary impetus for hospitals to improve performance, one might expect that pay for performance would have its largest effect relative to public reporting in the hospitals with the best chance of rising above the quality threshold (i.e., those in quintiles one and two). One might also expect the smallest relative improvement in hospitals that are farthest away from the threshold (in quintile five). However, the study by Lindenauer et al. does not show such patterns. Perhaps the findings are idiosyncratic, reflecting the low level of payments, the voluntary nature of the Premier demonstration, or the penalty that low-performing hospitals potentially suffer in year three. However, at least one other similar study showed behavior equally incompatible with expectations.14 Perhaps the explanation is that improvements in quality performance are easier to make at the low end and that the additional attention that financial rewards draw to performance catalyzes professional ethos. These explanations, if true, would be good news for those who are concerned about budget constraints undermining pay-for-performance programs.

Because of the federal sponsorship of this study and the vast resources required to carry it out, the results of the Premier demonstration have been eagerly awaited. However, the findings still leave us with many uncertainties concerning the level of financial incentives needed and the optimal formula for payment that might be used for attaining high levels of performance. Returning to the medical-advances analogy, we have learned through the years that medical therapies and procedures are not cost-effective per se but, rather, are more or less cost-effective for various populations and for various medical indications. Policymakers need similar fine-grained information about different aspects of pay for performance. The reality, however, is that we are at the tipping point with pay-for-performance programs, and such information is unlikely to be forthcoming before political pressure forces policymakers to act.

In this situation, the CMS may have much to gain from recognizing that pay for performance is fundamentally a social experiment likely to have only modest incremental value. Broad demonstration and evaluation will probably be helpful. Rather than adopt a single new payment system for all of Medicare, a series of regional models could accelerate learning and allow Medicare officials to find out more about the effect of differing levels of incentives and formulas for payment. No matter what the course, timely evaluation of any policy we adopt seems critical to ensure that we achieve high performance without unintended consequences.

No potential conflict of interest relevant to this article was reported.

From the Harvard School of Public Health, Boston.

Huntington’s Disease — Making Connections

J. Timothy Greenamyre, M.D., Ph.D.

When the causative gene for Huntington’s disease was identified in 1993, there was great anticipation and hope that key disease-causing mechanisms would be identified quickly and that rational neuroprotective treatments would soon follow. Fourteen years later, it is obvious we were wrong. Potential pathogenic mechanisms have proliferated, and their relative importance is unclear. Now, three recent studies\textsuperscript{1-3} have identified a protein and a mechanism that link two of the leading hypotheses of pathogenesis: transcriptional dysregulation and mitochondrial impairment (Fig. 1).

Huntington's disease is caused by the expansion of a CAG trinucleotide repeat in exon 1 of the gene that encodes huntingtin protein. If the gene contains more than 35 CAGs, each of which codes for a glutamine residue, Huntington's disease is likely to develop. Mutant huntingtin, by virtue of its lengthy polyglutamine stretch, binds to certain transcription factors (or proteins that interact with transcription factors), which themselves contain runs of polyglutamines.\textsuperscript{4} For example, mutant huntingtin binds and sequesters cyclic AMP response-element–binding protein (CREB)–binding protein (CBP), which alters the expression of genes regulated by the transcription factor CREB. In a similar way, mutant huntingtin interferes with Sp1-mediated gene transcription. So it seems that mutant huntingtin may alter the complement of proteins that are synthesized in a cell, a change that may lead to the pattern of neurodegeneration that characterizes Huntington’s disease.

A competing hypothesis is that neurodegeneration in Huntington’s disease results from mitochondrial impairment.\textsuperscript{5} Indeed, there is extensive evidence showing biochemical, morphologic, and functional mitochondrial abnormalities in patients with this disease. The activities of mitochondrial electron transport complexes II, III, and IV are reduced in Huntington’s disease, and magnetic resonance spectroscopy of the brain shows elevated lactate levels, a finding that is consistent with mitochondrial dysfunction. How mitochondrial impairment arises is uncertain, but a direct effect of mutant huntingtin on mitochondria may be partly to blame.\textsuperscript{6} Could mitochondrial dysfunction also be caused by altered gene transcription? After all, the vast majority of mitochondrial proteins are encoded by the nuclear genome and imported into the organelle — but what regulates their transcription?

The peroxisome proliferator-activated receptor-\(\gamma\) coactivator 1\(\alpha\) (PGC-1\(\alpha\)) is a transcriptional coactivator that controls many metabolic processes, including mitochondrial biogenesis, oxidative phosphorylation, and adaptive thermogenesis (the body’s response to cold temperatures). It regulates the expression of a large number of genes, including nuclear-encoded subunits of each of the electron-transport-chain complexes and several genes that provide protection against the effect of reactive oxygen species (ROS). As such, PGC-1\(\alpha\) seems well situated to bridge transcriptional dysregulation and mitochondrial impairment — as suggested by the three recent studies\textsuperscript{1-3}.\textsuperscript{1}

Cui et al.\textsuperscript{1} report that mutant huntingtin represses expression of the gene encoding PGC-1\(\alpha\) by binding to its promoter and interfering with its CREB-dependent transcription. Weydt et al.\textsuperscript{3} report reduced striatal expression of PGC-1\(\alpha\)–regulated genes in Huntington’s disease, including subunits of the electron-transport chain — an observation that is consistent with the finding reported by Cui et al. Moreover, St.-Pierre et al.\textsuperscript{2} show that cells that do not express PGC-1\(\alpha\) have an impaired ROS defense system because their expression of key PGC-1\(\alpha\)–regulated antioxidant enzymes is reduced. Perhaps the strongest and most direct evidence that reduced PGC-1\(\alpha\) activity results in neurodegeneration is the finding that overexpression of PGC-1\(\alpha\) ameliorates the...
atrophy of striatal neurons that normally occurs in transgenic mice with Huntington’s disease.1

Therefore, the basic idea is that in patients with this disease, PGC-1α–regulated gene transcription is defective. As a result, there is reduced expression of mitochondrial and antioxidant genes regulated by PGC-1α. In this way, PGC-1α provides a plausible link between what were previously unrelated mechanisms: transcriptional dysregulation and mitochondrial impairment. St.-Pierre et al.2 suggest that manipulation of PGC-1α may be therapeutic in other neurodegenerative diseases in which oxidative stress is pathogenic. They observed that mice lacking the PGC-1α gene have increased sensitivity to the parkinsonism-inducing neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) and, conversely, that neural cells that are made to express exogenous PGC-1α are comparatively resistant to oxidative stressors, such as paraquat.

The three studies3-5 underscore the role of PGC-1α in neurodegeneration. They raise the possibility that increasing PGC-1α expression or function might be therapeutic in Huntington’s disease and other neurodegenerative disorders. Tempering enthusiasm for this approach is the observation that overexpression of PGC-1α in the heart causes uncontrolled mitochondrial proliferation, destruction of sarcomeric structure, and dilated cardiomyopathy.7 That said, further study of PGC-1α will almost certainly lead to new therapeutic strategies to delay or retard the onslaught of neurodegenerative diseases.

No potential conflict of interest relevant to this article was reported.

From the Pittsburgh Institute for Neurodegenerative Diseases and the Department of Neurology, University of Pittsburgh.


Copyright © 2007 Massachusetts Medical Society.
Rituximab for Pemphigus Vulgaris

TO THE EDITOR: Ahmed et al. (Oct. 26 issue) report a resolution of pemphigus in 9 of 11 patients with refractory disease treated with a combination of 10 infusions of rituximab and 6 of intravenous immune globulin. On the basis of our experience with more than 20 patients with severe pemphigus who were treated with only one cycle of four infusions of rituximab, we confirm that rituximab has a dramatic effect: more than 90% of our patients had a complete remission, which was maintained after a 24-month follow-up period in two thirds of the patients. We did not treat the patients with intravenous immune globulin, and we observed only two cases of late infections. A single cycle of rituximab should be compared with the combination regimen proposed by Ahmed et al. in a clinical trial.

Pascal Joly, M.D., Ph.D.
Rouen University Hospital
76000 Rouen, France
pascal.joly@chu-rouen.fr

Michel D’Incan, M.D., Ph.D.
Clermont-Ferrand University Hospital
63000 Clermont-Ferrand, France

Philippe Musette, M.D., Ph.D.
Rouen University Hospital
76000 Rouen, France


THE AUTHORS REPLY: On the basis of the comments by Joly et al. about their patients with pemphigus, the patients treated with our published regimen appear to have fared better, although without the clinical details of their study in its entirety, it is difficult to make valid comparisons. Important questions with respect to their study include the definition of “severe pemphigus,” the duration of disease, immunosuppressive therapies used before rituximab therapy, and the ability to discontinue immunosuppressive therapies after rituximab therapy. The observations by Joly et al. appear to be similar to those in a report on 17 patients with pemphigus vulgaris who were treated with rituximab. We were able to discontinue all immunosuppressive agents in our patients before the cessation of rituximab therapy, and the patients have not needed them since. The role of intravenous immune globulin therapy is more than as prophylaxis against acute infections. The immunomodulatory effects of the drug may be synergistic with and complementary to rituximab in producing lasting periods of remission by

THIS WEEK’S LETTERS

521 Rituximab for Pemphigus Vulgaris
522 Effect of Ramipril on the Incidence of Diabetes
524 Prevention of Meningococcal Disease
526 Medical Mystery: Abnormal Abdominal Radiograph — The Answer
527 A Girl with Fever after a Visit to Africa
529 A Prison Inmate Intoxicated with a Hand Sanitizer
530 A Hospitalized Patient Intoxicated with a Hand Sanitizer
preventing the reappearance of pathogenic B-cell clones and blocking the stimulation of T-cell help.

Pemphigus is an excellent model for investigating autoimmunity and providing insights into the pathogenesis of more complex autoimmune diseases. The availability of this effective therapy should lead to randomized trials that will contribute to this process.

Marshall R. Posner, M.D.
Dana–Farber Cancer Institute
Boston, MA 02115

Zachary Spigelman, M.D.
A. Razzaque Ahmed, M.D.
New England Baptist Hospital
Boston, MA 02120


Effect of Ramipril on the Incidence of Diabetes

TO THE EDITOR: The Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication (DREAM) trial investigators (Oct. 12 issue)1 report that there was an increase in the incidence of regression to normoglycemia (a secondary outcome) of 16% among patients receiving ramipril as compared with those receiving placebo but that there was no effect on the incidence of diabetes or death. The authors appear to have overlooked the significant reduction in mean diastolic blood pressure in the ramipril group, as compared with the reduction in the placebo group, at 2 months (4.3 mm Hg vs. 1.6 mm Hg). We wonder whether the effect on the secondary outcome could merely be the result of decreased blood pressure rather than a favorable glycemic effect of ramipril itself.

Previous studies have shown that hypertension itself worsens insulin resistance2 and that insulin resistance can predispose patients to hypertension.3,4 In the International Verapamil Sustained Release–Trandolapril Study (INVEST) cohort of 16,176 patients, the risk of new-onset diabetes was directly and independently associated with blood pressure at follow-up.5 Given this interaction between insulin resistance and impaired glucose tolerance and hypertension, we propose that the effect of ramipril on glycemic control reflects, to a large extent, better blood-pressure control with a metabolically inert medication rather than an effect of the angiotensin-converting–enzyme (ACE) inhibitor itself.

Sripal Bangalore, M.D., M.H.A.
Franz H. Messerli, M.D.
St. Luke’s–Roosevelt Hospital
New York, NY 10025
sbangalo@chpnet.org


TO THE EDITOR: Diabetes is an ever-increasing problem in our society, associated with considerable morbidity, mortality, and financial costs. In the DREAM trial, ramipril failed to reduce the incidence of diabetes or death, but it did increase the incidence of regression to normoglycemia.

However, although listed as an exclusion criterion,1 140 participants in the placebo group in the DREAM trial received angiotensin-receptor blockers (ARBs). ARBs and ACE inhibitors probably have a similar metabolic effect,2-4 and it is therefore possible that the inclusion of these pa-
Patients diminished the effect of ramipril on the primary end point. Similarly—and more important—since the expected metabolic effect of even supranormal doses of ramipril is likely to be much less than that of rosiglitazone, it is conceivable that any effect of ramipril on the primary outcome was masked by the overwhelming benefit of rosiglitazone.

It would be interesting to know the results of a post hoc analysis for the primary outcome excluding participants who received either rosiglitazone or ARBs. Although we recognize the limitations of such an analysis, it might help guide future research on this important topic.

Brian J. Potter, M.D.
Jacques LeLorier, M.D., Ph.D.
Centre Hospitalier de l’Université de Montréal
Montreal, QC H2W 1T8, Canada
brianpotter@gmail.com


TO THE EDITOR: The DREAM trial extends findings from previous studies regarding the effect of inhibiting the renin–angiotensin system on new-onset diabetes, and the accompanying editorial places the implications of the study in perspective. The editorialists postulate that there are many mechanisms that could account for improved glucose metabolism in patients treated with inhibitors of the renin–angiotensin system, and the DREAM authors attempt to explain why, despite these mechanisms, ramipril did not decrease the incidence of new-onset diabetes. We believe one element that has been overlooked here and elsewhere is potassium homeostasis. It is accepted that thiazide diuretics decrease potassium levels, and a recent analysis of 59 clinical trials using diuretics identified a tight inverse correlation between changes in serum potassium levels and blood glucose levels. It thus seems reasonable that a similar, directionally opposite relationship could partially explain the protective effect of inhibitors of the renin–angiotensin system. If this is the case, then it would be valuable for readers to be cognizant of changes in serum potassium in patients treated with ramipril. Such data could help clinicians better understand the DREAM results with regard to the effect of ramipril.

Benjamin J. Epstein, Pharm.D.
Rhonda M. Cooper-DeHoff, Pharm.D.
University of Florida
Gainesville, FL 32610
epstein@cop.ufl.edu


THE AUTHORS REPLY: Bangalore and Messerli ask whether changes in blood pressure could explain the greater regression to normoglycemia with ramipril than with placebo. After adjustment for the change in blood pressure that occurred during the course of the trial, ramipril significantly increased the likelihood of regression to normoglycemia (adjusted hazard ratio for the ramipril group, 1.12; 95% confidence interval [CI], 1.03 to 1.23; P=0.008), indicating that the effect of ramipril on blood pressure does not explain its effect on regression of dysglycemia.

Potter and LeLorier ask whether the use of ARBs for uncontrolled hypertension could have affected our results. Only 286 of our 5269 patients (5.4%) received ARBs at baseline, with 5.6% receiving them at 2 years and and 7.9% receiving them at the end of the study. Adjustment for this small difference does not affect the primary outcome (adjusted hazard ratio for the ramipril group, 0.90; 95% CI, 0.80 to 1.02). Neither the effect of ramipril on the primary outcome (Fig. 1A) nor the effect of ramipril on regression to normo-
glycemia (Fig. 1B) was influenced by the presence or absence of rosiglitazone.

Epstein and Cooper-DeHoff ask whether the effect of potassium could have influenced our results. Unfortunately, we did not measure potassium levels after randomization.

Jackie Bosch, M.Sc.
Salim Yusuf, D.Phil., M.D.
Hertzel Gerstein, M.D.
McMaster University
Hamilton, ON L8S 1C7, Canada
dream@cardio.on.ca

Prevention of Meningococcal Disease

TO THE EDITOR: In his article on the prevention of meningococcal disease, Gardner (Oct. 5 issue) notes that both active and passive smoking may be risk factors for the disease but does not address the issue of exposure to smokers as differentiated from exposure to smoke. Contact with smokers rather than smoke is now recognized as a critical risk factor, most likely owing to higher rates of carriage and coughing among smokers.2-4 The case–control study by Coen et al. of 144 teenage survivors of meningococcal disease showed that older teens are more at risk from exposure to smokers than to smoke.4 In two studies, significant odds ratios for meningococcal disease (3.8 and 9.1) were reported for children whose mothers smoked.2,3 A remarkable 37% of cases were reported as being attributable to exposure to smokers in one of the studies,2 and we estimate that 60% of cases were attributable to exposure to smokers in the other study.3 Public health messages should underscore the need to stop smoking, not merely the need to limit smoking to outside the home.

Gardner also mentions that deficiency in the terminal complement pathway is responsible for increased risk but does not mention the much more common deficiency of mannose-binding lectin. Recent work has demonstrated that a deficiency of this protein, which is responsible for activation of the alternative complement pathway, is a critical factor. The frequency of homozygous variants was significantly higher among 194 children with meningococcal disease than among control subjects (odds ratio, 6.5; 95% con-
fidence interval, 2.0 to 27.2). The fraction of cases attributed to mannose-binding lectin variants was 32%. 5

Robert Booy, M.D.
Mary Iskander, M.B., B.S.
National Centre for Immunisation Research and Surveillance
Westmead 2145, Australia
robertb2@chw.edu.au

Russell Viner, M.D.
University College London Hospitals
London NW1 2PG, United Kingdom
Dr. Booy reports receiving research support from CSL, Sanofi, Roche, and Wyeth.


TO THE EDITOR: The article by Gardner identifies close contacts of an index patient with meningococcal disease who would need chemoprophylaxis. These contacts include people who have been directly exposed to the patient’s oral secretions, including through kissing. However, saliva itself is thought to have an inhibitory effect on meningococcus, probably owing to the presence of other oropharyngeal flora. 1 This has been supported by a study of 258 college students in the United Kingdom, in which meningococcal carriage in the tonsils, nasopharynx, and saliva was examined. The overall carriage rate was 34.9% (90 of 258 students), but only one swab from saliva (0.4%) was positive for meningococcus. 1 The Australian national guidelines now recommend that chemoprophylaxis not be used purely on the basis of activities such as nonintimate kissing (even on the mouth) or sharing of food, drinks, cigarettes, or bongs. 2 However, intimate kissing, especially with multiple partners, is a risk factor for meningococcal disease. 3 Certainly, as Gardner suggests, endotracheal intubation and mouth-to-mouth resuscitation would constitute sufficient exposure to warrant chemoprophylaxis; however, the risk associated with these activities is probably related to aerosolization of meningococci rather than exposure to saliva.

Sanjaya N. Senanayake, M.B., B.S.
Canberra Hospital
Canberra 2606, Australia
sanjaya.senanayake@act.gov.au


THE AUTHOR REPLIES: Clinical Practice articles in the Journal focus on management considerations, and space constraints allow for only a limited discussion of basic science and public health issues. Dr. Booy and colleagues provide additional insights in both areas. First, they note that in addition to the long-recognized increased risk of invasive meningococcal disease from active and passive smoking, there is an increased risk from exposure to people who smoke (presumably because these people have increased colonization with Neisseria meningitidis). It is notable that, despite the recognition of smoking as a risk factor for both invasive meningococcal disease 4 and invasive pneumococcal disease, 2 the Advisory Committee on Immunization Practices of the Centers for Disease Control and Prevention has failed to include smokers among the high-risk groups for which immunization against these diseases is recommended. 1,3

Second, their mention of a deficiency of mannose-binding lectin in the pathophysiology of invasive meningococcal disease is welcome.

Dr. Senanayake discusses the Australian guidelines for chemoprophylaxis, which define persons considered to be close contacts of a patient with meningococcal disease more specifically than do the U.S. guidelines. Noting that saliva is a much less likely source of N. meningitidis than material taken from the nasopharynx or tonsils, the Australians do not consider “nonintimate kissing” or shared ingested or smoked materials to constitute significant exposure. Accordingly, the Australian recommendations for chemoprophylaxis are more restrictive than the U.S recommenda-
Medical Mystery: Abnormal Abdominal Radiograph — The Answer

TO THE EDITOR: The medical mystery in the December 7, 2006, issue involved a 50-year-old woman who presented to the emergency department with obtundation and hypotension. An abdominal radiograph showed gas throughout the right kidney (Fig. 1A). Computed tomography (CT) of the abdomen revealed extensive destruction of the right renal parenchyma with associated gas, as well as gas in the retroperitoneal tissues (Fig. 1B). The patient's serum glucose level at presentation was 607 mg per deciliter (33.7 mmol per liter), and her glycated hemoglobin value was 12.2%. She did not have diabetic ketoacidosis. A diagnosis of emphysematous pyelonephritis in the setting of diabetes mellitus was made. The patient underwent urgent right nephrectomy, and *Escherichia coli* was cultured from the surgical site and from the blood. She had an uneventful recovery, with normalization of her renal function. Her newly diagnosed diabetes is well controlled through insulin therapy, and she is doing well.

Masahiko Kurabayashi, M.D., Ph.D.
Kumiko Tajima, M.D., Ph.D.
Gunma University Graduate School of Medicine
Gunma 371-8511, Japan
mkuraba@med.gunma-u.ac.jp


Editor's note: We received 1162 responses to this medical mystery — 55% from physicians in practice, 19% from physicians in training, 13% from medical students, and 13% from other readers. Responses were received from 82 countries. Many of the responses reflect a team effort — such as the results of a discussion of the case during a teaching conference.

Forty percent of the respondents correctly identified gas associated with the right kidney or emphysematous pyelonephritis. Eleven percent suggested a gallbladder disorder such as emphysematous cholecystitis, 12% suggested other infections (e.g., hydatid cyst or hepatic abscess), another 12% suggested cancer (e.g., renal, adrenal, or hepatic), and 19% suggested a variety of diagno-
Correspondence

Case 32-2006: A Girl with Fever after a Visit to Africa

To the Editor: In the Case Record regarding severe falciparum malaria in a 3-year-old girl (Oct. 19 issue), Fraser et al. discuss the use of exchange transfusion and attribute the patient’s clinical improvement to this treatment. Current evidence does not support this conclusion: a meta-analysis of eight comparative trials showed no significant benefit of adjunctive exchange transfusion over chemotherapy alone. Although there was systematic bias toward use of exchange transfusion in patients with severe malaria, subgroup analysis showed no additional benefit at any level of parasitemia.1

Treatment of severe malaria with artesunate, as compared with quinine, has been shown to reduce mortality by 35%.2 Unlike intravenous quinidine, artesunate is easy to administer and is well tolerated. Artesunate has not yet been approved by the Food and Drug Administration. Therefore, it is ironic that the drug is being used to great effect in much of Asia and Africa, even though in the United States, patients with severe falciparum malaria are denied the most effective treatment and may be exposed to unproven, potentially dangerous interventions.3

John Williams, M.R.C.P.
Tan Tock Seng Hospital
Singapore 308433, Singapore
john_williams@ttsh.com.sg


To the Editor: In the Case Record, a 3-year-old child with severe malaria was treated with intravenous quinidine and exchange transfusion. Both the treatment protocols are potentially life-threatening, and better alternatives are available. A review of the currently available data from trials that have compared quinine with artesunate suggests a 9% absolute reduction in the risk of death with the use of artesunate (number of patients who would need to be treated to prevent one death, 11), where available.1 Also, the use of artesunate is associated with lower infusion volumes and can potentially reduce the incidence of fluid overload, which is a common complication in children. Furthermore, the parasite clearance time is faster with artesunate than with quinine, and its use might have obviated the observed increase in parasitemia in this case.2 Moreover, artesunate is not associated with hypoglycemia and cardiac toxicity, both of which are commonly encountered with the use of quinidine. In a systematic review, exchange transfusion was not associated with a higher survival rate than was antimalarial chemotherapy alone, and the procedure is fraught with complications.2 In fact, the recent guidelines of the World Health Organization (WHO) do not endorse the use of exchange transfusion, even in patients with severe parasitemia.3

Ritesh Agarwal, M.D., D.M.
Rajagopala Srinivas, M.D.
Alok Nath, M.D.
Postgraduate Institute of Medical Education and Research
Chandigarh 160012, India
ritesh@indiachest.org


To the Editor: One must not generalize from the experience with the patient in the Case Record, especially in the developing world. The discussants state that there are risks involved in exchange transfusion but fail to mention an im-

ses, including pneumatosis coli, renal-vein thrombosis, intussusception, toxic megacolon, volvulus, pancreatic cyst, or a fecolith. The remaining 6% of respondents suggested a bezoar or an intra-abdominal pregnancy, including the possibility of a lithopedion.

To the Editor: In the Case Record regarding severe falciparum malaria in a 3-year-old girl (Oct. 19 issue), Fraser et al. discuss the use of exchange transfusion and attribute the patient’s clinical improvement to this treatment. Current evidence does not support this conclusion: a meta-analysis of eight comparative trials showed no significant benefit of adjunctive exchange transfusion over chemotherapy alone. Although there was systematic bias toward use of exchange transfusion in patients with severe malaria, subgroup analysis showed no additional benefit at any level of parasitemia.2

Treatment of severe malaria with artesunate, as compared with quinine, has been shown to reduce mortality by 35%.3 Unlike intravenous quinidine, artesunate is easy to administer and is well tolerated. Artesunate has not yet been approved by the Food and Drug Administration. Therefore, it is ironic that the drug is being used to great effect in much of Asia and Africa, even though in the United States, patients with severe falciparum malaria are denied the most effective treatment and may be exposed to unproven, potentially dangerous interventions.3

John Williams, M.R.C.P.
Tan Tock Seng Hospital
Singapore 308433, Singapore
john_williams@ttsh.com.sg


To the Editor: In the Case Record, a 3-year-old child with severe malaria was treated with intravenous quinidine and exchange transfusion. Both the treatment protocols are potentially life-threatening, and better alternatives are available. A review of the currently available data from trials that have compared quinine with artesunate suggests a 9% absolute reduction in the risk of death with the use of artesunate (number of patients who would need to be treated to prevent one death, 11), where available.1 Also, the use of artesunate is associated with lower infusion volumes and can potentially reduce the incidence of fluid overload, which is a common complication in children. Furthermore, the parasite clearance time is faster with artesunate than with quinine, and its use might have obviated the observed increase in parasitemia in this case.2 Moreover, artesunate is not associated with hypoglycemia and cardiac toxicity, both of which are commonly encountered with the use of quinidine. In a systematic review, exchange transfusion was not associated with a higher survival rate than was antimalarial chemotherapy alone, and the procedure is fraught with complications.2 In fact, the recent guidelines of the World Health Organization (WHO) do not endorse the use of exchange transfusion, even in patients with severe parasitemia.3

Ritesh Agarwal, M.D., D.M.
Rajagopala Srinivas, M.D.
Alok Nath, M.D.
Postgraduate Institute of Medical Education and Research
Chandigarh 160012, India
ritesh@indiachest.org


To the Editor: One must not generalize from the experience with the patient in the Case Record, especially in the developing world. The discussants state that there are risks involved in exchange transfusion but fail to mention an im-
The nonimmune patient with 5% parasitemia, described in the Case Record, should have received parenteral treatment from the outset. Furthermore, there was no need to give the second antimalarial agent, doxycycline (which is contraindicated in children), with the quinine. Blood films in the first 24 hours with a rise in the parasite count are not indicative of therapeutic failure and may indicate a more favorable outcome.

The article indicates that more mature forms and schizonts are not found in the peripheral blood in falciparum malaria. They can indeed be present in late ring forms or more mature trophozoite forms and indicates a poor prognosis, and hemozoin found in neutrophils, which also indicates a poor outcome. Although in malaria there may be evidence of activation of coagulation, thrombocytopenia is more often caused by widespread sequestration than by disseminated intravascular coagulation.

David L. Blazes, M.D., M.P.H.
Naval Medical Research Center Detachment
Lima, Peru
blazes@nmrcd.med.navy.mil

John W. Sanders, M.D., M.P.H.
National Naval Medical Center
Bethesda, MD 20889

Mark S. Riddle, M.D., M.P.H.
Naval Medical Research Center
Silver Spring, MD 20889

The views expressed in this letter are those of the authors and do not necessarily reflect the views or policies of the Department of the Navy or the Department of Defense.


TO THE EDITOR: The nonimmune patient with 5% parasitemia, described in the Case Record, should have received parenteral treatment from the outset. Furthermore, there was no need to give the second antimalarial agent, doxycycline (which is contraindicated in children), with the quinine. Blood films in the first 24 hours with a rise in the parasite count are not indicative of therapeutic failure and may indicate a more favorable outcome.

The article indicates that more mature forms and schizonts are not found in the peripheral blood in falciparum malaria. They can indeed be present in late ring forms or more mature trophozoite forms and indicates a poor prognosis, and hemozoin found in neutrophils, which also indicates a poor outcome. Although in malaria there may be evidence of activation of coagulation, thrombocytopenia is more often caused by widespread sequestration than by disseminated intravascular coagulation.

Geoffrey Pasvol, M.D., Ph.D.
Imperial College London
London HA1 3UJ, United Kingdom

Pasvol suggests that the patient should have received parenteral therapy from the outset on the basis of her parasitemia of 5%. Whereas the review he cites suggests the use of 2% parasitemia as a cutoff, we have used higher levels of parasitemia in making this determination, in part because of the toxicity of parenteral quinidine as compared with oral quinine. We recognize that doxycycline is not routinely used in children under the age of 8 years because of dental stain-
ing. Oral quinine plus clindamycin is an appropriate antimalarial combination for this age group, but oral clindamycin has significant gastrointestinal toxicity that may exacerbate the nausea and vomiting associated with malaria, thereby compromising oral therapy. Doxycycline is used to treat life-threatening infections in children of all ages, dosing is convenient, and short courses of therapy (as used in our case) carry a minimal risk of dental staining.\(^2\)

We agree with Blazes et al. that the real value of exchange-transfusion therapy in severe malaria is unknown and that exchange transfusion should not be universally applied until any benefit is proven to outweigh the risks. In the case we reported, the decision to perform an exchange transfusion was based on individual risk–benefit considerations for that single child. Indeed, the recent WHO guidelines (referred to by Agarwal et al.) are noncommittal on the topic of exchange transfusion. The guidelines state, “There is no consensus on the indications, benefits and dangers involved, or on practical details such as the volume of blood that should be exchanged. It is therefore not possible to make any recommendation regarding the use of exchange blood transfusion.”\(^3\)

Pasvol makes the point that the appearance of mature schizonts in the peripheral blood is an ominous sign in Plasmodium falciparum infection. In our patient, hemozoin was seen neither in the parasite nor in the neutrophils, and we agree that its presence portends a bad prognosis. We also agree with him that thrombocytopenia results principally from sequestration and rosette formation rather than from disseminated intravascular coagulation. According to the WHO, clinically significant disseminated intravascular coagulation develops in less than 5% of patients with severe malaria.\(^3\)

Iain P. Fraser, M.B., Ch.B., D.Phil.
Massachusetts General Hospital
Boston, MA 02114

Christine M. Cserti, M.D.
University Health Network
Toronto, ON M5G 2C4, Canada

Walter H. Dzik, M.D.
Massachusetts General Hospital
Boston, MA 02114


TO THE EDITOR: The Maryland Poison Center was called about a 49-year-old, usually calm prison inmate who was described as being “red-eyed,” “loony,” “combative,” and “intoxicated, lecturing everyone about life.” Other inmates and staff reported seeing this prisoner drinking from a gallon container of Purell hand sanitizer over the course of the evening. It was discovered that this sanitizer contains 62% ethanol by weight (more than 70% alcohol by volume). The inmate’s blood alcohol level was found to be 335 mg per deciliter. It was later confirmed that he had not consumed any other forms of ethanol or other illicit substances. The patient was treated with fluid repletion and haloperidol, with no complications.

The Centers for Disease Control and Prevention has recommended that before and after having any direct contact with patients, health care workers use an alcohol-based hand sanitizer “if hands are not visibly soiled”\(^1\) or wash their hands with an antimicrobial soap and water. The Joint Commission on Accreditation of Health Care Organizations and the Federal Bureau of Prisons have adopted these recommendations.\(^2,3\) As compared with the regular use of soap and water, alcohol-based hand rubs with moisturizer have been associated with a significant reduction in the number of microorganisms on skin; the sanitizers also act quickly and cause less skin irritation.

Alcohol-based hand sanitizers are widely used in the United States as low-viscosity rinses, gels, or foams. They contain 60 to 95% ethanol or isopropanol. Ethanol has greater activity against
viruses than does isopropanol, and the ethanol-based formulations are used much more commonly in the United States than are the isopropanol-based formulations, under trade names such as Avagard D, Avant, Nexcare, Prevacare, Germ-X, and Purell. Many of the ethanol-based sanitizers also contain small amounts of polyethylene glycol or isopropanol.

Although the potential fire hazard associated with such cleansers has been addressed through use and storage recommendations, there are no similar recommendations for the potential misuse of the cleansers as intoxicants. Given the widespread use and national endorsement of these products in the United States, health care providers and administrators in hospitals and correctional facilities should be aware of the potential misuse of the products as intoxicants and should take steps to minimize such use in high-risk populations.

Suzanne Doyon, M.D.
Maryland Poison Center
Baltimore, MD 21201
sdoyon@rx.umaryland.edu

Christopher Welsh, M.D.
University of Maryland School of Medicine
Baltimore, MD 21201


Intoxication of a Hospitalized Patient with an Isopropanol-Based Hand Sanitizer

TO THE EDITOR: A 43-year-old man with alcoholism was admitted to the hospital with chest pain, for which the workup was unremarkable. At discharge, the patient became acutely hypotensive and delirious. He was afebrile, oxygenating well, and had a nonfocal neurologic examination. Intravenous fluids and vasopressors were administered. The results of routine laboratory tests were normal, as were the results of arterial blood gas and serum ethanol measurements, toxicology screening, blood and urine cultures, and computed tomography of the head. Urinalysis showed a trace of acetone. The following day, the patient was hemodynamically stable, but his mental status did not improve. Because of a sweet, ketotic odor in the room, tests of serum isopropyl alcohol (isopropanol) and acetone levels were ordered. Before those results were obtained, the patient was seen in the bathroom drinking the alcohol-based hand wash from its dispenser.

The patient’s isopropanol level was 13.6 mg per deciliter, and his acetone level was 269.4 mg per deciliter (normal range, 0 to 1.9 for both). When asked why he ingested the hand cleaner, he pointed to the label, which read, “Active ingredient 63% v/v isopropyl alcohol.” He explained that this percentage is higher than that in vodka.

Ingestion of approximately 200 ml of isopropanol can be lethal owing to depression of both the central nervous system and myocardial function. A plasma concentration above 400 mg per deciliter is considered to be life-threatening. Unlike methanol and ethylene glycol, isopropanol is more toxic than its metabolites; hence, alcohol dehydrogenase inhibitors should not be given. Hemodialysis removes both isopropanol and acetone. Since in our patient, most of the isopropanol had already been converted to acetone (probably owing to the first-order kinetic metabolism), we decided not to pursue dialysis and instead continued with supportive therapy. He recovered fully.

Ingestion of methanol, ingestion of ethylene glycol, and ingestion of isopropanol all lead to an elevated plasma osmolal gap. However, isopropanol does not cause metabolic acidosis with an elevated anion gap. Primary alcohols are metabolized to aldehydes, followed by oxidation to
Correspondence

Carboxylic acids; deprotonation then causes acidosis with conjugate bases that elevate the anion gap. Secondary alcohols, such as isopropanol, oxidize to ketones, such as acetone, which under physiologic conditions cannot be further oxidized to acid.

Isopropanol-containing hand sanitizers are ubiquitous. Physicians should be aware of the potential for isopropanol intoxication, especially among alcoholics, in the hospital setting. Perhaps changing the description on the container from isopropyl alcohol to isopropanol or propane-2-ol would decrease the attraction of these hand sanitizers for potentially dangerous abuse.

Ashkan Emadi, M.D., Ph.D.
LeAnn Coberly, M.D.
University of Cincinnati
Cincinnati, OH 45267-0557
emadia@email.uc.edu


Correspondence Copyright © 2007 Massachusetts Medical Society.

INSTRUCTIONS FOR LETTERS TO THE EDITOR

Letters to the Editor are considered for publication, subject to editing and abridgment, provided they do not contain material that has been submitted or published elsewhere. Please note the following: • Letters in reference to a Journal article must not exceed 175 words (excluding references) and must be received within three weeks after publication of the article. Letters not related to a Journal article must not exceed 400 words. All letters must be submitted over the Internet at http://authors.nejm.org.
• A letter can have no more than five references and one figure or table. • A letter can be signed by no more than three authors.
• Financial associations or other possible conflicts of interest must be disclosed. (Such disclosures will be published with the letters. For authors of Journal articles who are responding to letters, this information appears in the original articles.) • Include your full mailing address, telephone number, fax number, and e-mail address with your letter.

Our Web site: http://authors.nejm.org

We cannot acknowledge receipt of your letter, but we will notify you when we have made a decision about publication. Letters that do not adhere to these instructions will not be considered. Rejected letters and figures will not be returned. We are unable to provide prepublishation proofs. Submission of a letter constitutes permission for the Massachusetts Medical Society, its licensees, and its assignees to use it in the Journal’s various print and electronic publications and in collections, revisions, and any other form or medium.
BOOK REVIEWS

PSYCHIATRIC ASPECTS OF HIV/AIDS

A QUARTER OF A CENTURY FROM THE BEGINNING of the epidemic of the human immunodeficiency virus (HIV) and AIDS, in a world where 40 million people are estimated to be living with the virus and where, in 2005, 5 million new infections occurred, this book is a welcome addition to our understanding of the manifestations and challenges of HIV. The psychiatrists, psychologists, epidemiologists, neurologists, and other contributors provide a thoughtful, evidence-based, practical resource for those working in the field.

The book moves effortlessly from basic science and clinical research, through the epidemiology of HIV (incidence and risk factors for mental disease), to a compilation of practical tools to measure, monitor, and treat the psychiatric aspects of HIV and AIDS. The specifics of HIV science, the epidemiology of HIV infection in the United States, and the psychiatric disorders, predispositions, and diseases associated with various stages of HIV in different populations are clearly outlined.

The populations discussed, including women, children, homosexual men, the homeless, and prisoners, are specific to the United States, but the information the chapters contain is applicable in many parts of the world. Of practical use to psychiatrists and others in HIV care will be the section on coexisting conditions, including the stress–distress spectrum and adjustment disorders, anxiety disorders, mood disorders, psychotic disorders, sleep disorders, and the difficult areas of cognitive and substance-abuse disorders. Accurate and accessible information on psychotropic drug interactions with antiretroviral medications, and on detailed management of coexisting psychiatric conditions in medically ill patients, will provide a valuable reference in many settings. Diagnostic tools, including psychiatric assessment, psychological and neuropsychological testing, and the use of electrophysiology and brain mapping, are also covered. All readers will appreciate the excellent reviews of the literature on the epidemiology of coexisting psychiatric conditions, postulated pathophysiological processes, and the available data on evidence for diagnosis and treatment.

Since the beginning, the HIV epidemic has challenged scientists and clinicians to look beyond a narrow specialty; to broaden their horizons into new communities; to face multiple personal, community, and societal challenges; and to overcome personal biases and assumptions. This book offers essays on legal, health care provider, and policy issues, covering topics such as complementary holistic medicine, suicide, and end-of-life care. It also presents challenging and diverse views on biosocial, psychiatric, and psychological aspects of care and prevention and on individual and community responses.

Themes running through many chapters are the need for multidisciplinary approaches and for the early recognition of psychiatric and psychological factors in HIV medicine. Many authors stress the evidence that appropriate and timely consideration of psychiatric, psychological, and social needs is required to achieve adherence to combination antiretroviral therapy and to facilitate better outcomes for patients with HIV. Throughout the book, the integration of assessment and an early response to psychiatric and psychological disorders is identified as a key criterion for success in care models. Another innovative and consistent theme is the incorporation of psychological and psychiatric responses in the prevention of HIV infection. Many authors argue for combination prevention programs that include biological, educational, and behavioral interventions adapted for individuals, partnerships, and communities.

There are some disappointments in this book. As always, epidemiologic data that were accurate at the time of writing were rapidly overcome by the speed of this epidemic. Some areas of basic science are overlooked, such as the new understandings of the early loss of gut-associated lymphoid tissue in the pathogenesis of immune dysfunction. Refer-
ences to evidence-based interventions for prevention, including motivational interviewing, are followed by comments with details that go beyond the scope of the text. However, these and other small disappointments — including minor repetitions and inconsistencies and a U.S.-centric perspective — do not detract from the usefulness of this book. Would I purchase it? Definitely. Will I read it? Yes, frequently. Would I recommend it? Absolutely.

Anne M. Mijch, M.B., B.S.
Alfred Hospital
3181 Melbourne, Australia
anne.mijch@med.monash.edu.au

THE CLINICAL NEUropsychiatry of Stroke: Cognitive, Behavioral, and Emotional Disorders Following Vascular Brain Injury


The usual care of patients affected by stroke is focused on limiting orremediating impairments in sensorimotor function and on facilitating improvements in the performance of activities of daily living. This focus, however, affords too limited a view of the clinically important sequelae of stroke. Cognitive, emotional, and behavioral disorders complicate sensorimotor and functional recovery and are themselves substantial sources of further complication and suffering for stroke survivors and their families. Nonetheless, the neuropsychiatric consequences of stroke remain uncommon subjects of evaluation and treatment in most general medical settings.

In this second edition of The Clinical Neuropsychiatry of Stroke, Robert G. Robinson discusses in detail the phenomenology, neurobiology, and treatment of post-stroke neuropsychiatric disorders, particularly depression. Psychological theories of post-stroke depression are considered in historical context as well as in response to published debates generated by Robinson’s own work. He thoughtfully acknowledges that psychosocial factors are involved in the development of post-stroke depression and other stroke-related emotional and behavioral problems. However, he also makes a strong argument — with volumes of supporting data — for the view that neurobiologic distur-

bances are primary contributors to post-stroke neuropsychiatric disorders.

Working from a review of his own work as well as the world literature, Robinson makes several important and clinically relevant observations. Post-stroke depressions are phenomenologically similar to idiopathic depressive disorders and can be identified using standard psychiatric diagnostic criteria. Depression, particularly in the period shortly after a stroke, is a major contributor to cognitive and functional impairments, rather than a purely psychological response to such problems. Although the issue of the association between lesion location and post-stroke depression is contentious, anterior left-hemisphere stroke is identified as a clear risk factor for episodes of post-stroke major depression. Serotonergic depression and noradrenergic dysfunction resulting from such lesions are further identified as contributors to post-stroke emotional disturbances and as targets for pharmacotherapy. Finally, the treatment of post-stroke depression not only improves recovery from this condition but also improves stroke outcome more generally. Conversely, failure to treat post-stroke depression not only impedes recovery from stroke but also increases long-term post-stroke mortality.

Robinson concludes that depression is a neurobiologically understandable consequence of stroke, one for which treatments are both available and necessary. The evidence presented in this book...
indicates that post-stroke depression responds to pharmacotherapies in a manner similar to that of depressions resulting from other causes, including idiopathic ones. By contrast, and in opposition to purely “psychological” views of this condition, cognitive behavioral psychotherapy — an intervention widely regarded as effective for the treatment of idiopathic major depressive disorder — is identified as an ineffective treatment for post-stroke depression.

Other post-stroke neuropsychiatric disturbances, including mania, anxiety disorders, irritability and aggression, psychosis, and pathologic laughing and crying, among others, are also addressed in this book, albeit more briefly. Vascular cognitive impairments arising independently of post-stroke emotional and behavioral disorders are not addressed, but Robinson acknowledges this issue in the final chapter and suggests that vascular cognitive impairments merit consideration in a separate volume.

In this edition, Robinson demonstrates clearly and convincingly that a neurobiopsychosocial approach to the study and treatment of post-stroke neuropsychiatric disorders affords unprecedented opportunities to improve the lives of stroke survivors and their families. As a clinician, educator, and scientist involved daily in stroke neurorehabilitation, I found that reading this book improved my understanding of the neuropsychiatry of stroke and is already changing for the better my practice, teaching, and research. This book is essential reading for anyone interested in stroke rehabilitation, and it deserves a place in the libraries of all clinics and clinicians involved in the care of people affected by stroke.

David B. Arciniegas, M.D.
University of Colorado School of Medicine
Denver, CO 80262
david.arciniegas@uchsc.edu

McAlpine’s Multiple Sclerosis


Multiple sclerosis is the most common inflammatory disorder of the central nervous system and a leading cause of disability in young adults. It usually begins with a relapsing and remitting course but eventually evolves into a chronic, progressive disease with an accumulation of neurologic dysfunction, permanent deficits, and increasing disability. Life expectancy is shortened. For these reasons, multiple sclerosis is a heavy burden on patients, their families and caregivers, the health system, and society.

McAlpine’s Multiple Sclerosis has been the leading single-book source of information on this disease since it was first published in 1985, based on works from as early as 1955 by Douglas McAlpine, Nigel Compston (the father of the current edition’s editor-in-chief, Alastair Compston), and Charles Lumsden. For this edition, Alastair Compston enlisted leading experts in multiple sclerosis research. Over a period of 3 years, these authors — world-renowned specialists in neurobiology, neuroimmunology, neuropathology, neuroimaging, genetics, neurophysiology, and neurology — produced an encyclopedic reference work on multiple sclerosis. What sets this book apart from others with several authors is the coherence in the approach, style, and layout of the chapters. The book’s 19 chapters are organized into five sections, discussing the history of multiple sclerosis, its cause and course, its clinical features and diagnosis, its pathogenesis, and its treatment.

The first section provides an excellent account of the history of the disease. The second includes well-written, comprehensive, analytical commentaries on epidemiology, geographic distribution, natural history, and prognostic factors. In the third section, readers will find extremely valuable descriptions of the clinical features, the diagnostic approach, the invaluable contribution of magnetic resonance imaging, and the broad differential diagnosis of a bewildering array of disorders that mimic multiple sclerosis. The chapters in the fourth section discuss the cause, pathology, and pathogenesis of the disease; the final section discusses the ever-increasing therapeutic armamentarium. Included in this section is a relevant and extremely useful critical appraisal of trial methodology. Management of multiple sclerosis obviously mandates a multidisciplinary approach and encompasses neurorehabilitation at its core, apart from the use of drugs to modify the disease and alleviate its symptoms.

The introduction of disease-modifying drugs in the 1990s stimulated and intensified research into the causes of multiple sclerosis, which has in turn generated a wealth of information challeng-
ing old dogmas and changed paradigms. Among the areas subject to these changing points of view are the growing acknowledgment of the pathologic heterogeneity of the disease, the role of axonal damage (in addition to the cardinal feature of demyelination), and the involvement of the cortex and gray matter. Gratifying to note are the designer drugs currently in the pipeline that hold promise for further improvement in the treatment of this disabling disease.

All this knowledge has been admirably captured in this excellent book. The factual contents are greatly enhanced by the inclusion of highly instructive, aesthetically pleasing, almost artistic illustrations. The scientific literature incorporated into the text is current almost up to the book’s date of publication. I congratulate the authors and editors and in particular their spiritus rector, Alastair Compston, on their achievements. I can enthusiastically recommend this latest edition of McAlpine’s Multiple Sclerosis.

Hans-Peter Hartung, M.D.
Heinrich-Heine University
D-40225 Düsseldorf, Germany
hans-peter.hartung@uni-duesseldorf.de

TEXTBOOK OF NEURAL REPAIR AND REHABILITATION, VOLUME I: NEURAL REPAIR AND PLASTICITY

TEXTBOOK OF NEURAL REPAIR AND REHABILITATION, VOLUME II: MEDICAL NEUROREHABILITATION


The discipline of neuroscience continues to expand. From early anatomy-based investigations through the cellular and molecular biology revolutions, neuroscience has stayed at the forefront and at times has even led to paradigm shifts in biology. One emerging area that moves away from reductionist biology and toward the view of the organism as a whole is neurorehabilitation. Although often considered an area that is not grounded in “hard science,” rehabilita-

tion has been transformed over the past two decades with the integration of advances in cellular and molecular neurobiology, bioengineering, and computer science. Bringing such disparate disciplines together coherently is a challenge, and the editors and authors of this substantial two-volume textbook have risen to it with remarkable success.

Selzer and colleagues have compiled contributions from leaders in the field of neural repair and rehabilitation. Whereas each of the two volumes can stand alone, together they cover a breadth of information that is unavailable in any other single source.

The first volume concerns the biology of neural function and dysfunction. Not surprisingly, several sections discuss plasticity in the central nervous system and the molecular aspects of axonal regeneration and pathfinding. Remarkably, little space is devoted to the current excitement about neural stem cells or the use of cellular therapies in the nervous system, but the reason for this seeming oversight may be that these fields are in their infancy. The final section of this volume, which addresses the application of translational research to neural injury, is not well integrated with the rest of the volume and, disappointingly, does not address the translation of basic advances to clinical applications. The exception to this criticism is the chapter on cell therapies by Olle Lindvall and Peter Hagell. A broader perspective would have been most welcome.

The second volume, a comprehensive overview of neurorehabilitation, builds from the knowledge base of the first volume. It includes sections on technological approaches as diverse as functional mapping, functional electrical stimulation, wheelchair design, and cell transplantation. There are also discussions of symptom-specific and disease-specific rehabilitation. The logic behind this organization is clear; the first volume is directed largely at basic scientists and the second at practitioners of neurorehabilitation. Still, there is the feeling that an opportunity for synthesis has been missed.

Even so, the quality and utility of these volumes are considerable. The individual contributions in the two volumes are uniformly excellent. The editors have done a good job of bringing leaders of the different disciplines together, and the result is a book that will be a valuable resource to neu-
Similarly, rate or at (www.nejm.org/meetings). The listings can be viewed in another ment. We regret that we are unable to publish all notices submitted for publication should contain a mailing address and telephone number of a contact person or department. We regret that we are unable to publish all notices received. Notices also appear on the Journal’s Web site (www.nejm.org/meetings). The listings can be viewed in their entirety or searched by location, month, or key word.

INFANTILE SEIZURE SOCIETY

The 10th Annual Meeting, entitled “International Symposium on Biology of Seizure Susceptibility,” will be held in Tokyo, April 7 and 8.

Contact Dr. Takao Takahashi, Dept. of Pediatrics, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan; or call (81) 3 3563 3815; or fax (81) 3 335 7022; or e-mail ttakahashi@x3.keio.jp; or see http://www.iss-jpn.info.

10TH ANNUAL OBSTETRIC ULTRASOUND: SETTING THE STANDARD FOR 2007

The meeting will be held in Toronto, Feb. 16–18.

Contact Elizabeth Gan, CME, Department of Obstetrics and Gynecology, Mount Sinai Hospital, 600 University Ave., Room 1255, Toronto, ON M5G 1X5, Canada; or call (416) 586-4800, extension 2489; or e-mail egan@mtsinai.on.ca; or see http://www.mtsinai.on.ca/seminars/cc.

MAYO CLINIC COLLEGE OF MEDICINE

The following courses will be offered in Amelia Island, FL, unless otherwise indicated: “Electromyography and Electrocen- cephalography in Clinical Practice” (March 7–11); “17th Annual Clinical Nutrition” (Savannah, GA, April 11–15); “Clinical Reviews and Primary Care Update” (June 17–22); “17th Annual Hema- tology Oncology Reviews” (July 31–Aug. 4); “Advances and Chang- ing Trends in Medicine” (Miami Beach, FL, Aug. 6–8); and “Cancer and Cardiology in Primary Care” (Asheville, NC, Oct. 11–13).

Contact Mayo Clinic College of Medicine, School of Con- tinuing Medical Education, Davis Bisig, 172-W, 4500 San Pablo Rd., Jacksonville, FL 32224; or call (800) 462-9633; or fax (904) 953-2954; or e-mail cme-jax@mayo.edu; or see http://www.mayo.edu/cme.

UNIVERSITY OF TORONTO

The following courses and conferences will be held in To- ronto: “Diabetes Update 2007: Innovative and Emerging Therapies in the Management of Type 1 and Type 2 Diabetes” (March 23); “Update in General Surgery 2007” (March 29–31); “18th Annual Jack Crawford Day Pediatric Ophthalmology Confer- ence: Strabismus for the Community” (March 30); “7th Annual Toronto Breast Surgery Symposium” (May 3); and “37th Annual Aesthetic Plastic Surgery Symposium” (May 4 and 5).

Contact Office of Continuing Education and Professional Development, Faculty of Medicine, University of Toronto, 500 University Ave., Suite 650, Toronto, ON M5G 1V7, Canada; or call (888) 512-8173 or (416) 978-2719; or fax (416) 966-7028; or e-mail ce.med@utoronto.ca; or see http://www.cme.utoronto.ca.

INTERNATIONAL SOCIETY FOR CLINICAL DENSITOMETRY

The following courses will be offered: “ISCD Bone Densi- tometry Course” (Tampa, FL, March 13 and 14; Destin, FL, April 1 and 2; Washington, DC, April 22 and 23; Hartford, CT, May 5 and 6; Atlanta, May 19 and 20; Vancouver, BC, Canada, May 19 and 20; Allentown, PA, June 9 and 10; Montreal, June 23 and 24; Madison, WI, June 23 and 24; Kiawah Island, SC, July 7 and 8; Lansdowne, VA, July 22 and 23); “International Society for Clinical Densitometry 13th Annual Meeting” (Tampa, FL, March 14–17); and “ISCD Vertebral Fracture Assessment Course” (Tampa, FL, March 17; Washington, DC, April 23; Hartford, CT, May 4; Allentown, PA, June 8; Madison, WI, June 24; Lansdowne, VA, July 22).

Contact Anabela Gomes, International Society for Clinical Densitometry, 342 N. Main St., West Hartford, CT 06117-2507; or call (860) 586-7563, extension 583; or fax (860) 586-7550; or e-mail agomes@iscd.org; or see http://www.iscd.org.

NOTICES

Corrections

Prevention of Meningococcal Disease (October 5, 2006;355: 1466-73). In the third paragraph under the heading “Epidemiology of Neisseria meningitidis” (page 1467), the fourth sentence should have read “The highest rate of disease occurs among the young (Fig. 1), but 62% of cases occur in individuals 11 years of age or older,” rather than “The highest rates of disease occur among the very young (Fig. 1), and 62% of cases occur in children younger than 11 years of age.” The text has been corrected on the Journal’s Web site at www.nejm.org.

Release from Prison — A High Risk of Death for Former Inmates (January 11, 2007;356:157-65). The second sentence of the first paragraph should have read “At the end of 2001, there were approximately 5.6 million adults who had ever been incarcer- ated in a state or federal prison,” not including stays in local jails,” rather than “At the end of 2001, approximately 5.6 million adults were incarcerated in a state or federal prison” as printed. Also, in the second paragraph in the Results section, the fifth sentence should have read “In contrast, the calculated mortality rate for Washington State residents of the same age, sex, and race as the former inmates was 223 deaths per 100,000 person-years,” rather than “266 deaths per 100,000 person-years.” Similarly, in the legend to Figure 1, the first sentence should have read “The dashed line represents the adjusted mortality rate for residents of the State of Washington (223 deaths per 100,000 person-years), rather than “266 deaths per 100,000 person-years.” The text has been corrected on the Journal’s Web site at www.nejm.org.

The Journal’s Web and E-mail Addresses:

For letters to the Editor: authors.nejm.org
For information about the status of a submitted manuscript: authors.nejm.org
To submit a meeting notice: meetingnotices@nejm.org
The Journal’s Web pages: www.nejm.org

Book Reviews Copyright © 2007 Massachusetts Medical Society.