Article Summaries

PERTSPECTIVE

Taking TRIPS to India — Novartis, Patent Law, and Access to Medicines
J. M. Mueller

Thailand and the Compulsory Licensing of Efavirenz
R. Steinbrook

Focus on Research: Pulmonary Alveolar Proteinosis — Is Host Defense Awry?
C. M. Doerschuk

ORIGINAL ARTICLES

Clomiphene, Metformin, or Both for Infertility in the Polycystic Ovary Syndrome
R. S. Legro and Others

GM-CSF Autoantibodies and Neutrophil Dysfunction in Pulmonary Alveolar Proteinosis
K. Uchida and Others

A Human Interleukin-12/23 Monoclonal Antibody for the Treatment of Psoriasis
G. G. Krueger and Others

SPECIAL ARTICLES

Religion, Conscience, and Controversial Clinical Practices
F. A. Curlin, R. E. Lawrence, M. H. Chin, and J. D. Lantos

CLINICAL PRACTICE

The Incidentally Discovered Adrenal Mass
W. F. Young Jr.
IMAGES IN CLINICAL MEDICINE

Urine Fluorescence in Ethylene Glycol Poisoning
C. M. McStay and P. E. Gordon

Hydrofluoric Acid Burn
M. W. Dünser and J. Rieder

CASE RECORDS OF THE MASSACHUSETTS GENERAL HOSPITAL

Case 4-2007 — A 56-Year-Old Woman with Rapidly Progressive Vertigo and Ataxia
J. Dalmau, R. G. Gonzalez, and M. F. Lerwill

EDITORIALS

Treating the Polycystic Ovary Syndrome the Old-Fashioned Way
D. S. Guzick

HEALTH POLICY REPORTS

Regulatory and Judicial Oversight of Nonprofit Hospitals
D. M. Studdert, M. M. Mello, C. M. Jedrey, and T. A. Brennan

CORRESPONDENCE

Colonoscopy Screening for Detection of Advanced Neoplasia

Antithymocyte Globulin versus Basiliximab in Renal Transplantation

DHEA and Testosterone in the Elderly

Spinal Epidural Abscess

Medical Education — Professionalism

Thoracentesis

Elderly Survivors with Homozygous Sickle Cell Disease

BOOK REVIEWS

Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong

Surgically Shaping Children: Technology, Ethics, and the Pursuit of Normality

Melancholia: The Diagnosis, Pathophysiology, and Treatment of Depressive Illness
Taking TRIPS to India — Novartis, Patent Law, and Access to Medicines

Janice M. Mueller, J.D.

In August and September 2006, patients with cancer, lawyers for patient advocacy groups, and representatives of nongovernmental organizations (NGOs) converged on the offices of Novartis in Mumbai, India, to protest the company’s efforts to obtain an Indian patent on Gleevec, the company’s brand-name version of imatinib mesylate. Gleevec (spelled Glivec outside the United States) is used to treat chronic myeloid leukemia, and Novartis has patented the drug in 35 countries. The protesters also decried the drug’s high price: Novartis sells it in India (where only 5% of people have private health insurance) for $26,000 per year; generic-drug manufacturers offer the drug at less than one tenth that price.¹

Citing its right to recoup enormous research-and-development expenditures, Novartis refuses to drop the legal petitions it filed in the Chennai High Court in May 2006, challenging the Indian Patent Office’s denial of a patent. According to Novartis, there is “no faster way to kill access to the latest life-saving drugs for people in India than to avoid offering patent protection.”² The company also emphasizes that 99% of Indian patients now receiving the drug get it free through the company’s patient-assistance program.

The Gleevec challenge is the latest controversy facing India since its January 1, 2005, implementation of substantially enhanced patent protection for pharmaceuticals. India’s membership in the World Trade Organization (WTO) means that for the first time in 35 years, drug products (the pharmaceutical compositions themselves, rather than merely the processes for making them) must be considered potentially patentable in India. The Indian government supports the expanded availability of patent protection as a catalyst that may enable India’s enormous drug-manufacturing sector to evolve from reverse engineering to innovation.

It will take years, of course, to evaluate the effects of enhanced patent-based incentives on India’s pharmaceutical industry. The immediate concern is patients’ access to essential medicines that are manufactured in India and exported around the world. In the absence of notable patent-law restraints before 2005, India devel-
oped a world-class generic-drug-manufacturing sector, spawning major generics firms such as Ranbaxy, Cipla, and Dr. Reddy’s, in addition to hundreds of smaller firms. India boasts more drug-manufacturing facilities that have been approved by the U.S. Food and Drug Administration than any country other than the United States. Indian generics companies, for instance, supply 84% of the AIDS drugs that Doctors without Borders uses to treat 60,000 patients in more than 30 countries.3

Will India’s patenting of medicines put patients around the world at risk of losing a critical source of lifesaving generic drugs? The risk is currently minimal, thanks to public health safeguards developed by the Indian government. For example, the government has imposed price controls on essential medicines since 1970, and recent reports suggest that it may be expanding the list of drugs that are subject to such controls.4

More to the point, a number of safeguards have been built into the new patent law itself. These provisions resulted from years of intense public debate, government study, and legislative compromise.

First, patent coverage for pharmaceutical products will apply only prospectively to applications filed with the Indian Patent Office on or after January 1, 1995. Second, the law imposes powerful limitations on patents applied for between that date and December 31, 2004. Any Indian generics firm that began before 2005 to manufacture a drug that was subsequently covered by an Indian patent can continue to make and sell that drug, though it might have to pay royalties established by the government to the patent holder.

The law also includes the world’s most extensive provisions on “compulsory licensing.” Generics firms can legally copy patented drugs for export to the least-developed countries, which lack domestic manufacturing capability. Furthermore, generics firms and patient-advocacy groups are already making active use of robust “opposition” provisions in the law; indeed, it was opposition by a group of patients with cancer that led to the patent office’s rejection of the Gleevec application. And clearly, the culture engendered by 35 years of prohibition of the patenting of pharmaceuticals will not be changed overnight. Two years into the new patents regime, the government has granted only one patent on a pharmaceutical product — to Hoffmann–La Roche, for its hepatitis C therapy, peginterferon alfa-2a (Pegasys).

Still other protections included in the law ensure that only truly innovative advances will be patented. The Novartis lawsuit is the first legal challenge to the most controversial safeguard, a provision against “evergreening” that targets attempts to patent minor improvements to old drugs. Section 3(d) of India’s Patents Act forbids the patenting of derivative forms of known substances (e.g., salts, polymorphs, metabolites, and isomers) unless they are substantially more effective than the known substance. Neither the Indian patent statute nor its implementing rules define “efficacy.” They give the patent office no guidelines for applying the new test. Novartis has asked the Chennai High Court to strike down this section as inconsistent with the WTO’s Agreement on Trade-Related Aspects of Intellectual Property (TRIPS). TRIPS requires that patentable inventions be new and involve an “inventive step.” Novartis contends that TRIPS
gives WTO members the option of providing patent rights more generous than these basic criteria would mandate but does not allow members to go in the opposite direction by implementing stricter requirements for obtaining a patent.

The counterargument is that TRIPS does not define "inventive step." It permits (but does not require) WTO members to equate this criterion with the "nonobviousness" requirement of U.S. patent law — and thus gives member countries the flexibility to fine-tune their inventive-step criteria to reflect national socioeconomic conditions.

Moreover, Section 3(d) of India’s patent law does not necessarily impose stricter requirements than are used elsewhere; it may be seen as simply creating a general presumption of nonpatentability for modifications of known chemical compositions — and shifting to patent applicants the burden of rebutting this presumption in each particular case. For example, the U.S. Patent and Trademark Office may reject a claimed drug as "prima facie obvious" on the basis of its structural similarity to existing chemical compositions. A classic way to overcome the rejection is to demonstrate the drug’s unexpectedly good results. India’s new efficacy test might well operate in a similar fashion.

The Chennai High Court considered these issues of sufficient importance to merit referral to a two-judge panel. By late January 2007, the panel had not issued a decision. NGOs were disappointed by the court’s refusal to dismiss Novartis’s challenge outright. But the Indian judiciary must analyze and rule on the viability and uncertain contours of the new patentability test. Until it does so, the patent office retains virtually complete discretion in its application of Section 3(d). The court must also determine whether the patent office followed correct administrative procedures in rejecting Novartis’s application. The company contends that among other errors, patent examiners ignored data demonstrating that Gleevec has greater manufacturing stability than does the imatinib free-base form, as well as 30% greater bioavailability.5

India has an independent judiciary and an established rule-of-law tradition. Novartis’s litigation needs to run its course, and the system will benefit from the judicial analysis. And even if Novartis ultimately obtains an Indian patent on Gleevec, the current safeguards give the government multiple options for ensuring public access to this and other lifesaving drugs.

An interview with Ms. Mueller is available at www.nejm.org.

Ms. Mueller is a professor of law at the University of Pittsburgh School of Law, Pittsburgh.


Copyright © 2007 Massachusetts Medical Society.
Thailand and the Compulsory Licensing of Efavirenz

Robert Steinbrook, M.D.

Of the many medicines for human immunodeficiency virus (HIV) infection, efavirenz, a nonnucleoside reverse-transcriptase inhibitor that became available in the late 1990s, is one of the most important. For the initial treatment of adults, the combination of efavirenz and two nucleoside reverse-transcriptase inhibitors “has become a standard-of-care comparator in clinical trials,” according to Hammer et al. Moreover, efavirenz is available in a fixed-dose combination tablet with the nucleoside analogues emtricitabine and tenofovir; this tablet is taken only once a day. Efavirenz can cause birth defects when taken during the first trimester of pregnancy, so its use is restricted in women of childbearing potential, but its other side effects, such as central nervous system symptoms and rash, usually resolve, allowing treatment to continue.

Efavirenz, however, is relatively expensive, particularly for low- and middle-income countries, in part because it is more complicated to make than some other antiretroviral drugs. In 2006, first-line treatment regimens containing 600 mg of efavirenz, the usual daily dose for adults, had a median cost of nearly $500 per patient per year in middle-income countries; such regimens, including those containing generic efavirenz, were more than three times as costly as some alternatives (see graph). The pattern in low-income countries is similar.

The importance of efavirenz was highlighted again in November 2006, when Thailand’s Ministry of Public Health issued a compulsory license for the drug. The license, which took effect immediately and lasts until December 31, 2011, permits the Thai Government Pharmaceutical Organization to import generic efavirenz from India, where the drug is not patented, and to produce it—-at a time when Merck still has a patent on it in Thailand. The implications of this move go beyond efavirenz and beyond HIV. Thailand or other countries could use a similar rationale to license additional HIV drugs; drugs for other infectious diseases with public health consequences such as malaria, tuberculosis, or influenza; or drugs for cancer or other chronic diseases.

In the United States, Canada, and some European countries, Bristol-Myers Squibb markets efavirenz as Sustiva; in the rest of the world, Merck markets it as Stocrin. In the United States, a 1-year supply costs about $6,000. In 2006, Merck’s official price for the 600-mg formulation in the least developed countries was $277 per patient per year. In middle-income countries with an HIV prevalence rate among adults of 1% or greater, such as Thailand, it was also $277.

Since October 2003, Thailand has had a policy of universal coverage for antiretroviral treatment; patients receive their drugs free of charge. Thailand’s goal is to reduce the price of efavirenz, allowing the government to provide it to additional patients while remaining within its budget for the treatment of HIV and AIDS. An estimated 580,000 Thais (range, 330,000 to 920,000) were living with HIV infection at the end of 2005, according to the Joint United Nations Program on HIV/AIDS. As of June 2006, an estimated 89,000 Thais, or about four fifths of those who needed it, were receiving antiretroviral therapy, according to the World Health Organization. Less than one quarter received regimens that included efavirenz. As a result of the compulsory license, fewer Thai patients would be expected to receive regimens that include nevirapine, a less costly nonnucleoside reverse-transcriptase inhibitor that can cause a severe rash and hepatotoxic effects.

Thailand is not the first country to issue a compulsory license for an HIV drug—Indonesia, Malaysia, Mozambique, and Zambia have also done so. Objections of the pharmaceutical industry and international pressure, however, have ensured that such licenses remain rare. In addition, administering policies regarding intellectual property rights can be challenging. Thailand’s action has received considerable attention because the country has a leadership role in fighting AIDS, it has a domestic pharmaceutical industry, and it has licensed a high-profile medication. The government simply announced the “public use” of the patent without discussing the matter with Merck first.

Brazil has taken a different approach. Although it has threatened to issue compulsory licenses for antiretroviral drugs, so far it
has not actually done so. Instead, it has negotiated price discounts and voluntary licenses covering the local manufacture of patented drugs. Like Thailand, Brazil has a domestic pharmaceutical industry and provides free treatment to people with AIDS. As of June 2006, an estimated 186,000 people in Brazil — nearly 100% of those who needed it — were receiving antiretroviral therapy for HIV. But the generic AIDS drugs that Brazil manufactures predate its signing of the international Trade-Related Aspects of Intellectual Property Rights (TRIPS) agreement. Now, the cost of antiretroviral therapy is increasing substantially in Brazil — a reflection of the effects of drug resistance and the expanding requirements for newer, and more expensive, second-line agents.\(^5\)
Despite objections from Merck, Thailand has continued to issue compulsory licenses for antiretroviral drugs, including efavirenz. This has sparked a debate about the role of compulsory licensing in promoting access to essential medications. Thailand has made use of the TRIPS agreement to protect public health, but this has also led to tensions with Merck, who is the patent holder for efavirenz.

In issuing the compulsory license, Thailand cited its own laws and the declaration signed in Doha, Qatar, at a 2001 meeting of the World Trade Organization (WTO) regarding the relation between the TRIPS agreement and public health. The doha declaration clarified that the agreement "contains flexibilities that allow countries to enable both the import and production of generic versions of antiretroviral drugs under patent to protect public health." Nonetheless, Merck has objected to Thailand's unilateral action and wants the Thais to consider other options. For example, Merck might sell efavirenz at a lower price or it might provide a voluntary license for the production of a generic version, as it has done in South Africa.

Thailand has placed an order for efavirenz with Ranbaxy, a large Indian pharmaceutical company, as an interim measure. Meanwhile, the Thai Government Pharmaceutical Organization is testing the bioequivalence of the generic efavirenz that it has already formulated. Production is expected to start later this year. Merck is to receive a royalty fee of 0.5% of the total sale value of the medication that Thailand imports or produces, according to the Thai Ministry of Public Health. Nonetheless, Thailand could reach an agreement with Merck. Or the United States, which is negotiating a bilateral trade agreement with Thailand, could ask the government to agree to greater restrictions on compulsory licensing. Of course, Thailand may also issue additional compulsory licenses. In late January 2007, the Ministry of Public Health issued compulsory licenses for lopinavir–ritonavir (Kaletra, Abbott Laboratories), an antiretroviral that is a co-formulation of two protease inhibitors, and for clopidogrel (Plavix, Bristol-Myers Squibb and Sanofi-Aventis), an antiplatelet agent used for prevention of cardiovascular events. According to a spokesperson, the Office of the U.S. Trade Representative does "not question that WTO rules permit countries to issue compulsory licenses provided that they follow certain steps." Although the office has "not provided specific advice to Thai authorities on this matter," it has encouraged Thailand "to engage in dialogue...with all concerned parties."

Millions of people require antiretroviral medications, including second-line regimens that still cost $1,500 per year or more around the world. Production efficiencies and increased competition among manufacturers of generic drugs would be expected to decrease the price of medications such as efavirenz over time. Another way to encourage competition is for a pharmaceutical company to issue nonexclusive licenses to multiple manufacturers of generics. This is what Gilead has done in India, where it has granted 11 companies the rights to produce generic versions of tenofovir and distribute them to 95 low-income countries. There is no cost for these licenses, but if the company makes tenofovir and sells it, Gilead is to receive a 5% royalty on sales. Prices can also be reduced through negotiations with pharmaceutical companies — an approach taken by the William J. Clinton Foundation, most recently for pediatric formulations of antiretroviral drugs. Although the future of compulsory licensing for HIV drugs is uncertain, the need for greater availability of less expensive medications is not.

Dr. Steinbrook (rsteinbrook@attglobal.net) is a national correspondent for the Journal.


Copyright © 2007 Massachusetts Medical Society.
Pulmonary alveolar proteinosis is a rare disorder caused by abundant accumulation of surfactant-derived components in the lungs. The incidence is estimated to be 0.36 case per million population, and the prevalence, 3.70 cases per million. About 500 cases have been recorded in the literature.

The condition usually presents as progressive dyspnea and a minimally productive cough. There are elevated serum levels of products derived from pulmonary epithelial cells, including cytokinin 19, the mucin KL-6, and surfactant proteins A, B, and D. Pulmonary-function testing reveals a restrictive pattern and a disproportionate reduction in diffusing capacity. Patients often have hypoxemia, and the partial pressure of oxygen in the arteries is elevated as compared with that in the alveoli. The appearance of the lungs on high-resolution computed tomography is often a homogeneous ground-glass haze overlaid by thickened interlobular septa forming geometric shapes — the “crazy paving” pattern.

Diagnosis usually requires lung biopsy. The characteristic pathological feature is the filling of alveoli and distal bronchioles with surfactant-derived material that is granular, acidophilic, acellular, and amorphous; periodic acid–Schiff staining is positive and diastase-resistant. The material consists of approximately 90% lipid (primarily phospholipids), 10% protein, and 1% carbohydrate and is usually sterile. Large foamy alveolar macrophages are present, and the alveolar structure is preserved until late in the course of disease. Mild interstitial lymphocytic infiltrates may be present.

Pulmonary alveolar proteinosis results either from decreased degradation of surfactant or from surfactant dysfunction and is classified as primary, secondary, or congenital. The congenital type has been attributed to one of two rare mutations: one causes a deficiency of surfactant protein B, and the other causes abnormalities in the β chain of the receptor for granulocyte–macrophage colony-stimulating factor (GM-CSF). Secondary pulmonary alveolar proteinosis occurs in patients with systemic inflammatory diseases or cancers (most commonly hematologic) or in association with other causes.

Primary pulmonary alveolar proteinosis accounts for nearly 90% of cases and is virtually always attributable to neutralizing GM-CSF autoantibodies, which prevent the binding of GM-CSF to GM-CSF receptors on alveolar macrophages (see diagram). GM-CSF–initiated signaling plays a unique, nonredundant role in alveolar macrophage function and pulmonary homeostasis.

The hypothesis that pulmonary alveolar proteinosis is due to ineffective signaling by GM-CSF receptors was first put forward in 1994, when Stanley et al. and Dranoff et al. simultaneously reported on mice in which both alleles of the gene for GM-CSF are disabled. These GM-CSF−/− mice have striking pulmonary pathologic characteristics that closely resemble those of patients with pulmonary alveolar proteinosis, suggesting that the intracellular signaling initiated by the binding of GM-CSF to its receptor is critical to pulmonary surfactant homeostasis. Signaling initiated by GM-CSF receptors is mediated through PU.1, a transcription factor modulating the expression of many genes that are important in the terminal differentiation of alveolar macrophages. Functions regulated by PU.1 include surfactant degradation, expression of pathogen pattern-recognition receptors, toll-like–receptor signaling, phagocytosis, and bacterial killing. Reconstitution of PU.1 in GM-CSF–deficient alveolar macrophages restores most of these functions. GM-CSF signaling also enhances the function of peroxisome-proliferator–activated receptor γ (PPARγ), another transcription factor that regulates many cellular functions, including intracellular lipid metabolism. These findings explain how inhibiting the binding of GM-CSF to its receptor causes decreased clearance of surfactant from the alveolar spaces. GM-CSF is required for the terminal differentiation and function of alveolar macrophages.
but not for those of other tissue macrophages — which may explain why pulmonary alveolar proteinosis is primarily a lung disease.⁴

The discovery that nearly all patients with primary pulmonary alveolar proteinosis have high titers of GM-CSF antibodies has led to the development of new therapies based on granulocyte colony-stimulating factor (G-CSF) supplementation. Previously, therapy consisted of whole-lung lavage, which results in numerous complications and does not address the pathogenetic mechanisms. Administration of exogenous GM-CSF appears to help many patients, and its potential as a subcutaneous or aerosolized therapy is being evaluated.

Whether patients with pulmonary alveolar proteinosis have defects in host defense and innate immunity is not clear, nor is the effect of such defects. GM-CSF−/− mice clearly have a defect in pulmonary defense against pathogens, as well as extrapulmonary abnormalities and a decreased susceptibility to experimentally induced autoimmune disorders. Patients with pulmonary alveolar proteinosis do not have deficient expression of GM-CSF, but they have neutralizing GM-CSF autoantibodies, which may account for differences between host defense defects in such patients and defects in mice. Although secondary infections have been described in these patients, our understanding of host defense in this disorder is limited by its rarity, the variability of its outcomes, and reporting biases. Patients may be at risk for secondary infections, but bacteria that commonly cause respiratory infections are not often the inciting

---

**Pathophysiology of Pulmonary Alveolar Proteinosis.**

GM-CSF–initiated signaling (Panel A) plays unique, nonredundant roles in alveolar macrophage function and pulmonary homeostasis, including terminal differentiation and survival of macrophages, intracellular lipid metabolism, surfactant catabolism and recycling, expression of pathogen receptors, and phagocytosis and killing. These functions are inhibited when neutralizing GM-CSF autoantibodies bind to GM-CSF (Panel B), block its binding to the α chain of the GM-CSF receptor, and prevent assembly of the GM-CSF–receptor complex. Thus, inhibition of GM-CSF binding to its receptor by autoantibodies results in decreased clearance of surfactant from the alveolar spaces, the hallmark of pulmonary alveolar proteinosis.
agents.1 Opportunistic pathogens—most commonly nocardia, but occasionally cryptococcus, histoplasma, aspergillus, or Mycobacterium tuberculosis—have been reported in about 13% of patients,4 with systemic infection documented in some cases. In a review of 65 cases in which death was attributable to pulmonary alveolar proteinosis, respiratory failure was the apparent cause of death in 47 cases (72%), whereas uncontrolled infection was the cause in 12 (18%).1 These findings suggest that clinically important lung infections occur in some patients and that systemic infections are less common.

Few researchers have investigated whether GM-CSF autoantibodies cause defects in cells, other than alveolar macrophages, that express GM-CSF receptors. In this issue of the Journal, Uchida et al. report on the effects of GM-CSF autoantibodies on neutrophil functions (pages 567–579). Their observations show that neutrophils from patients with pulmonary alveolar proteinosis have defects in both basal and GM-CSF–primed antimicrobial functions. In particular, the phagocytic index and phagocytic capacity of neutrophils isolated from these patients were approximately 90% and 30% lower, respectively, than those of neutrophils from healthy control subjects. The basal capacities for adhesion to plastic, the oxidative burst in whole blood, and the killing of Staphylococcus aureus were also reduced. Furthermore, GM-CSF priming in vitro was impaired. Similar defects in basal neutrophil functions were observed in GM-CSF−/− mice, although no defect in GM-CSF priming was observed. The defect in affected human neutrophils was mimicked by treating blood from healthy control subjects with GM-CSF autoantibodies. Thus, neutrophils from patients with pulmonary alveolar proteinosis have functional defects when tested in vitro.

How the signaling of GM-CSF through its receptor on neutrophils augments neutrophil function remains an important question.5 PU.1 is apparently critical in GM-CSF–initiated signaling in alveolar macrophages but not in neutrophils, since Uchida et al. show that neutrophils from patients with pulmonary alveolar proteinosis do not have lower PU.1 expression than those from healthy control subjects. Furthermore, defects in neutrophil function are apparent in vitro, but it is not clear how they contribute to a defect in host defense in vivo. Humans and mutant mice with defects in innate immunity, including abnormal neutrophil function, often have increased blood levels of neutrophils, G-CSF, and cytokines; the failure to destroy pathogens leads to the persistence of pathogen-induced stimuli. In contrast, in the study by Uchida et al., patients with pulmonary alveolar proteinosis who had defective neutrophil function in vitro did not have increased neutrophil counts and serum G-CSF levels, suggesting that their host defense and innate immunity were sufficient for managing daily exposures to common pathogens and commensal organisms. Clinically apparent infections, especially with opportunistic microbes, do occur in some patients; microbes were identified at presentation in more than half the patients studied by Uchida et al.

Whether host defense mechanisms are able to compensate for defects in neutrophil and macrophage function until the environment delivers a particular pathogen or a large load of pathogens, or until a combined assault is made on the immune system, remains to be elucidated. The absence of repeated infections, particularly in the lungs, is surprising, given the magnitude of the neutrophil defects seen in vitro; perhaps products of surfactant degradation have important antimicrobial effects. Clearly, the study by Uchida et al. raises important, provocative questions that beg to be pursued.
Clomiphene, Metformin, or Both for Infertility in the Polycystic Ovary Syndrome

This multicenter, randomized trial compared the effects of clomiphene citrate plus placebo, metformin plus placebo, and combination therapy in infertile women with the polycystic ovary syndrome. The rate of live birth was significantly higher with clomiphene than with metformin; there was no significant difference between the rates with combination therapy and with clomiphene alone. Multiple birth was a complication associated with clomiphene but was infrequent. These data support the use of clomiphene over metformin for the treatment of infertility in women with the polycystic ovary syndrome.

GM-CSF Autoantibodies and Neutrophil Dysfunction in Pulmonary Alveolar Proteinosis

Infection, especially with opportunistic microbes, is a prominent feature of pulmonary alveolar proteinosis; extrapulmonary infection suggests a systemic susceptibility. The authors show that neutrophil functions (phagocytosis, adhesion, oxidative burst, and bactericidal activity) are depressed in patients with pulmonary alveolar proteinosis and that the cause is autoantibodies against granulocyte–macrophage colony-stimulating factor (GM-CSF). These findings clearly demonstrate the essential role of GM-CSF in the antimicrobial activities of neutrophils.

Interleukin-12/23 Monoclonal Antibody for the Treatment of Psoriasis

Type 1 cytokines are overexpressed in psoriatic plaques. This trial evaluated a monoclonal antibody against interleukin-12 and interleukin-23 in patients with psoriasis. Response rates at 12 weeks were significantly higher in patients treated with interleukin-12/23 monoclonal antibody than in those treated with placebo. Four percent of patients who received interleukin-12/23 monoclonal antibody and 1% of those who received placebo had serious adverse events. Larger studies of longer duration are needed to assess the effectiveness and safety of interleukin-12/23 monoclonal antibody for psoriasis.

The Incidentally Discovered Adrenal Mass

A 68-year-old woman is incidentally found to have a left adrenal mass, 2.8 cm in diameter, on abdominal computed tomography that was ordered to evaluate right lower abdominal discomfort (which has since resolved). Her medical history is notable only for hypertension that has been well controlled with hydrochlorothiazide, at a dose of 25 mg daily. She reports no sweating, palpitations, headache, weight gain, or proximal muscle weakness. Her physical examination is unremarkable. How should she be evaluated?

Health Policy Report

Legal Oversight of Nonprofit Hospitals

Nonprofit hospitals agree to operate for charitable purposes in exchange for exemption from taxes. As financial pressures facing hospitals have intensified, the business decisions of nonprofit hospitals are being challenged by the Internal Revenue Service, state attorneys general, and patients. Uninsured patients have joined class-action lawsuits arguing that nonprofit hospitals are abandoning their charitable missions and accusing them of overcharging uninsured patients and of using aggressive debt-collection measures when uninsured patients cannot pay their bills.
Clomiphene, Metformin, or Both for Infertility in the Polycystic Ovary Syndrome

Richard S. Legro, M.D., Huiman X. Barnhart, Ph.D., William D. Schlaff, M.D., Bruce R. Carr, M.D., Michael P. Diamond, M.D., Sandra A. Carson, M.D., Michael P. Steinkampf, M.D., Christos Coutifaris, M.D., Ph.D., Peter G. McGovern, M.D., Nicholas A. Cataldo, M.D., Gabriella G. Gosman, M.D., John E. Nestler, M.D., Linda C. Giudice, M.D., Ph.D., Phyllis C. Leppert, M.D., Ph.D., and Evan R. Myers, M.D., M.P.H., for the Cooperative Multicenter Reproductive Medicine Network*

ABSTRACT

BACKGROUND
The polycystic ovary syndrome is a common cause of infertility. Clomiphene and insulin sensitizers are used alone and in combination to induce ovulation, but it is unknown whether one approach is superior.

METHODS
We randomly assigned 626 infertile women with the polycystic ovary syndrome to receive clomiphene citrate plus placebo, extended-release metformin plus placebo, or a combination of metformin and clomiphene for up to 6 months. Medication was discontinued when pregnancy was confirmed, and subjects were followed until delivery.

RESULTS
The live-birth rate was 22.5% (47 of 209 subjects) in the clomiphene group, 7.2% (15 of 208) in the metformin group, and 26.8% (56 of 209) in the combination-therapy group (P<0.001 for metformin vs. both clomiphene and combination therapy; P=0.31 for clomiphene vs. combination therapy). Among pregnancies, the rate of multiple pregnancy was 6.0% in the clomiphene group, 0% in the metformin group, and 3.1% in the combination-therapy group. The rates of first-trimester pregnancy loss did not differ significantly among the groups. However, the conception rate among subjects who ovulated was significantly lower in the metformin group (21.7%) than in either the clomiphene group (39.5%, P=0.002) or the combination-therapy group (46.0%, P<0.001). With the exception of pregnancy complications, adverse-event rates were similar in all groups, though gastrointestinal side effects were more frequent, and vasomotor and ovulatory symptoms less frequent, in the metformin group than in the clomiphene group.

CONCLUSIONS
Clomiphene is superior to metformin in achieving live birth in infertile women with the polycystic ovary syndrome, although multiple birth is a complication. (ClinicalTrials.gov number, NCT00068861.)

From Pennsylvania State University College of Medicine, Hershey (R.S.L.); Duke University Medical Center, Durham, NC (H.X.B., E.R.M.); University of Colorado, Denver (W.D.S.); University of Texas Southwestern Medical Center, Dallas (B.R.C.); Wayne State University, Detroit (M.P.D.); Baylor College of Medicine, Houston (S.A.C.); University of Alabama, Birmingham (M.P.S.); University of Pennsylvania School of Medicine, Philadelphia (C.C.); University of Medicine and Dentistry of New Jersey, Newark (P.G.M.); Stanford University, Stanford, CA (N.A.C.); University of Pittsburgh, Pittsburgh (G.G.G.); Virginia Commonwealth University School of Medicine, Richmond (J.E.N.); University of California at San Francisco, San Francisco (L.C.G.); and the National Institute of Child Health and Human Development, Bethesda, MD (P.C.L.). Address reprint requests to Dr. Legro at the Department of Obstetrics and Gynecology, Pennsylvania State University College of Medicine, M.S. Hershey Medical Center, 500 University Dr., H103, Hershey, PA 17033, or at rsl1@psu.edu.

*Other members of the Cooperative Multicenter Reproductive Medicine Network are listed in the Appendix.

The polycystic ovary syndrome affects 7 to 8% of women\(^1\) and may be the most common cause of female infertility.\(^2\) Anovulation,\(^2\) early pregnancy loss,\(^3\) and later pregnancy complications\(^4\) have all been implicated in the low fecundity of women with this disorder. Obesity is also common in such women,\(^5\) and this condition alone appears to have an adverse effect on reproduction.\(^6,7\) The cause of the polycystic ovary syndrome is poorly understood, and both the diagnosis and treatment of the disorder are controversial.\(^5,8,9\)

Women with this syndrome have hyperandrogenism,\(^10\) morphologic changes in the ovary (polycystic),\(^10\) inappropriate gonadotropin secretion (elevated levels of circulating luteinizing hormone),\(^11\) and insulin resistance with accompanying compensatory hyperinsulinemia.\(^12\) Targeting these metabolic abnormalities has been noted to improve ovulation and fertility in women with this syndrome.\(^13-17\) Results from small head-to-head trials have suggested that the efficacy of treatment with insulin sensitizers such as metformin (alone or in combination with clomiphene citrate) is equal or superior to that of clomiphene alone for infertility.\(^13,16,17\)

We designed a trial to test the hypothesis that treatment of women with the polycystic ovary syndrome with extended-release metformin is more likely to result in a live birth than is treatment with clomiphene citrate and that the combination of the two therapies will result in the highest live-birth rate.

**STUDY DESIGN**

We have previously described the rationale for choosing live birth as the primary outcome,\(^18\) the power analysis and main statistical methods,\(^19\) the use of infertility screening in the study,\(^20\) and the study design and baseline characteristics of the subjects.\(^21\)

The institutional review board at each center approved the protocol, and all subjects gave written informed consent. All subjects had received the diagnosis of the polycystic ovary syndrome, which was defined as oligomenorrhea (with a history of no more than eight spontaneous menses per year) and hyperandrogenemia (with an elevated testosterone level documented within the previous year in an outpatient setting on the basis of local laboratory results, with a predetermined cutoff level set by the principal investigator at each study site). Subjects were excluded if they had hyperprolactinemia, congenital adrenal hyperplasia, thyroid disease, or other causes of amenorrhea, including premature ovarian failure. Clinically suspected Cushing’s syndrome and androgen-secreting neoplasm were additional exclusion criteria.\(^21\)

We randomly assigned 626 infertile women with the polycystic ovary syndrome to one of three study groups by means of an interactive voice system. The assignments were stratified according to the study site and the presence or absence of previous exposure to either of the study drugs.\(^21\) Subjects with other causes of infertility were excluded on the basis of documentation of a normal uterine cavity and at least one patent fallopian tube; analysis of the semen of each woman’s current partner was performed within 1 year before participation in the study, and a sperm concentration of at least 20 million per milliliter was required.\(^21\) All subjects were in good health with no major medical disorders.\(^21\)

**STUDY DRUGS**

We used extended-release metformin because of its increased tolerability and proven efficacy in the treatment of type 2 diabetes.\(^22,23\) Extended-release metformin (Glucophage XR) plus identical placebo were provided by Bristol-Myers Squibb. Overencapsulated clomiphene citrate tablets (purchased from Teva Pharmaceuticals) and matching placebo capsules were packaged and tested by a commercial pharmacy supply company (CTS) specifically for the study. Neither manufacturer had any other role in the study.

Baseline laboratory testing was performed after the subjects had fasted overnight. All specimens were analyzed in a core laboratory.\(^21\) In subjects without recent menses, withdrawal bleeding was induced with a course of oral medroxyprogesterone acetate before the initiation of study medication. Each subject received a monthly medication package consisting of bottle M (metformin in 500-mg tablets or matching placebo) and blister pack C (clomiphene in 50-mg tablets or matching placebo). The two drugs were begun concurrently. Subjects gradually increased the dose of the study drug in bottle M until reaching the maximum dose of four tablets (two tablets twice a day). Subjects took one tablet a day from
blister pack C for 5 days, beginning on day 3 of menses; this dose was maintained if adequate ovulation was documented. However, in subjects who had no response or a poor response, the dose was increased by one tablet a day on a treatment-cycle basis (either after 5 weeks of anovulation or after a menses until the maximum dose of three tablets per day was reached).

After the baseline visit, subjects returned each month for a visit with a limited physical examination, urine pregnancy test, and repeated fasting blood tests. Subjects were instructed to have regular intercourse every 2 to 3 days and to keep a diary recording intercourse, vaginal bleeding, and symptoms. The progesterone levels in all subjects were measured weekly or every other week in local laboratories in order to document ovulation. If two consecutive measurements showed elevated levels of progesterone (above 5 ng per milliliter [16 nmol per liter]), a weekly pregnancy test was administered until a positive result or menses occurred. Induction of withdrawal bleeding with progestin was scheduled at the discretion of the principal investigator at each site. Ultrasonography for follicular and endometrial response was not included in the protocol, and ovulation triggering with human chorionic gonadotropin and intrauterine insemination were not permitted.

Subjects were treated for up to six cycles, or 30 weeks. All study medication was discontinued if a pregnancy test was positive. Pregnant subjects were followed until ultrasonography documented fetal viability and were then referred for prenatal care. Investigators reviewed all obstetrical records to obtain data on birth outcomes. We did not collect data on the use of other medications during pregnancy.

OUTCOMES
The primary outcome of the trial was the rate of live births. Secondary outcomes included the rate of pregnancy loss, singleton birth, and ovulation (a serum progesterone level above 5 ng per milliliter during a cycle). A serious adverse event was defined as any event that was fatal, immediately life-threatening, or severely or permanently disabling; an event that required or prolonged hospitalization; an overdose (intentional or accidental); a congenital anomaly; pregnancy loss after 12 weeks of gestation; or an event that was deemed to be serious by the principal investigator at each site.

DATA MANAGEMENT
All data entry, data management, and analyses were performed at the Data Coordinating Center at the Duke Clinical Research Institute. Subjects were enrolled in the study from November 2002 to December 2004. The last conception was in June 2005, the last subject finished medication in August 2005, and the last birth was reported in February 2006. Data were analyzed according to the intention-to-treat principle.

STATISTICAL ANALYSIS
We assumed a dropout rate of 15% and the following rates of live birth: 45% in the combination-therapy group, 30% in the metformin group, and 25% in the clomiphene group. On the basis of these assumptions, we needed to enroll 678 subjects for the study to have a power of 80% with a type I error rate of 0.05 to detect a 15% absolute difference in live-birth rates for the following two primary comparisons: the combination-therapy group versus the next best group and the metformin group versus the clomiphene group.

Because of limitations in the supply of metformin and matching placebo, the number of subjects was reduced to 626 after the data safety and monitoring board reviewed blinded data in November 2004. Because the live-birth rate was lower than projected, the final number of subjects provided adequate power (≥80%) to detect the same 15% absolute difference in live-birth rates for the original two primary comparisons because of the increased power for detecting the same difference in proportions when the magnitude of the proportion was decreased.

Either a chi-square test or Fisher’s exact test was used for testing differences among the three study groups for categorical variables. A Wilcoxon rank-sum test was used for testing differences between two groups, and a Kruskal–Wallis test was used for testing differences among groups of three or more. Kaplan–Meier curves were used for time-to-event analyses. Generalized estimating equations were used for analysis of the ovulation rate to account for correlation of multiple ovulation cycles for each subject. Post hoc stratification of outcomes was performed on the basis of the body-mass index (BMI; the weight in kilograms divided by the square of the height in meters) and presence or absence of previous exposure to medication. All analyses were performed with SAS software, version 8.2 (SAS Institute).
RESULTS

STUDY POPULATION

There were no significant differences in baseline variables among the study groups (Table 1). The numbers of subjects who dropped out of the study were 55 of 209 (26.3%) in the clomiphene group, 72 of 208 (34.6%) in the metformin group, and 49 of 209 (23.4%) in the combination-therapy group (P = 0.07 for the metformin group vs. the clomiphene group, and P = 0.01 for the metformin group vs. the combination-therapy group) (Fig. 1). The reasons for dropout were similar among the three groups, except that the metformin group had a higher rate of loss to follow-up than did the other two groups (P = 0.03 for the comparison with the clomiphene group, and P = 0.01 for the comparison with the combination-therapy group).

PRIMARY OUTCOME

The rate of live birth was significantly lower in the metformin group than in the clomiphene group and the combination-therapy group (P < 0.001 for both comparisons), and there was no significant advantage of the combination therapy over clomiphene (Table 2 and Fig. 2A). However, independently of treatment, subjects with a BMI below 30 had a significantly higher rate of live births than did women whose BMI was 30 or more (P < 0.001 by univariate analysis) (Fig. 2B). The relationship between treatment and live birth was similar in post hoc analyses of subgroups stratified according to BMI (<30, 30 to 34, and ≥35) (see Table 1 of the Supplementary Appendix, available with the full text of this article at www.nejm.org) and according to previous treatment (Table 2 of the Supplementary Appendix).

Table 1. Baseline Characteristics of the Subjects.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Clomiphene Group (N = 209)</th>
<th>Metformin Group (N = 208)</th>
<th>Combination-Therapy Group (N = 209)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biometric features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age — yr</td>
<td>27.9±4.0</td>
<td>28.1±4.0</td>
<td>28.3±4.0</td>
</tr>
<tr>
<td>BMI</td>
<td>36.0±8.9</td>
<td>35.6±8.5</td>
<td>34.2±8.4</td>
</tr>
<tr>
<td>Clinically significant hirsutism (FG &gt;16) — no. (%)</td>
<td>81 (38.8)</td>
<td>76 (36.5)</td>
<td>86 (41.1)</td>
</tr>
<tr>
<td>Waist circumference — cm</td>
<td>105.0±22.3</td>
<td>102.4±17.6</td>
<td>100.2±18.2</td>
</tr>
<tr>
<td>Race or ethnic group — no. (%)†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>147/208 (70.7)</td>
<td>140/207 (67.6)</td>
<td>148/208 (71.2)</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>53/209 (25.4)</td>
<td>61/208 (29.3)</td>
<td>50/209 (23.9)</td>
</tr>
<tr>
<td>Black</td>
<td>37/208 (17.8)</td>
<td>40/207 (19.3)</td>
<td>32/208 (15.4)</td>
</tr>
<tr>
<td>Asian</td>
<td>5/208 (2.4)</td>
<td>5/207 (2.4)</td>
<td>7/208 (3.4)</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>21/208 (10.1)</td>
<td>27/207 (13.0)</td>
<td>24/208 (11.5)</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>1/208 (0.5)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fertility history</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of time subject had been attempting conception — mo</td>
<td>41.4±39.4</td>
<td>39.0±31.9</td>
<td>40.7±36.0</td>
</tr>
<tr>
<td>Previous therapy for infertility — no. (%)</td>
<td>116 (55.5)</td>
<td>111 (53.4)</td>
<td>116 (55.5)</td>
</tr>
<tr>
<td>Previous pregnancy — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conception</td>
<td>77 (36.8)</td>
<td>66 (31.7)</td>
<td>67 (32.1)</td>
</tr>
<tr>
<td>Live birth</td>
<td>33 (15.8)</td>
<td>33 (15.9)</td>
<td>28 (13.4)</td>
</tr>
<tr>
<td>Pregnancy loss</td>
<td>53 (25.4)</td>
<td>40 (19.2)</td>
<td>45 (21.5)</td>
</tr>
<tr>
<td>Previous exposure to study drug — no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>89 (42.6)</td>
<td>87 (41.8)</td>
<td>86 (41.1)</td>
</tr>
<tr>
<td>Metformin only</td>
<td>14 (6.7)</td>
<td>16 (7.7)</td>
<td>24 (11.5)</td>
</tr>
<tr>
<td>Clomiphene only</td>
<td>67 (32.1)</td>
<td>68 (32.7)</td>
<td>53 (25.4)</td>
</tr>
<tr>
<td>Combination therapy</td>
<td>39 (18.7)</td>
<td>37 (17.8)</td>
<td>46 (22.0)</td>
</tr>
</tbody>
</table>
SECONDARY OUTCOMES
The rate of ovulation was significantly higher in the combination group than in either of the single-agent groups (Table 2). Over the course of the study, the mean (±SD) number of ovulations per subject was 2.22±1.87 in the clomiphene group, 1.43±1.72 in the metformin group, and 2.80±2.04 in the combination-therapy group (P<0.001 for the comparisons of the clomiphene and combination-therapy groups with the metformin group). However, as previously noted, the differences in ovulation rates did not translate into an increase in the live-birth rate among subjects receiving combination therapy. Rates of conception and live birth per cycle in which ovulation occurred and per subject who ovulated were significantly higher in the clomiphene group and the combination-therapy group than in the metformin group. All multiple pregnancies occurred in either the clomiphene group or the combination-therapy group, although rates were low, and the differences among the three groups were not significant (P=0.56 for the metformin group vs. the clomiphene group, and P=1.0 for the metformin group vs. the combination-therapy group).

Over the course of the study, there were no documented ovulations in 52 of 209 women (24.9%) in the clomiphene group, 93 of 208 (44.7%) in the metformin group, and 35 of 209 (16.7%) in the combination-therapy group (P<0.001 for both comparisons with the metformin group). There was no significant linear effect of time on the rate of ovulation in the metformin group and the clomiphene group, but there was such an effect on ovulation and live birth in the combination-therapy group (P=0.002 and P=0.05, respectively).

OTHER TREATMENT EFFECTS
We analyzed metabolic and hormonal effects associated with the medications by comparing baseline data with data recorded at the last study visit.

Table 1. (Continued.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Clomiphene Group (N=209)</th>
<th>Metformin Group (N=208)</th>
<th>Combination-Therapy Group (N=209)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasonographic findings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphologic features of polycystic ovary — no. (%)‡</td>
<td>192 (91.9)</td>
<td>189 (90.9)</td>
<td>192 (91.9)</td>
</tr>
<tr>
<td>Ovarian volume — cm³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left ovary</td>
<td>10.9±6.9</td>
<td>11.2±6.1</td>
<td>11.2±6.2</td>
</tr>
<tr>
<td>Right ovary</td>
<td>11.8±6.9</td>
<td>11.8±5.9</td>
<td>12.5±8.1</td>
</tr>
<tr>
<td>Fasting serum levels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proinsulin — pmol/liter</td>
<td>25.4±24.3</td>
<td>25.5±27.5</td>
<td>23.9±25.7</td>
</tr>
<tr>
<td>Insulin — µU/ml</td>
<td>22.6±20.7</td>
<td>24.0±28.4</td>
<td>24.4±30.0</td>
</tr>
<tr>
<td>Glucose — mg/dl</td>
<td>89.2±16.5</td>
<td>88.8±17.1</td>
<td>88.9±18.6</td>
</tr>
<tr>
<td>SHBG — nmol/liter</td>
<td>29.8±18.7</td>
<td>27.5±14.4</td>
<td>31.8±20.3</td>
</tr>
<tr>
<td>Testosterone — ng/dl</td>
<td>61.3±32.0</td>
<td>61.6±25.0</td>
<td>63.1±28.4</td>
</tr>
<tr>
<td>Free androgen index§</td>
<td>9.4±7.1</td>
<td>9.9±6.2</td>
<td>9.4±6.8</td>
</tr>
<tr>
<td>HOMA-IR¶</td>
<td>5.2±5.3</td>
<td>5.6±8.9</td>
<td>5.6±10.2</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. Percentages may not total 100 because of rounding. To convert the values for glucose to millimoles per liter, multiply by 0.05551. To convert the values for insulin to picomoles per liter, multiply by 6. To convert the values for testosterone to nanomoles per liter, multiply by 0.03467. BMI denotes body-mass index (the weight in kilograms divided by the square of the height in meters), FG Ferriman–Gallwey score, and SHBG sex hormone–binding globulin.
† Race or ethnic group was designated by the subjects. Some subjects chose more than one category, including Hispanic or Latino.
‡ One or both ovaries were affected.
§ The free androgen index was calculated according to the following formula: (total testosterone [nanomoles per liter] ÷ SHBG [nanomoles per liter]) ×100.
¶ A homeostasis model assessment of insulin resistance (HOMA-IR) was calculated according to the following formula: (insulin × glucose) ÷ 405.
before pregnancy was documented or at the final study visit, whichever came first (Table 3). As compared with the baseline values, the clomiphene group had a significant increase in BMI ($P = 0.05$), levels of insulin ($P = 0.01$), insulin resistance as determined by homeostasis model assessment (HOMA) ($P = 0.01$), and levels of sex hormone–binding globulin ($P < 0.001$) and a corresponding decrease in the free androgen index ($P < 0.001$). Conversely, the metformin group had a significant decrease in BMI and total testosterone and a significant increase in sex hormone–binding globulin levels, with a corresponding decrease in the free androgen index ($P < 0.001$ for all comparisons).
The combination-therapy group had changes similar to those in the metformin group, including a significant decrease in BMI, levels of testosterone, and the free androgen index and a significant increase in sex hormone–binding globulin levels (P<0.001 for all comparisons), and a significant decrease in waist circumference (P=0.004).

Over the course of the study, as compared with the metformin group, both the clomiphene group and the combination-therapy group had significant increases in sex hormone–binding globulin levels (P<0.001 for both comparisons) and decreases in the free androgen index (P<0.001 for the combination-therapy group vs. the metformin group, and P=0.01 for the clomiphene group vs. the metformin group). As compared with the clomiphene group, the combination-therapy group had a significant decrease in BMI (P<0.001) and in levels of testosterone (P<0.001), proinsulin (P=0.04), insulin (P=0.01), and insulin resistance as determined by HOMA (P=0.006).

**ADVERSE EVENTS AND PREGNANCY COMPLICATIONS**

Serious adverse events, most of which were complications of pregnancy, were more common among subjects in the clomiphene group and the combination-therapy group than among those in the metformin group: 7 of 209 (3.3%), 11 of 209 (5.3%; 1 subject had two events), and 2 of 208 (1.0%), respectively (P=0.12 for metformin vs. clomiphene, and P=0.02 for metformin vs. combination therapy) (Table 4). Gastrointestinal symptoms were more frequent in the groups receiving metformin, whereas hot flashes and symptoms associated with ovarian enlargement and ovulation were more common in the groups receiving clomiphene.

The rates of compliance with the recommended frequency of sexual intercourse ranged from 74 to 76% at the first visit and declined to 54 to 56% at the visit at 6 months. The compliance rates were similar across groups at all cycles except cycle 4, during which the rates were 71% in the clomiphene group, 66% in the metformin group, and 76% in the combination-therapy group (P=0.04 for the combination-therapy group vs. the metformin group).

**DISCUSSION**

Our findings do not support the hypothesis that extended-release metformin, either alone or in combination with clomiphene citrate, improves the rate of live birth in women with the polycystic ovary syndrome. Conception, pregnancy, and live birth were significantly more likely to occur after treatment with clomiphene alone than after metformin alone. Adverse-event rates were similar among the study groups, although serious adverse events, primarily related to pregnancy, tended to occur in the groups receiving clomiphene (either alone or in combination therapy); these groups had correspondingly higher rates of pregnancy than did the metformin group.

These results are inconsistent with data from several other studies reporting benefits of metformin, especially in combination with clomiphene, in stimulating ovulation in women with the polycystic ovary syndrome. Previous data have generally come from small, primarily single-center trials that did not assess pregnancy rates but rather focused on metabolic and hormonal measures, rates of ovulation, or both. As in the earlier studies, we found that the groups receiving metformin (both the metformin group and the combination-therapy group) had improved insulin sensitivity (including effects on BMI, proinsulin and insulin levels, and insulin resistance as determined by HOMA), as compared with the clomiphene group. However, these effects did not translate into increased live-birth rates. Instead, increases in sex hormone–binding globulin levels were associated with improved live-birth rates.

Our results also differed from those of a previous randomized trial by Palomba et al., which included 100 subjects and used a design similar to ours. In that study, the live-birth rate was 52% after 6 months of metformin, as compared with 18% after clomiphene. In contrast to our results, levels of fecundity improved over time among subjects receiving metformin, as compared with clomiphene. Unlike our trial, the study by Palomba et al. excluded subjects whose BMI was greater than 30. However, our post hoc analysis of women with a BMI of less than 30 also showed an increased live-birth rate with clomiphene, as compared with metformin.

Our findings regarding the effects of the combination of metformin and clomiphene are consistent with those of another large, multicenter, randomized trial, reported by Moll et al., in which the rate of ovulation was the primary outcome. Among 228 subjects with the polycystic ovary syndrome who were randomly assigned to...
### Table 2. Rates of Ovulation, Pregnancy, and Pregnancy Loss.\(^a\)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Clomiphene Group (N = 209) no./total no. (%)</th>
<th>Metformin Group (N = 208) no./total no. (%)</th>
<th>Combination-Therapy Group (N = 209) no./total no. (%)</th>
<th>Absolute Difference between Combination Therapy and Metformin % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovulation</td>
<td>462/942 (49.0)</td>
<td>296/1019 (29.0)</td>
<td>582/964 (60.4)</td>
<td>31.4 (24.7 to 38.0)</td>
</tr>
<tr>
<td>Conception</td>
<td>62/209 (29.7)</td>
<td>25/208 (12.0)</td>
<td>80/209 (38.3)</td>
<td>26.3 (18.4 to 34.2)</td>
</tr>
<tr>
<td>Pregnancy</td>
<td>50/209 (23.9)</td>
<td>18/208 (8.7)</td>
<td>65/209 (31.1)</td>
<td>22.4 (15.0 to 29.8)</td>
</tr>
<tr>
<td>Singleton</td>
<td>47/50 (94.0)</td>
<td>18/18 (100.0)</td>
<td>63/65 (96.9)</td>
<td>−3.1 (−7.3 to 1.1)</td>
</tr>
<tr>
<td>Twins</td>
<td>2/50 (4.0)</td>
<td>0</td>
<td>2/65 (3.1)</td>
<td>−3.1 (−10.1 to 16.3)</td>
</tr>
<tr>
<td>Triplets</td>
<td>1/50 (2.0)</td>
<td>0</td>
<td>0</td>
<td>0 (−12.7 to 12.7)</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 (−12.7 to 12.7)</td>
</tr>
<tr>
<td>Live birth</td>
<td>47/209 (22.5)</td>
<td>15/208 (7.2)</td>
<td>56/209 (26.8)</td>
<td>19.6 (12.6 to 26.6)</td>
</tr>
<tr>
<td>Pregnancy loss</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total losses among subjects who conceived</td>
<td>16/62 (25.8)</td>
<td>10/25 (40.0)</td>
<td>24/80 (30.0)</td>
<td>−10.0 (−31.7 to 11.7)</td>
</tr>
<tr>
<td>Loss in first trimester</td>
<td>14/62 (22.6)</td>
<td>10/25 (40.0)</td>
<td>20/80 (25.0)</td>
<td>−14.5 (−35.9 to 6.9)</td>
</tr>
<tr>
<td>Biochemical factor or no fetal heart motion</td>
<td>10/62 (16.1)</td>
<td>7/25 (28.0)</td>
<td>13/80 (16.2)</td>
<td>−11.7 (−31.1 to 7.7)</td>
</tr>
<tr>
<td>Ectopic pregnancy</td>
<td>2/62 (3.2)</td>
<td>0</td>
<td>2/80 (2.5)</td>
<td>2.5 (−7.8 to 12.8)</td>
</tr>
<tr>
<td>Loss after observed heart motion</td>
<td>2/62 (3.2)</td>
<td>3/25 (12.0)</td>
<td>5/80 (6.2)</td>
<td>−5.7 (−19.5 to 8.1)</td>
</tr>
<tr>
<td>Loss in second or third trimester</td>
<td>2/62 (3.2)</td>
<td>0</td>
<td>4/80 (5.0)</td>
<td>5.0 (−5.7 to 15.7)</td>
</tr>
<tr>
<td>Events among ovulated cycles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conception</td>
<td>62/462 (13.4)</td>
<td>25/296 (8.4)</td>
<td>80/582 (13.7)</td>
<td>5.4 (1.2 to 9.6)</td>
</tr>
<tr>
<td>Singleton pregnancy</td>
<td>47/462 (10.2)</td>
<td>18/296 (6.1)</td>
<td>63/582 (10.8)</td>
<td>4.7 (1.0 to 8.4)</td>
</tr>
<tr>
<td>Singleton live birth</td>
<td>47/462 (10.2)</td>
<td>15/296 (5.1)</td>
<td>56/582 (9.6)</td>
<td>4.5 (1.0 to 8.0)</td>
</tr>
<tr>
<td>Events among subjects who ovulated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conception</td>
<td>62/157 (39.5)</td>
<td>25/115 (21.7)</td>
<td>80/174 (46.0)</td>
<td>24.3 (13.7 to 34.9)</td>
</tr>
<tr>
<td>Singleton pregnancy</td>
<td>47/157 (29.9)</td>
<td>18/115 (15.7)</td>
<td>63/174 (36.2)</td>
<td>20.5 (10.7 to 30.3)</td>
</tr>
<tr>
<td>Singleton live birth</td>
<td>47/157 (29.9)</td>
<td>15/115 (13.0)</td>
<td>56/174 (32.2)</td>
<td>19.2 (9.9 to 28.5)</td>
</tr>
<tr>
<td>Ovulation per monthly visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit 1</td>
<td>90/209 (43.1)</td>
<td>45/208 (21.6)</td>
<td>109/209 (52.2)</td>
<td>30.6 (21.8 to 39.4)</td>
</tr>
<tr>
<td>Visit 2</td>
<td>90/181 (49.7)</td>
<td>72/189 (38.1)</td>
<td>114/185 (61.6)</td>
<td>23.5 (13.6 to 33.4)</td>
</tr>
<tr>
<td>Visit 3</td>
<td>86/159 (54.1)</td>
<td>44/169 (26.0)</td>
<td>97/163 (59.5)</td>
<td>33.5 (23.5 to 43.5)</td>
</tr>
<tr>
<td>Visit 4</td>
<td>70/141 (49.6)</td>
<td>49/153 (32.0)</td>
<td>92/144 (63.9)</td>
<td>31.9 (21.1 to 42.7)</td>
</tr>
<tr>
<td>Visit 5</td>
<td>58/119 (48.7)</td>
<td>38/138 (27.5)</td>
<td>81/125 (64.8)</td>
<td>37.3 (26.1 to 48.5)</td>
</tr>
<tr>
<td>Visit 6</td>
<td>55/99 (55.6)</td>
<td>38/120 (31.7)</td>
<td>74/106 (69.8)</td>
<td>38.1 (26.0 to 50.2)</td>
</tr>
<tr>
<td>After Visit 6</td>
<td>13/34 (38.2)</td>
<td>10/42 (23.8)</td>
<td>15/32 (46.9)</td>
<td>23.1 (1.5 to 44.7)</td>
</tr>
<tr>
<td>Live birth per monthly visit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit 1</td>
<td>10/209 (4.8)</td>
<td>3/208 (1.4)</td>
<td>9/209 (4.3)</td>
<td>2.9 (−0.3 to 6.1)</td>
</tr>
<tr>
<td>Visit 2</td>
<td>10/199 (5.0)</td>
<td>3/205 (1.5)</td>
<td>5/200 (2.5)</td>
<td>1.0 (−1.7 to 3.7)</td>
</tr>
<tr>
<td>Visit 3</td>
<td>9/189 (4.8)</td>
<td>1/202 (0.5)</td>
<td>11/195 (5.6)</td>
<td>5.1 (1.7 to 8.5)</td>
</tr>
<tr>
<td>Visit 4</td>
<td>10/180 (5.6)</td>
<td>2/201 (1.0)</td>
<td>8/184 (4.3)</td>
<td>3.3 (0.1 to 6.5)</td>
</tr>
<tr>
<td>Visit 5</td>
<td>2/170 (1.2)</td>
<td>4/199 (2.0)</td>
<td>6/176 (3.4)</td>
<td>1.4 (−1.9 to 4.7)</td>
</tr>
<tr>
<td>Visit 6</td>
<td>6/168 (3.6)</td>
<td>2/195 (1.0)</td>
<td>15/170 (8.8)</td>
<td>7.8 (3.3 to 12.3)</td>
</tr>
<tr>
<td>After Visit 6</td>
<td>0</td>
<td>0</td>
<td>2/155 (1.3)</td>
<td>1.3 (−1.5 to 4.1)</td>
</tr>
</tbody>
</table>

\(^a\) Ovulation was defined as a serum progesterone level of more than 5 ng per milliliter. Conception was defined as any positive serum level of human chorionic gonadotropin. Pregnancy was defined as an intrauterine pregnancy with fetal heart motion, as determined by transvaginal ultrasonography. Live birth was defined as the delivery of a viable infant.
<table>
<thead>
<tr>
<th>P Value</th>
<th>Absolute Difference between Combination Therapy and Clomiphene</th>
<th>P Value</th>
<th>Absolute Difference between Clomiphene and Metformin</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.001</td>
<td>11.4 (4.2 to 18.4)</td>
<td>0.003</td>
<td>20.0 (9.1 to 30.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>8.6 (−0.4 to 17.6)</td>
<td>0.06</td>
<td>17.7 (10.1 to 25.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>7.2 (−1.3 to 15.7)</td>
<td>0.10</td>
<td>15.2 (8.3 to 22.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0.96</td>
<td>2.9 (−4.9 to 10.7)</td>
<td>0.45</td>
<td>−6.0 (−12.6 to 0.6)</td>
<td>0.95</td>
</tr>
<tr>
<td>1.0</td>
<td>−0.9 (−9.8 to 8.0)</td>
<td>1.0</td>
<td>4.0 (−9.9 to 17.9)</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>−2.0 (−12.7 to 12.7)</td>
<td>1.0</td>
<td>2.0 (−11.5 to 15.5)</td>
<td>1.0</td>
</tr>
<tr>
<td>1.0</td>
<td>0 (−6.4 to 6.4)</td>
<td>1.0</td>
<td>0 (−13.0 to 13.0)</td>
<td>1.0</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>4.3 (−4.0 to 12.6)</td>
<td>0.31</td>
<td>15.3 (8.6 to 22.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0.35</td>
<td>4.2 (−10.6 to 19.0)</td>
<td>0.58</td>
<td>−14.2 (−36.3 to 7.9)</td>
<td>0.19</td>
</tr>
<tr>
<td>0.15</td>
<td>2.9 (−11.2 to 17.0)</td>
<td>0.74</td>
<td>−17.4 (−39.2 to 4.4)</td>
<td>0.10</td>
</tr>
<tr>
<td>0.18</td>
<td>0.2 (−12.0 to 12.4)</td>
<td>0.98</td>
<td>−11.9 (−31.7 to 7.9)</td>
<td>0.21</td>
</tr>
<tr>
<td>0.57</td>
<td>−0.7 (−8.0 to 6.6)</td>
<td>0.80</td>
<td>3.2 (−7.7 to 14.1)</td>
<td>0.97</td>
</tr>
<tr>
<td>0.35</td>
<td>3.1 (−3.8 to 10.0)</td>
<td>0.42</td>
<td>−8.8 (−22.3 to 4.7)</td>
<td>0.14</td>
</tr>
<tr>
<td>0.57</td>
<td>1.8 (−6.2 to 9.8)</td>
<td>0.61</td>
<td>3.2 (−7.7 to 14.1)</td>
<td>0.97</td>
</tr>
<tr>
<td>0.02</td>
<td>0.4 (−3.8 to 4.6)</td>
<td>0.88</td>
<td>5.0 (0.6 to 9.4)</td>
<td>0.03</td>
</tr>
<tr>
<td>0.02</td>
<td>0.6 (−3.1 to 4.3)</td>
<td>0.73</td>
<td>4.1 (0.2 to 8.0)</td>
<td>0.05</td>
</tr>
<tr>
<td>0.02</td>
<td>−0.6 (−4.3 to 3.1)</td>
<td>0.77</td>
<td>5.1 (1.4 to 8.8)</td>
<td>0.01</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>6.5 (−4.1 to 17.1)</td>
<td>0.23</td>
<td>17.8 (7.1 to 28.5)</td>
<td>0.002</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>6.3 (−3.8 to 16.4)</td>
<td>0.23</td>
<td>14.2 (4.4 to 24.0)</td>
<td>0.006</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>2.3 (−7.7 to 12.3)</td>
<td>0.66</td>
<td>16.9 (7.5 to 26.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>9.1 (−0.4 to 18.6)</td>
<td>0.06</td>
<td>21.5 (12.8 to 30.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>11.9 (1.8 to 22.0)</td>
<td>0.02</td>
<td>11.6 (1.6 to 21.6)</td>
<td>0.02</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>5.4 (−5.4 to 16.2)</td>
<td>0.33</td>
<td>28.1 (17.9 to 38.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>14.3 (2.9 to 25.7)</td>
<td>0.02</td>
<td>17.6 (6.5 to 28.7)</td>
<td>0.002</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>16.1 (3.8 to 28.4)</td>
<td>0.01</td>
<td>21.2 (9.5 to 32.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>14.2 (1.1 to 27.3)</td>
<td>0.03</td>
<td>23.9 (11.1 to 36.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0.04</td>
<td>8.7 (−15.1 to 32.5)</td>
<td>0.48</td>
<td>14.4 (−6.4 to 35.2)</td>
<td>0.17</td>
</tr>
<tr>
<td>0.07</td>
<td>−0.5 (−4.5 to 3.5)</td>
<td>0.81</td>
<td>3.4 (0.1 to 6.7)</td>
<td>0.04</td>
</tr>
<tr>
<td>0.50</td>
<td>−2.5 (−6.2 to 1.2)</td>
<td>0.18</td>
<td>3.5 (0 to 7.0)</td>
<td>0.04</td>
</tr>
<tr>
<td>0.001</td>
<td>0.8 (−3.6 to 5.2)</td>
<td>0.70</td>
<td>4.3 (1.1 to 7.5)</td>
<td>0.009</td>
</tr>
<tr>
<td>0.05</td>
<td>−1.3 (−5.8 to 3.2)</td>
<td>0.60</td>
<td>4.6 (1.0 to 8.2)</td>
<td>0.008</td>
</tr>
<tr>
<td>0.53</td>
<td>2.2 (−0.9 to 5.3)</td>
<td>0.28</td>
<td>−0.8 (−3.3 to 1.7)</td>
<td>0.69</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>5.2 (0.1 to 10.3)</td>
<td>0.04</td>
<td>2.6 (−0.5 to 5.7)</td>
<td>0.15</td>
</tr>
<tr>
<td>1.0</td>
<td>1.3 (−1.6 to 4.2)</td>
<td>1.0</td>
<td>0 (−2.1 to 2.1)</td>
<td>1.0</td>
</tr>
</tbody>
</table>
receive either clomiphene alone or a combination of metformin and clomiphene for up to six ovulatory cycles, there were no significant differences in ovulation rates or pregnancy rates between the combination-therapy group and the group that received clomiphene alone, with a cumulative pregnancy rate of 40% in the combination-therapy group and 46% in the clomiphene group (absolute change in the combination-therapy group, −6%; 95% confidence interval [CI], −20 to 7).28

Although we found no significant benefit of the combination of metformin and clomiphene, as compared with clomiphene alone, the possibility of some benefit cannot be excluded. On the basis of the 95% CIs, plausible differences in the live-birth rate between groups range from a 12.6% absolute increase to a 4.2% absolute decrease in
<table>
<thead>
<tr>
<th></th>
<th>Clomiphene</th>
<th>Metformin</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHBG — nmol/liter</td>
<td>&lt;0.001</td>
<td>0.14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>204</td>
<td>202</td>
<td>204</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>13.2±15.8 (11.0 to 15.4)</td>
<td>3.3±13.4 (1.5 to 5.2)</td>
<td>18.3±25.0 (14.8 to 21.7)</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total testosterone — ng/dl</td>
<td>0.23</td>
<td>&lt;0.001</td>
<td>0.01</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>203</td>
<td>201</td>
<td>204</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>−1.8±29.8 (−5.9 to 2.3)</td>
<td>−7.2±26.3 (−10.9 to −3.6)</td>
<td>−10.1±30.9 (−14.3 to −5.8)</td>
</tr>
<tr>
<td>P value</td>
<td>0.39</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Free androgen index†</td>
<td>&lt;0.001</td>
<td>0.22</td>
<td>0.01</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>203</td>
<td>201</td>
<td>204</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>−3.3±5.3 (−4.0 to −2.6)</td>
<td>−1.6±4.5 (−2.3 to −1.0)</td>
<td>−3.7±5.9 (−4.5 to −2.9)</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Insulin — µU/ml</td>
<td>0.15</td>
<td>0.01</td>
<td>0.29</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>204</td>
<td>202</td>
<td>204</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>7.0±39.7 (1.5 to 12.4)</td>
<td>2.2±34.0 (−2.4 to 6.9)</td>
<td>−0.3±34.7 (−5.1 to 4.4)</td>
</tr>
<tr>
<td>P value</td>
<td>0.01</td>
<td>0.35</td>
<td>0.89</td>
</tr>
<tr>
<td>Proinsulin — pmol/liter</td>
<td>0.26</td>
<td>0.04</td>
<td>0.36</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>203</td>
<td>202</td>
<td>204</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>2.7±28.0 (−1.2 to 6.5)</td>
<td>−0.8±28.1 (−4.7 to 3.0)</td>
<td>−0.1±33.7 (−4.7 to 4.5)</td>
</tr>
<tr>
<td>P value</td>
<td>0.18</td>
<td>0.68</td>
<td>0.96</td>
</tr>
<tr>
<td>Glucose — mg/dl</td>
<td>0.69</td>
<td>0.31</td>
<td>0.17</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>204</td>
<td>201</td>
<td>203</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>2.9±22.9 (−0.3 to 6.0)</td>
<td>−0.4±19.3 (−3.0 to 2.3)</td>
<td>0.7±20.5 (−2.1 to 3.6)</td>
</tr>
<tr>
<td>P value</td>
<td>0.08</td>
<td>0.78</td>
<td>0.61</td>
</tr>
<tr>
<td>HOMA-IR‡</td>
<td>0.23</td>
<td>0.006</td>
<td>0.14</td>
</tr>
<tr>
<td>No. of subjects</td>
<td>204</td>
<td>201</td>
<td>203</td>
</tr>
<tr>
<td>Mean change (95% CI)</td>
<td>2.2±12.6 (0.5 to 4.0)</td>
<td>0.7±10.9 (−0.9 to 2.2)</td>
<td>−0.1±11.8 (−1.7 to 1.5)</td>
</tr>
<tr>
<td>P value</td>
<td>0.01</td>
<td>0.40</td>
<td>0.91</td>
</tr>
</tbody>
</table>

* If conception occurred, the last visit was the one before pregnancy was documented. Plus–minus values are means ±SD. To convert the values for glucose to millimoles per liter, multiply by 0.05551. To convert the values for insulin to picomoles per liter, multiply by 6. To convert the values for testosterone to nanomoles per liter, multiply by 0.03467. BMI denotes body-mass index, and SHBG sex hormone–binding globulin.

† The free androgen index was calculated according to the following formula: (total testosterone [nanomoles per liter] ÷ SHBG [nanomoles per liter]) × 100.

‡ A homeostasis model assessment of insulin resistance (HOMA-IR) was calculated according to the formula (insulin×glucose) ÷ 405.
### Table 4. Adverse Events.\textsuperscript{57}

<table>
<thead>
<tr>
<th>Event</th>
<th>Clomiphene Group</th>
<th>Metformin Group</th>
<th>Combination-Therapy Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before conception in subjects who received a study drug</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total no. of subjects</td>
<td>209</td>
<td>208</td>
<td>209</td>
</tr>
<tr>
<td><strong>Serious adverse event</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemorrhagic corpus luteum cyst\textsuperscript{†}</td>
<td>1 (0.5)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hypersensitivity reaction\textsuperscript{‡}</td>
<td>0</td>
<td>1 (0.5)</td>
<td>0</td>
</tr>
<tr>
<td>Bronchitis or back pain\textsuperscript{§}</td>
<td>1 (0.5)</td>
<td>0</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Death\textsuperscript{¶}</td>
<td>0</td>
<td>1 (0.5)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Other adverse event</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal distention\textsuperscript{‖}</td>
<td>45 (21.5)</td>
<td>56 (26.9)</td>
<td>39 (18.7)</td>
</tr>
<tr>
<td>Abdominal pain or discomfort\textsuperscript{**}</td>
<td>110 (52.6)</td>
<td>123 (59.1)</td>
<td>137 (65.6)</td>
</tr>
<tr>
<td>Constipation</td>
<td>32 (15.3)</td>
<td>21 (10.1)</td>
<td>22 (10.5)</td>
</tr>
<tr>
<td>Diarrhea\textsuperscript{**††}</td>
<td>48 (23.0)</td>
<td>135 (64.9)</td>
<td>126 (60.3)</td>
</tr>
<tr>
<td>Dyspepsia\textsuperscript{††}</td>
<td>9 (4.3)</td>
<td>24 (11.5)</td>
<td>14 (6.7)</td>
</tr>
<tr>
<td>Flatulence</td>
<td>38 (18.2)</td>
<td>37 (17.8)</td>
<td>39 (18.7)</td>
</tr>
<tr>
<td>Nausea\textsuperscript{**††}</td>
<td>82 (39.2)</td>
<td>128 (61.5)</td>
<td>138 (66.0)</td>
</tr>
<tr>
<td>Stomach discomfort</td>
<td>8 (3.8)</td>
<td>15 (7.2)</td>
<td>16 (7.7)</td>
</tr>
<tr>
<td>Vomiting\textsuperscript{**††}</td>
<td>28 (13.4)</td>
<td>62 (29.8)</td>
<td>72 (34.4)</td>
</tr>
<tr>
<td>Decreased appetite\textsuperscript{**}</td>
<td>17 (8.1)</td>
<td>27 (13.0)</td>
<td>33 (15.8)</td>
</tr>
<tr>
<td>Back pain</td>
<td>25 (12.0)</td>
<td>22 (10.6)</td>
<td>22 (10.5)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>26 (12.4)</td>
<td>35 (16.8)</td>
<td>34 (16.3)</td>
</tr>
<tr>
<td>Impaired sense of taste</td>
<td>10 (4.8)</td>
<td>11 (5.3)</td>
<td>10 (4.8)</td>
</tr>
<tr>
<td>Headache</td>
<td>92 (44.0)</td>
<td>88 (42.3)</td>
<td>87 (41.6)</td>
</tr>
<tr>
<td>Altered mood or mood swings</td>
<td>32 (15.3)</td>
<td>36 (17.3)</td>
<td>27 (12.9)</td>
</tr>
<tr>
<td>Hot flashes\textsuperscript{††}</td>
<td>58 (27.8)</td>
<td>32 (15.4)</td>
<td>59 (28.2)</td>
</tr>
<tr>
<td>Adnexal pain\textsuperscript{‖}</td>
<td>10 (4.8)</td>
<td>4 (1.9)</td>
<td>12 (5.7)</td>
</tr>
<tr>
<td>Anovulatory bleeding\textsuperscript{‖††}</td>
<td>6 (2.9)</td>
<td>18 (8.7)</td>
<td>7 (3.3)</td>
</tr>
<tr>
<td>Breast tenderness or pain</td>
<td>41 (19.6)</td>
<td>36 (17.3)</td>
<td>47 (22.5)</td>
</tr>
<tr>
<td>Dysmenorrhea or cramps\textsuperscript{‖††}</td>
<td>42 (20.1)</td>
<td>26 (12.5)</td>
<td>43 (20.6)</td>
</tr>
<tr>
<td>Sore throat</td>
<td>13 (6.2)</td>
<td>16 (7.7)</td>
<td>8 (3.8)</td>
</tr>
<tr>
<td>Respiratory tract infection</td>
<td>27 (12.9)</td>
<td>24 (11.5)</td>
<td>16 (7.7)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>38 (18.2)</td>
<td>42 (20.2)</td>
<td>45 (21.5)</td>
</tr>
<tr>
<td><strong>After conception (with observed fetal heart motion)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total no. of subjects</td>
<td>50</td>
<td>18</td>
<td>65</td>
</tr>
<tr>
<td><strong>Serious adverse event before birth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy loss after 12 weeks</td>
<td>2 (4.0)</td>
<td>0</td>
<td>4 (6.2)</td>
</tr>
<tr>
<td>Ectopic pregnancy</td>
<td>2 (4.0)</td>
<td>0</td>
<td>2 (3.1)</td>
</tr>
<tr>
<td>Cervical incompetence or preterm labor\textsuperscript{‡‡}</td>
<td>1 (2.0)</td>
<td>0</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Severe preeclampsia</td>
<td>0</td>
<td>0</td>
<td>2 (3.1)</td>
</tr>
<tr>
<td>Congenital anomaly\textsuperscript{§§}</td>
<td>0</td>
<td>0</td>
<td>2 (3.1)</td>
</tr>
</tbody>
</table>
the combination-therapy group. Furthermore, when ovulation is used as the outcome, the combination of metformin and clomiphene was superior to either clomiphene alone or metformin alone. The pregnancy rates in our trial were lower than those reported by others, perhaps reflecting the inclusion of obese women and the fact that many of the subjects had a long-standing history of infertility. These factors may also have contributed to a high rate of pregnancy complications. Our selection criteria were consistent with both National Institutes of Health criteria and the revised Rotterdam diagnostic criteria for the polycystic ovary syndrome, and more than 90% of our subjects had polycystic ovaries on baseline ultrasonography.

Table 4. (Continued.)

<table>
<thead>
<tr>
<th>Event</th>
<th>Clomiphene Group</th>
<th>Metformin Group</th>
<th>Combination-Therapy Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>number (percent)</td>
<td>number (percent)</td>
<td>number (percent)</td>
</tr>
<tr>
<td>Other adverse event before birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preterm labor</td>
<td>4 (8.0)</td>
<td>1 (5.6)</td>
<td>5 (7.7)</td>
</tr>
<tr>
<td>Mild preeclampsia</td>
<td>6 (12.0)</td>
<td>1 (5.6)</td>
<td>7 (10.8)</td>
</tr>
<tr>
<td>HELLP syndrome</td>
<td>1 (2.0)</td>
<td>0</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Gestational diabetes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet controlled (class A1)</td>
<td>6 (12.0)</td>
<td>1 (5.6)</td>
<td>4 (6.2)</td>
</tr>
<tr>
<td>Insulin required (class A2)</td>
<td>3 (6.0)</td>
<td>1 (5.6)</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Intrauterine growth restriction</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Preterm premature rupture of membranes</td>
<td>1 (2.0)</td>
<td>1 (5.6)</td>
<td>3 (4.6)</td>
</tr>
<tr>
<td>Placental abruption</td>
<td>2 (4.0)</td>
<td>0</td>
<td>2 (3.1)</td>
</tr>
<tr>
<td>Placenta accreta</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Placenta previa</td>
<td>1 (2.0)</td>
<td>0</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Other placental abnormality</td>
<td>1 (2.0)</td>
<td>1 (5.6)</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Other pregnancy complication</td>
<td>6 (12.0)</td>
<td>2 (11.1)</td>
<td>4 (6.2)</td>
</tr>
<tr>
<td>Serious adverse event after birth</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other adverse event after birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postpartum depression requiring intervention</td>
<td>1 (2.0)</td>
<td>0</td>
<td>2 (3.1)</td>
</tr>
<tr>
<td>Endometritis</td>
<td>0</td>
<td>0</td>
<td>3 (4.6)</td>
</tr>
<tr>
<td>Postpartum hemorrhage</td>
<td>2 (4.0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other disorder</td>
<td>3 (6.0)</td>
<td>1 (5.6)</td>
<td>3 (4.6)</td>
</tr>
</tbody>
</table>

* Diagnoses after pregnancy were made by the treating physician. HELLP syndrome denotes hemolysis, elevated liver-enzyme levels, and a low platelet count.
† This event resulted in hospitalization and surgery.
‡ One subject in the metformin group had an anaphylactic reaction during a dinner of shellfish and tuna, resulting in a visit to the emergency department, during which patient was treated with Benadryl and a corticosteroid and discharged home. She took a dose of metformin that evening and continued in the study.
§ The subjects with bronchitis (in the clomiphene group) and back pain (in the combination-therapy group) were hospitalized.
¶ One patient in the metformin group had a fatal subarachnoid hemorrhage. She had received the drug for one cycle and was not pregnant, according to the autopsy report.
‖ P<0.05 for the comparison between combination therapy and metformin.
*** P<0.05 for the comparison between combination therapy and clomiphene.
**** P<0.05 for the comparison between clomiphene and metformin.
††† One subject in the clomiphene group had cervical incompetence and delivered at 37 weeks, and one subject in the combination-therapy group had preterm labor.
§§ One subject, who had severe preeclampsia and nephrolithiasis during her pregnancy, delivered an infant with the Prader–Willi syndrome, and one patient delivered an infant with a congenital diaphragmatic hernia.
¶¶ Preterm premature rupture of membranes is membrane rupture before contractions begin and at less than 37 weeks’ gestation.
was similar in age and BMI to the cohort in a large, multicenter trial that showed a benefit of the insulin sensitizer troglitazone on ovulatory frequency in the polycystic ovary syndrome.15

Our study demonstrates the limitations of relying on ovulation rates as a surrogate for live-birth rates.18,27 We found that pregnancy was approximately twice as likely when ovulation was induced by clomiphene as when it was induced by metformin. Our study did not address mechanisms for improved fecundity per ovulation with clomiphene, as compared with metformin. Multiple follicular recruitment, which is characteristic of the induction of ovulation with clomiphene,31 may have resulted in an increased opportunity for fertilization and successful implantation (as evidenced by multiple pregnancies only in the groups receiving clomiphene), as compared with the presumed monofollicular ovulation rate with metformin. We did not perform routine ultrasonography to monitor follicular development because the addition of such a procedure exceeds the normal standard of care in this setting and because it might have led to unblinding in the presence of multiple follicles.31

Early pregnancy loss may be another mechanism for subfecundity in women with the polycystic ovary syndrome. The observed rate of loss of intrauterine pregnancies in our study was similar to or lower than that observed after in vitro fertilization among women of a similar age range using their own eggs (approximately 13%) on the basis of 2003 data from the Society for Assisted Reproductive Technology.32 Our study was not adequately powered to detect a difference in the first-trimester loss rate between the clomiphene group (22.6%) and the metformin group (40.0%), but our results appear to be inconsistent with those of Palomba et al.,17,20 in which study medication was likewise discontinued after a positive pregnancy test. These investigators reported significantly lower first-trimester loss rates with metformin than with clomiphene (9.7% vs. 37.5%)17 or fertility treatment with laparoscopic ovarian diathermy (9.3% vs. 29%),29 although these results were based on small numbers (six events17 and four events,29 respectively). Another group reported no significant difference in rates of spontaneous abortion between groups treated with clomiphene (11%) and combination therapy (12%).27

Our study cannot address the effects of continuing metformin throughout pregnancy,33 though our findings raise concern and highlight the need for randomized trials for this indication.34 Our results, however, support other studies suggesting an increased rate of pregnancy complications in women with the polycystic ovary syndrome, such as gestational diabetes (12% among subjects in our study, as compared with 2 to 5% in the U.S. population35) and preeclampsia (12% in our study, as compared with 3 to 8% in the U.S. population36), although obesity clearly contributes to these risks.7,30

Subjects in our study received extended-release metformin, and this form of the drug may be less efficacious in women with the polycystic ovary syndrome than is immediate-release metformin.37 Our study demonstrates that the tolerability of extended-release metformin is similar to that of clomiphene, although metformin had more gastrointestinal side effects and fewer vascular side effects and ovulation-related symptoms.

In summary, our study supports the use of clomiphene citrate alone as first-line therapy for infertility in women with the polycystic ovary syndrome. We did not find a significant benefit of combination therapy with clomiphene and metformin over clomiphene alone with respect to the live-birth rate. In addition, the results of our study underscore the limitations of the use of ovulation as a surrogate marker for live birth in infertility trials.

Supported by grants from the National Institutes of Health (U10 HD27049, to Dr. Coutifaris; U01 HD38997, to Dr. Myers; U10 HD39005, to Dr. Diamond; U10 HD27011, to Dr. Carson; U10 HD33172, to Dr. Steinkampf; U10 HD38988, to Dr. Carr; U10 HD38992, to Dr. Legro; U10 HD38998, to Dr. Schlaff; U10 HD38999, to Dr. McGovern; U54 HD29834, to the University of Virginia Center for Research in Reproduction Ligand Assay and Analysis Core; GCRC MO1R00056, to the University of Pittsburgh; and MO1R010732 and C06 RR016499, to Pennsylvania State University). Glucophage XR and matching placebo were provided by Bristol-Myers Squibb.

Dr. Legro reports receiving consulting fees from GlaxoSmithKline, Ferring, and Abbott, lecture fees from Serono, and grant support from Pfizer; Dr. Barnhart, consulting fees from TAP and grant support from Ortho Biotech; Dr. Schlaff, grant support from Organon and Wyeth; Dr. Diamond, consulting fees from TAP and Serono and grant support from Serono, TAP, GlaxoSmithKline, and Merck; Dr. McGovern, grant support from Ferring and Serono; and Dr. Cataldo, consulting fees from Organon. Dr. Nestler reports holding an equity interest in Bristol-Myers Squibb. Dr. Leppert reports receiving grant support from TAP; and Dr. Myers, consulting fees from Merck and TAP and grant support from Merck. No other potential conflict of interest relevant to this article was reported.

We thank the staff at the Ligand Assay and Analysis Core Laboratory at the University of Virginia Center for Research and Reproduction, under the direction of D. Hasenleider, for their contributions.
APPENDIX


REFERENCES

25. Palomba S, Orio F Jr, Nardo LG, et al. Metformin administration versus laparoscopic ovarian diathermy in clomiphene citrate-resistant women with polycystic...
Clomiphene versus Metformin in the Polycystic Ovary Syndrome


Copyright © 2007 Massachusetts Medical Society.
GM-CSF Autoantibodies and Neutrophil Dysfunction in Pulmonary Alveolar Proteinosis

Kanji Uchida, M.D., Ph.D., David C. Beck, M.D., Ph.D., Takashi Yamamoto, M.D., Ph.D., Pierre-Yves Berclaz, M.D., Ph.D., Shuichi Abe, M.D., Ph.D., Margaret K. Staudt, M.S., Brenna C. Carey, Ph.D., Marie-Dominique Filippi, Ph.D., Susan E. Wert, Ph.D., Lee A. Denson, M.D., Jonathan T. Puchalski, M.D., Diane M. Hauck, B.A., M.T., and Bruce C. Trapnell, M.D.

ABSTRACT

BACKGROUND
Increased mortality from infection in patients with pulmonary alveolar proteinosis occurs in association with high levels of autoantibodies against granulocyte–macrophage colony-stimulating factor (GM-CSF). We tested the hypothesis that neutrophil functions are impaired in patients with pulmonary alveolar proteinosis and that GM-CSF autoantibodies cause the dysfunction.

METHODS
We studied 12 subjects with pulmonary alveolar proteinosis, 61 healthy control subjects, and 12 control subjects with either cystic fibrosis or end-stage liver disease. We also studied GM-CSF−/− mice and wild-type mice. We evaluated basal neutrophil functions, neutrophil functions after priming by GM-CSF to augment antimicrobial functions, and the effects of highly purified GM-CSF autoantibodies on neutrophil functions in vitro and in vivo.

RESULTS
Neutrophils from subjects with pulmonary alveolar proteinosis had normal ultrastructure and differentiation markers but impaired basal functions and antimicrobial functions after GM-CSF priming. GM-CSF−/− mice also had reduced basal neutrophil functions, but functions after GM-CSF priming were unimpaired. The neutrophil dysfunction characteristic of pulmonary alveolar proteinosis was reproduced in a dose-dependent fashion in blood specimens from healthy control subjects after incubation with affinity-purified GM-CSF autoantibodies isolated from patients with pulmonary alveolar proteinosis. The injection of mouse GM-CSF antibodies into wild-type mice also caused neutrophil dysfunction.

CONCLUSIONS
The antimicrobial functions of neutrophils are impaired in patients with pulmonary alveolar proteinosis, owing to the presence of GM-CSF autoantibodies. The effects of these autoantibodies show that GM-CSF is an essential regulator of neutrophil functions.
PULMONARY ALVEOLAR PROTEINOSIS IS A RARE DISORDER IN WHICH SURFACTANT ACCUMULATES WITHIN PULMONARY ALVEOLI, CAUSING RESPIRATORY INSUFFICIENCY. THE DISEASE IS SPECIFICALLY ASSOCIATED WITH HIGH LEVELS OF AUTOANTIBODIES AGAINST GRANULOCYTE–MACROPHAGE COLONY-STIMULATING FACTOR (GM-CSF) IN BLOOD AND TISSUES, INCLUDING PULMONARY ALVEOLI. THESE AUTOANTIBODIES NEUTRALIZE THE BILOGIC ACTIVITY OF GM-CSF. IN MICE, GM-CSF STIMULATES THE TERMINAL DIFFERENTIATION OF ALVEOLAR MACROPHAGES, PRIMARILY THROUGH THE ACTION OF THE TRANSCRIPTION FACTOR PU.1. THE HOMOZYGOUS DELETION OF GM-CSF GENES CAUSES PULMONARY ALVEOLAR PROTEINOSIS IN MICE BY IMPAIRING THE CLEARANCE OF PULMONARY SURFACTANT BY ALVEOLAR MACROPHAGES THAT ARE DEPENDENT ON GM-CSF. IN PATIENTS WITH PULMONARY ALVEOLAR PROTEINOSIS, PU.1 LEVELS ARE REDUCED IN ALVEOLAR MACROPHAGES, BUT LEVELS INCREASE IN RESPONSE TO GM-CSF THERAPY. THERE IS EVIDENCE THAT THE REGULATION OF ALVEOLAR MACROPHAGES BY GM-CSF IS SIMILAR IN HUMANS AND MICE.

Infections, which are caused predominantly by opportunistic pathogens, account for 18% of reported deaths attributable to pulmonary alveolar proteinosis. Some of these infections occur at extrapulmonary sites, an indication that the predisposition to infection is systemic, rather than confined to the lungs. Similarly, in GM-CSF−/− mice, mortality from infection and susceptibility to bacterial, fungal, and mycobacterial pathogens are increased. These findings suggest that pulmonary alveolar proteinosis is characterized by defective immune function, especially because GM-CSF is important during infection.

GM-CSF augments the antimicrobial functions of neutrophils by means of a mechanism known as priming. GM-CSF priming increases levels of the adhesion molecule CD11b, which promotes the adhesion of neutrophils to the vascular endothelium, a critical event in the recruitment of neutrophils into infected tissues. GM-CSF also primes phagocytosis, oxidative burst, and bactericidal activity in neutrophils. Since GM-CSF autoantibodies may impair these functions, we studied neutrophils from patients with pulmonary alveolar proteinosis and from GM-CSF−/− mice. We also evaluated the functions of normal human or mouse neutrophils after incubation with highly purified GM-CSF autoantibodies.

METHODS

PARTICIPANTS
The institutional review board of the Cincinnati Children’s Hospital Medical Center approved the study. All participants or their legal guardians gave written informed consent, and minors gave assent. Between July 24, 2004, and June 9, 2006, all 12 subjects with pulmonary alveolar proteinosis who were referred to our center for evaluation or treatment were enrolled in the study. We also studied 61 healthy control subjects, 6 control subjects with cystic fibrosis, and 6 control subjects with end-stage liver disease. (The case histories of the subjects with pulmonary alveolar proteinosis are given in the Supplementary Appendix, available with the full text of this article at www.nejm.org.) GM-CSF autoantibodies were quantified in serum with the use of an enzyme-linked immunosorbent assay (ELISA), as previously described. Granulocyte colony-stimulating factor (G-CSF) was quantified in serum by means of an ELISA (Quantikine Kit, R&D Systems).

HUMAN NEUTROPHILS
The isolation of human neutrophils in blood on Ficoll (MP Biomedicals) gradients was initiated within 1 hour after phlebotomy. After the lysis of red cells, neutrophils were resuspended in phosphate-buffered saline containing 10 mM D-glucose. Mean (±SE) viability, determined immediately after isolation by means of trypan-blue staining, was 97.5±1.2% in neutrophils from subjects with pulmonary alveolar proteinosis and 97.9±0.6% in neutrophils from healthy control subjects. In some experiments, neutrophils were isolated for evaluation with the use of Mono-Poly resolving medium (MP Biomedicals), washed in phosphate-buffered saline (without hypotonic red-cell lysis), and resuspended in Hank’s balanced salt solution, 20% fetal bovine serum, and 25 mM HEPES buffer (pH 7.4); these neutrophils are referred to as washed neutrophils. Cell-surface markers of apoptosis were assessed by flow cytometry (Apoptosis Detection Kit I, Pharmingen).

Neutrophil differentiation was evaluated with the use of cell-surface differentiation markers and flow cytometry. We measured PU.1 messenger RNA levels in neutrophils using reverse transcription and quantitative polymerase-chain-reaction
amplification and evaluated PU.1 protein using Western-blot analysis. Neutrophil ultrastructure was evaluated in 20 neutrophils from each of two subjects with pulmonary alveolar proteinosis (Patients 6 and 7) and from each of two healthy control subjects.

The basal capacity of isolated neutrophils to phagocytose opsonized fluorescent microspheres was evaluated as previously described for macrophages, except that internalization of the microspheres was quantified with the use of confocal microscopy at a magnification of 63×. The phagocytic index was calculated as the percentage of neutrophils that had phagocytosed microspheres multiplied by the number of microspheres per cell. For each human (or mouse) studied, 400 neutrophils were evaluated.

Phagocytosis by neutrophils in whole blood (hereafter called the phagocytic capacity) was measured by flow cytometry to eliminate the potential influence of reduced adhesion. Triplicate samples of heparinized human blood (200 μl) were incubated (at 37°C for 60 minutes) with IgG-opsonized fluorescent microspheres (7.5×10⁶, prepared as previously described) in capped, siliconized microcentrifuge tubes with gentle orbital rotation (Thermomixer, Eppendorf). After red-cell lysis, flow cytometry was performed to evaluate neutrophils.

Phagocytic capacity was calculated as the percentage of neutrophils containing internalized microspheres multiplied by the mean fluorescence intensity of phagocytic neutrophils (both determined with the use of flow cytometry) and multiplied by the neutrophil count in the blood.

Cellular adhesion was evaluated by seeding isolated neutrophils into low-adhesion plastic dishes (Corning, catalog no. 3471). After 2 hours, the plates were washed twice with phosphate-buffered saline, and adherent cells were counted with the use of a microscope. The production of hydrogen peroxide was measured in neutrophils in whole blood as previously described. All data for the basal phagocytic index, phagocytic capacity, oxidative burst, and cellular adhesion were normalized by dividing by the mean value for the healthy control group and multiplying by 100.

The amount of Staphylococcus aureus killed (American Type Culture Collection no. 49476) in 1 hour by neutrophils in whole blood or by isolated neutrophils was determined as previously described, except that interferon-γ and lysozyme, respectively, were omitted.

We studied neutrophil functions after GM-CSF priming by incubating heparinized whole blood for 30 minutes in the absence or presence of 10 ng of human GM-CSF per milliliter (Leukine, Berlex) or 10 ng of mouse GM-CSF per milliliter (R&D Systems). The increase in CD11b levels on neutrophils after GM-CSF priming (CD11b stimulation index) was calculated as the mean fluorescence intensity of CD11b on neutrophils primed by GM-CSF minus that of CD11b on nonprimed neutrophils, divided by the mean fluorescence intensity of nonprimed neutrophils and multiplied by 100. Phagocytic capacity after GM-CSF priming was calculated as the phagocytic capacity of neutrophils primed by GM-CSF minus that of nonprimed neutrophils.

**MOUSE NEUTROPHILS**

Experiments in mice were conducted according to protocols approved by the local institutional animal care and use committee. GM-CSF mice were backcrossed for more than 10 generations into C57BL/6J mice, which served as the wild-type control mice. Mouse neutrophils were isolated and evaluated as described for human neutrophils, except as follows. Neutrophils were isolated from bone marrow on Percoll (Amersham) gradients; more than 90% of the isolated cells were neutrophils, and more than 95% were viable. Phagocytic capacity was determined with the use of 100 μl of blood, 30-minute periods of incubation, and immunostaining of neutrophils with Ly6G.

**EFFECTS OF GM-CSF ANTIBODY**

We made two GM-CSF autoantibody preparations using GM-CSF affinity chromatography as previously described: one from a single subject with pulmonary alveolar proteinosis (Patient 6) and one from serum pooled from 11 subjects with the disease (purified GM-CSF autoantibodies and purified pooled GM-CSF autoantibodies, respectively). Affinity-purified autoantibodies were concentrated with the use of ultrafiltration (Microcon 30 K, Amicon), resuspended in phosphate-buffered saline, and assessed for purity by means of electrophoresis on polyacrylamide gels. We measured the ability of GM-CSF autoantibodies to neutralize GM-CSF, using TF1 cells as previously described.
For in vitro studies, purified GM-CSF autoantibody or human immune globulin (Gammagard, Baxter) was incubated with heparinized blood or washed neutrophils (0.5 μg per milliliter, except as noted) before CD11b levels or phagocytic capacity was assessed. Since human GM-CSF and mouse GM-CSF are not immunologically cross-reactive, for in vivo studies, monoclonal anti–mouse GM-CSF antibody (22E9, Endogen) or isotype-control antibody (rat IgG2a, Pharmingen; 200 μg per mouse) was injected intraperitoneally into C57BL/6J mice (five in each group). The phagocytic capacity of neutrophils was assessed 3 days after injection.

**STATISTICAL ANALYSIS**

We evaluated the numerical data for a normal distribution using the Kolmogorov–Smirnov test and for equal variance using the Levene median test; parametric data are presented as means (±SE) and nonparametric data are presented as medians and interquartile ranges. Statistical comparisons of parametric data were made with Student’s t-test for two-group comparisons and with one-way analysis of variance with post hoc analysis according to the Holm–Sidak method for multiple-group comparisons. Nonparametric data were compared with the use of the Mann–Whitney rank-sum test. P values of less than 0.05 were considered to indicate statistical significance. All experiments were repeated at least twice, with similar results.

**RESULTS**

**NEUTROPHIL COUNTS AND DIFFERENTIATION MARKERS**

Neutrophil counts in blood and G-CSF levels in serum were normal in the 12 subjects with pulmonary alveolar proteinosis (Table 1). Neutrophils from subjects with pulmonary alveolar proteinosis and those from healthy control subjects had similar ultrastructure (Fig. 1A), cell-surface differentiation markers (Table 2), and levels of PU.1 expression (data not shown). Differentiation markers on neutrophils from GM-CSF−/− and wild-type mice were also similar (Table 2).

**NEUTROPHIL FUNCTION**

**Patients**

The basal phagocytic functions of isolated neutrophils, evaluated with the use of confocal microcopy, were reduced in subjects with pulmonary alveolar proteinosis as compared with healthy control subjects (Fig. 1B and Table 2). In whole blood, the percentage of phagocytic neutrophils and the number of phagocytosed microspheres per neutrophil were also reduced in subjects with pulmonary alveolar proteinosis (Fig. 1C and Table 2). Basal cellular adhesion, oxidative burst, and bactericidal activity were reduced in subjects with pulmonary alveolar proteinosis as compared with healthy control subjects (Table 2).

The phagocytic capacity of washed neutrophils incubated in Hank’s balanced salt solution without added GM-CSF declined during a 3-hour period (100.0±0.6 at 0 minutes, 92.3±0.3 at 90 minutes, and 71.8±5.0 at 180 minutes; 3 determinations per time point) (90 minutes vs. 0 minutes, P=0.04; 180 minutes vs. 0 minutes, P<0.001). The decline was not due to apoptosis, which occurred in less than 2% of neutrophils at each time point. In Patient 6, who had pulmonary alveolar proteinosis and received GM-CSF therapy for 13 weeks, the basal phagocytic capacity of neutrophils (normalized to the capacity in 20 healthy control subjects) improved from 44.1±2.2% before therapy to 104.5±1.9% after therapy. Although we studied only three children with pulmonary alveolar proteinosis (Patients 2, 3, and 5), their pattern of neutrophil impairment was similar to that of the adult subjects.

GM-CSF priming in vitro increased CD11b levels on neutrophils in the blood of healthy control subjects to maximum levels at low levels of GM-CSF (Fig. 1E and Table 2). In contrast, GM-CSF priming of CD11b levels on neutrophils from subjects with pulmonary alveolar proteinosis was severely impaired, with levels increasing by only modest amounts at high GM-CSF levels. The phagocytic capacity of neutrophils after GM-CSF priming was also severely impaired in the subjects with pulmonary alveolar proteinosis (Fig. 1F and Table 2).

We also examined basal and GM-CSF–primed neutrophil functions in control subjects with cystic fibrosis or end-stage liver disease, neither of which is associated with GM-CSF autoantibodies. The basal phagocytic capacity and GM-CSF priming of CD11b levels of neutrophils was normal in both disorders (Table 3). Thus, neutrophil dysfunction is not a characteristic feature of these chronic diseases.
**Table 1. Clinical Characteristics of Subjects with Pulmonary Alveolar Proteinosis.**

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Sex</th>
<th>Smoking Status†</th>
<th>Presentation</th>
<th>Microbial Isolate‡</th>
<th>Age at Diagnosis yr</th>
<th>Age at Enrollment mo</th>
<th>Duration of Symptoms mo</th>
<th>Serum GM-CSF Autoantibody§ μg/ml</th>
<th>Serum G-CSF¶ ng/ml</th>
<th>Neutrophils in Blood‖ cells/mm³</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>Former smoker</td>
<td>Brain abscess, lung infiltrates</td>
<td><em>Staphylococcus epidermidis</em></td>
<td>37</td>
<td>41</td>
<td>46</td>
<td>227</td>
<td>0.0</td>
<td>2688</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>Never smoked</td>
<td>Pneumonia, lung abscess</td>
<td><em>Haemophilus influenzae</em></td>
<td>12</td>
<td>12</td>
<td>20</td>
<td>482</td>
<td>4.2</td>
<td>1039</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>Never smoked</td>
<td>Pneumonia</td>
<td>Influenza B</td>
<td>14</td>
<td>15</td>
<td>21</td>
<td>86</td>
<td>15.4</td>
<td>2968</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>Never smoked</td>
<td>Dyspnea</td>
<td>None identified</td>
<td>25</td>
<td>25</td>
<td>18</td>
<td>289</td>
<td>16.2</td>
<td>5184</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>Never smoked</td>
<td>Dyspnea</td>
<td>None identified</td>
<td>15</td>
<td>15</td>
<td>27</td>
<td>39</td>
<td>3.3</td>
<td>3564</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>Former smoker</td>
<td>Pneumonia</td>
<td>None identified</td>
<td>32</td>
<td>36</td>
<td>42</td>
<td>676</td>
<td>15.4</td>
<td>3396</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>Never smoked</td>
<td>Dyspnea</td>
<td>MAC</td>
<td>52</td>
<td>52</td>
<td>13</td>
<td>306</td>
<td>12.8</td>
<td>1686</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>Current smoker</td>
<td>Dyspnea</td>
<td>None identified</td>
<td>39</td>
<td>39</td>
<td>35</td>
<td>760</td>
<td>36.1</td>
<td>3909</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>Current smoker</td>
<td>Asymptomatic, lung infiltrates</td>
<td><em>Nocardi a asteroid es</em></td>
<td>39</td>
<td>43</td>
<td>77</td>
<td>297</td>
<td>0.0</td>
<td>2074</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>Never smoked</td>
<td>Pneumonia</td>
<td><em>N. asteroid es, MAC</em></td>
<td>43</td>
<td>48</td>
<td>69</td>
<td>159</td>
<td>0.0</td>
<td>2744</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>Former smoker</td>
<td>Pneumonia</td>
<td>MAC</td>
<td>65</td>
<td>67</td>
<td>91</td>
<td>451</td>
<td>0.0</td>
<td>3872</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>Former smoker</td>
<td>Pneumonia</td>
<td><em>Mycobacterium tuberculosis</em></td>
<td>58</td>
<td>60</td>
<td>29</td>
<td>118</td>
<td>0.0</td>
<td>3434</td>
</tr>
</tbody>
</table>

* Details of the case histories are provided in the Supplementary Appendix.
† Former smokers had discontinued smoking before evaluation; current smokers were smoking at the time of evaluation.
‡ Pathogens were isolated from a culture of cerebrospinal fluid (Patient 1), lung-lavage fluid (Patients 2, 3, 7, 10, 11, and 12), or chest-wall–biopsy specimen (Patient 9). MAC denotes *Mycobacterium avium* complex.
§ The upper limit of the confidence interval (99th percentile) for the serum GM-CSF autoantibody level in 61 healthy control subjects was 3.52 μg per milliliter.
¶ The median serum G-CSF level in 59 healthy control subjects was 3.8 (interquartile range, 0.0 to 16.8), which did not differ significantly from that in the subjects with pulmonary alveolar proteinosis (P=0.48).
‖ The mean (±SE) neutrophil count in blood specimens from 61 healthy control subjects was 3876±221, which did not differ significantly from that in the subjects with pulmonary alveolar proteinosis (P=0.11).
**GM-CSF−/− Mice**

Basal neutrophil functions were reduced in GM-CSF−/− mice in a pattern similar to that among subjects with pulmonary alveolar proteinosis. Phagocytosis was reduced the most, followed by adhesion, oxidative burst, and bactericidal activity ($r^2=0.94$, $P=0.03$) (Fig. 1D and Table 2).

In contrast to basal functions, GM-CSF–primed neutrophil functions — including CD11b levels and phagocytic capacity — were not impaired in GM-CSF−/− mice (Table 2). This finding contrasts with that in subjects with pulmonary alveolar proteinosis in the absence of GM-CSF priming, but the mechanism underlying the disruption of GM-CSF functions differs: in GM-CSF−/− mice, it is the absence of GM-CSF production; in humans, it is the high levels of neutralizing GM-CSF autoantibodies.

**Effect of GM-CSF Autoantibodies**

Affinity-purified GM-CSF autoantibodies produced a single band similar in size to purified human IgG on polyacrylamide gels under nonreducing conditions and two bands corresponding to the expected sizes of immunoglobulin heavy and light chains under reducing conditions (Fig. 2A). The amounts of autoantibodies required to inhibit the activity of GM-CSF by 50% were similar with purified GM-CSF autoantibodies from a single patient and with purified pooled GM-CSF autoantibodies (10.3 and 10.6 mol of IgG per mole of GM-CSF, respectively); these values were also similar to previously reported values.

In vitro incubation of the GM-CSF autoantibodies with whole blood from healthy control subjects or from wild-type mice blocked GM-CSF priming of CD11b levels on human neutrophils but not mouse neutrophils (Fig. 2B). Phagocytic capacity primed by GM-CSF was also reduced in a dose-dependent fashion by the presence of GM-CSF autoantibodies (Fig. 2C).

A low GM-CSF level (10 pg per milliliter) maintained normal neutrophil phagocytic capacity at 90 minutes as compared with 0 minutes ($P=0.11$) (Fig. 2D). The addition of GM-CSF autoantibodies reduced the phagocytic capacity of washed neutrophils and blocked the ability of GM-CSF to maintain phagocytic functions in neutrophils during short-term, ex vivo incubation (Fig. 2D). Purified GM-CSF autoantibodies from a single patient and purified pooled GM-CSF autoantibodies had similar inhibitory effects on CD11b levels primed with GM-CSF (mean of three replicates, 9.6±4.90 and 0.00±4.83, respectively; $P=0.24$) and on GM-CSF–primed phagocytic capacity (mean of three replicates, −0.84±6.17 and 4.21±0.53, respectively; $P=0.46$).

Injection of a monoclonal mouse GM-CSF antibody into wild-type mice reproduced the impaired neutrophil phagocytic capacity observed in GM-CSF−/− mice (Fig. 2E). The mean serum GM-CSF antibody level at the time of evaluation was 69.3±8.1 μg per milliliter.

**DISCUSSION**

Our study showed impairment of basal phagocytosis, adhesion, oxidative burst, and bactericidal activity of neutrophils from subjects with pulmonary alveolar proteinosis and neutrophils from GM-CSF−/− mice. Neutrophils from the subjects also had impaired responses to GM-CSF priming.
The dysfunction was reproduced in normal human neutrophils by incubating them with highly purified GM-CSF antibodies from subjects with pulmonary alveolar proteinosis and also by injecting mouse GM-CSF autoantibodies into normal mice. The observed constellation of neutrophil abnormalities can explain the increased risk of infection, especially infection with opportunistic organisms, that is associated with pulmonary alveolar proteinosis\(^2,3\) and with other phagocyte...
Table 2. Differentiation Markers and Functions of Neutrophils from Healthy Control Subjects, Subjects with Pulmonary Alveolar Proteinosis (PAP), Wild-Type Mice, and GM-CSF−/− Mice.*

<table>
<thead>
<tr>
<th>Marker or Function</th>
<th>Control Value†</th>
<th>Value for GM-CSF Deficiency‡</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expression of CD antigen (% of cells positive for antigen)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD11b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>13</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>99.0</td>
<td>99.9</td>
<td>0.06</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>98.3 to 99.7</td>
<td>99.4 to 100.0</td>
<td></td>
</tr>
<tr>
<td>CD16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>99.0</td>
<td>99.9</td>
<td>0.28</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>99.0 to 99.7</td>
<td>98.9 to 99.9</td>
<td></td>
</tr>
<tr>
<td>CD18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>94.0</td>
<td>99.0</td>
<td>0.24</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>90.0 to 98.5</td>
<td>87.9 to 99.9</td>
<td></td>
</tr>
<tr>
<td>CD114</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>97.0</td>
<td>98.9</td>
<td>0.18</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>94.3 to 99.1</td>
<td>98.4 to 99.5</td>
<td></td>
</tr>
<tr>
<td>CD116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>75.0</td>
<td>80.3</td>
<td>0.30</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>47.0 to 80.3</td>
<td>65.5 to 94.1</td>
<td></td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD11b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>100</td>
<td>100</td>
<td>0.69</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>92.4 to 100</td>
<td>100 to 100</td>
<td></td>
</tr>
<tr>
<td>CD16 or CD32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>98.1</td>
<td>99.1</td>
<td>0.40</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>97.5 to 98.4</td>
<td>98.1 to 99.2</td>
<td></td>
</tr>
<tr>
<td>Ly6G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>94.7</td>
<td>98.3</td>
<td>0.03</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>90.1 to 96.2</td>
<td>97.9 to 98.7</td>
<td></td>
</tr>
<tr>
<td><strong>Basal phagocytic index§</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±8.2</td>
<td>11.7±1.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±4.7</td>
<td>10.0±0.9</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table 2. (Continued.)

<table>
<thead>
<tr>
<th>Marker or Function</th>
<th>Control Value†</th>
<th>Value for GM-CSF Deficiency‡</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal phagocytic capacity¶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>40</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±3.4</td>
<td>69.4±7.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±6.6</td>
<td>55.8±7.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Adhesion of isolated neutrophils (no. of adherent cells/plate)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±2.1</td>
<td>23.8±2.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±1.5</td>
<td>42.0±1.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Oxidative burst in whole blood (% of mean control value)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>14</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±4.3</td>
<td>54.8±8.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>100.0±1.6</td>
<td>65.6±3.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Bacterial killing in whole blood (% of inoculum)‖</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>8</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>82.6±3.0</td>
<td>61.4±6.8</td>
<td>0.007</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>87.7±1.8</td>
<td>73.6±4.1</td>
<td>0.008</td>
</tr>
<tr>
<td>Bacterial killing time of isolated neutrophils (min)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>21.3±2.6</td>
<td>35.7±4.8</td>
<td>0.03</td>
</tr>
<tr>
<td>CD11b stimulation index††</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>429</td>
<td>−0.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>320 to 572</td>
<td>−7.5 to 4.8</td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>17.1</td>
<td>13.6</td>
<td>0.69</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>12.7 to 22.2</td>
<td>8.73 to 22.5</td>
<td></td>
</tr>
</tbody>
</table>
immunodeficiencies, as well as the increased mortality from infection in GM-CSF−/− mice.

Neutrophil dysfunction in patients with pulmonary alveolar proteinosis was associated with high levels of GM-CSF autoantibodies and could not be attributed to a nonspecific effect of chronic illness. Neutralizing GM-CSF autoantibodies have been reported in patients with myasthenia gravis, but the antibody levels are relatively low, and pulmonary alveolar proteinosis has not been reported in this setting.

G-CSF can also prime the antimicrobial function of neutrophils by means of a mechanism distinct from that in GM-CSF priming. In our study, however, serum G-CSF levels and blood neutrophil counts, which are regulated by G-CSF, were similar in subjects with pulmonary alveolar proteinosis and in healthy subjects. In some patients with Felty’s syndrome, G-CSF autoantibodies are thought to cause neutropenia by neutralizing G-CSF activity; however, pulmonary alveolar proteinosis has not been reported in these pa-

Table 2. (Continued.)

<table>
<thead>
<tr>
<th>Function</th>
<th>Control Value†</th>
<th>Value for GM-CSF Deficiency‡</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phagocytic capacity after GM-CSF priming‡‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>37.1±8.4</td>
<td>1.1±2.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. evaluated</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Mean ±SE</td>
<td>22.1±9.3</td>
<td>13.1±5.3</td>
<td>0.43</td>
</tr>
</tbody>
</table>

* Data for the basal phagocytosis of isolated neutrophils (basal phagocytic index), the basal phagocytosis by neutrophils in whole blood (basal phagocytic capacity), adhesion of isolated neutrophils, and oxidative burst in whole blood (hydrogen peroxide production) were normalized by dividing the mean of the control group and multiplying by 100.
† Control values are given for either healthy control subjects (for comparison with subjects with PAP) or wild-type mice (for comparison with GM-CSF−/− mice).
‡‡ Values for GM-CSF deficiency are given for either subjects with PAP or GM-CSF−/− mice.
§ The phagocytic index was calculated as the percentage of neutrophils that had phagocytosed microspheres multiplied by the number of microspheres per cell.
¶ Phagocytic capacity was calculated as the percentage of neutrophils containing internalized microspheres multiplied by the mean fluorescence intensity of phagocytic neutrophils and multiplied by the neutrophil count in the blood.
‖ Bacterial killing in whole blood was defined as the percentage of Staphylococcus aureus killed in 1 hour by neutrophils in whole blood.
** Bacterial killing time of isolated neutrophils was defined as the time required to kill 50% of the S. aureus inoculum.
†† The CD11b stimulation index was calculated as the mean fluorescence intensity of CD11b on neutrophils primed by GM-CSF minus that of CD11b on nonprimed neutrophils, divided by the mean fluorescence intensity of nonprimed neutrophils and multiplied by 100.
§§ Phagocytic capacity after GM-CSF priming was calculated as the phagocytic capacity of GM-CSF–primed neutrophils minus that of nonprimed neutrophils.

Table 3. Neutrophil Functions in Control Subjects with Cystic Fibrosis or End-Stage Liver Disease, as Compared with Healthy Control Subjects.*

<table>
<thead>
<tr>
<th>Neutrophil Function</th>
<th>Mean Value (±SE)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal phagocytic capacity†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects with cystic fibrosis</td>
<td>91.1±4.0</td>
<td>0.33</td>
</tr>
<tr>
<td>Subjects with end-stage liver disease</td>
<td>90.1±14.3</td>
<td>0.34</td>
</tr>
<tr>
<td>Healthy subjects</td>
<td>100.0±3.4</td>
<td></td>
</tr>
<tr>
<td>CD11b stimulation index‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects with cystic fibrosis</td>
<td>292±179</td>
<td>0.30</td>
</tr>
<tr>
<td>Subjects with end-stage liver disease</td>
<td>400±101</td>
<td>0.59</td>
</tr>
<tr>
<td>Healthy subjects</td>
<td>460±61</td>
<td></td>
</tr>
</tbody>
</table>

* We studied 6 control subjects with cystic fibrosis, 6 with end-stage liver disease, and 40 healthy control subjects for phagocytic capacity and 10 for CD11b level.
† Phagocytic capacity was calculated as the percentage of neutrophils containing internalized microspheres multiplied by the mean fluorescence intensity of phagocytic neutrophils and multiplied by the neutrophil count in the blood.
‡ The CD11b stimulation index was calculated as the mean fluorescence intensity of CD11b on neutrophils primed by GM-CSF minus that of CD11b on nonprimed neutrophils, divided by the mean fluorescence intensity of nonprimed neutrophils and multiplied by 100.
Figure 2. Effect of GM-CSF Autoantibodies on Neutrophil Dysfunction in Vitro and in Vivo.
Panel A shows the results of polyacrylamide-gel electrophoresis of GM-CSF autoantibodies isolated from subjects with pulmonary alveolar proteinosis. Panel B shows the CD11b stimulation index normalized by dividing by the mean value for the pooled human immune globulin control. Panel C shows the mean phagocytic capacity of neutrophils from healthy control subjects after GM-CSF priming and incubation with GM-CSF autoantibodies. The dashed line indicates the capacity of nonprimed neutrophils; phagocytic capacity after GM-CSF priming was calculated as that of primed neutrophils minus that of nonprimed neutrophils. Panel D shows the phagocytic capacity of washed neutrophils from healthy control subjects incubated in Hank’s balanced salt solution for 90 minutes as a percentage of the normal phagocytic capacity at 0 minutes (dashed line). All comparisons indicated by brackets in Panel D were significant (P<0.001). Panel E shows the basal phagocytic capacity of neutrophils in whole blood. In Panels B, C, D, and E, T bars and I bars indicate standard errors.
patients. Therefore, G-CSF and GM-CSF have distinct effects on neutrophil functions. Levels of macrophage colony-stimulating factor (M-CSF), which has regulatory effects on myeloid cells that overlap with the regulatory effects of GM-CSF, are increased in patients with pulmonary alveolar proteinosis and in GM-CSF−/− mice. However, since M-CSF receptors are not present on neutrophils, elevated M-CSF levels in patients with pulmonary alveolar proteinosis are unlikely to influence neutrophil functions.

We found that GM-CSF priming of neutrophil functions was blocked in patients with pulmonary alveolar proteinosis but not in GM-CSF−/− mice, although the pattern of basal neutrophil dysfunction was similar between the two. In humans and mice, abrogation of GM-CSF signaling (by means of gene knockout in mice and by means of antibodies in patients with pulmonary alveolar proteinosis) reduces, but does not abolish, multiple neutrophil functions. Moreover, GM-CSF appears to be unnecessary for neutrophil differentiation: neutrophil morphology, cell-surface differentiation markers, and levels of expression of PU.1 (a transcription factor that is critical for neutrophil differentiation) were similar in subjects with pulmonary alveolar proteinosis and healthy control subjects. This evidence contrasts with the fact that GM-CSF stimulates the terminal differentiation of alveolar macrophages, primarily by increasing levels of PU.1 expression.

The GM-CSF receptor mediates the dose-dependent effects of GM-CSF in a mutually exclusive, reciprocal fashion through two β-chain residues: Ser585 at low GM-CSF levels (in general, <300 pg per milliliter) and Tyr577 at high levels (>300 pg per milliliter). Thus, the GM-CSF receptor acts as a binary switch that promotes cell survival at low GM-CSF levels but also stimulates proliferation and antimicrobial functions (including an increased CD11b level on the cell surface) at high levels. This arrangement of the GM-CSF receptor places the level of GM-CSF bioactivity at the center of the mechanism that regulates neutrophil functions.

Our findings, together with those in previously published reports, provide support for the feasibility of GM-CSF therapy to augment innate immune defenses in patients with serious infections. Conversely, therapy with a humanized monoclonal GM-CSF antibody could be of use in reducing neutrophil priming (and functional capacity) in patients with chronic inflammatory disorders characterized by increased numbers of activated neutrophils, such as severe neutrophilic asthma, cystic fibrosis, and the adult respiratory distress syndrome. It is relevant that GM-CSF antibodies reduce neutrophilic pulmonary inflammation in a dose-dependent fashion in endotoxin-exposed mice. Finally, assessment of blood neutrophil functions may provide convenient, functional surrogate-outcome measures for use in clinical trials evaluating new therapies for pulmonary alveolar proteinosis.

Supported in part by grants from the National Heart, Lung, and Blood Institute (HL071823 and HL69459, to Dr. Trapnell), and T32HD043005, to Dr. Puchalski and the National Center for Research Resources and the National Institutes of Health Office of Rare Diseases (RR019498, to Dr. Trapnell).

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

We thank Drs. Elizabeth Allen (Ohio State University, Columbus), Hugh Black (Carolinas Medical Center, Charlotte, NC), J.-P. Clancy (University of Alabama, Birmingham), David Dortin (Medical Associates of Cincinnati, Cincinnati) Jeffrey Hammersley (Medical University of Ohio, Toledo), David Kamelhar (New York University Medical Center, New York) Vijay Patel (Newnan Hospital, Newnan, GA), and Robert Smith (Dayton Chest Medicine, Dayton, OH) for referring their patients to us; Drs. Robert Wood (Cincinnati Children’s Hospital Medical Center, Cincinnati), John Howington, and Michael Reed (University of Cincinnati Medical Center, Cincinnati) for contributions to the care of these patients; Diane Black for technical support; Susan Radtke and Carrie Stevens (Translational Research Trials Office, Cincinnati Children’s Research Foundation, Cincinnati) for help with the clinical protocols; Carrie Jennings (University of Cincinnati Medical Center, Cincinnati) for identifying patients with chronic liver disease; Dr. Glenn Dranoff (Harvard Medical School, Boston) for providing the GM-CSF−/− mice; and Drs. David Williams and Jeffrey Whitsett (Divisions of Experimental Hematology and Pulmonary Biology, respectively, Cincinnati Children’s Hospital Medical Center, Cincinnati) for critical reading of the manuscript.

REFERENCES


578 N ENGL J MED 356:6 WWW.NEJM.ORG FEBRUARY 8, 2007

Copyright © 2007 Massachusetts Medical Society.
A Human Interleukin-12/23 Monoclonal Antibody for the Treatment of Psoriasis

Gerald G. Krueger, M.D., Richard G. Langley, M.D., Craig Leonardi, M.D., Newman Yeilding, M.D., Cynthia Guzzo, M.D., Yuhua Wang, Ph.D., Lisa T. Dooley, Dr.P.H., and Mark Lebwohl, M.D., for the CNTO 1275 Psoriasis Study Group*

From the University of Utah, Salt Lake City (G.G.K.); Dalhousie University, Halifax, NS, Canada (R.G.L.); Saint Louis University, St. Louis (C.L.); Centocor, Malvern, PA (N.Y., C.G., Y.W., L.T.D.); and Mount Sinai School of Medicine, New York (M.L.). Address reprint requests to Dr. Krueger at 30 N. 1900 East, Suite 4B-454, Salt Lake City, UT 84132, or to gerald.krueger@derm.med.utah.edu.

*The investigators for the CNTO 1275 Psoriasis Study Group are listed in the Appendix.

ABSTRACT

BACKGROUND
Skin-infiltrating lymphocytes expressing type 1 cytokines have been linked to the pathophysiology of psoriasis. We evaluated the safety and efficacy of a human interleukin-12/23 monoclonal antibody in treating psoriasis.

METHODS
In this double-blind, placebo-controlled trial, 320 patients with moderate-to-severe plaque psoriasis underwent randomization to treatment with the interleukin-12/23 monoclonal antibody (one 45-mg dose, one 90-mg dose, four weekly 45-mg doses, or four weekly 90-mg doses) or placebo; 64 patients were randomly assigned to each group. Patients assigned to the interleukin-12/23 monoclonal antibody received one additional dose at week 16 if needed. Patients assigned to placebo crossed over to receive one 90-mg dose of interleukin-12/23 monoclonal antibody at week 20.

RESULTS
There was at least 75% improvement in the psoriasis area-and-severity index at week 12 (the primary end point) in 52% of patients who received 45 mg of the interleukin-12/23 monoclonal antibody, in 59% of those who received 90 mg, in 67% of those who received four weekly 45-mg doses, and in 81% of those who received four weekly 90-mg doses, as compared with 2% of those who received placebo (P<0.001 for each comparison), and there was at least 90% improvement in 23%, 30%, 44%, and 52%, respectively, of patients who received the monoclonal antibody as compared with 2% of patients who received placebo (P<0.001 for each comparison). Adverse events occurred in 79% of patients treated with the interleukin-12/23 monoclonal antibody as compared with 72% of patients in the placebo group (P=0.19). Serious adverse events occurred in 4% of patients who received the monoclonal antibody and in 1% of those who received placebo (P=0.69).

CONCLUSIONS
This study demonstrates the therapeutic efficacy of an interleukin-12/23 monoclonal antibody in psoriasis and provides further evidence of a role of the interleukin-12/23 p40 cytokines in the pathophysiology of psoriasis. Larger studies are needed to determine whether serious adverse events might limit the clinical usefulness of this new therapeutic target. (ClinicalTrials.gov number, NCT00320216.)
Psoriasis is a chronic inflammatory skin disorder affecting 2 to 3% of the world’s population.¹⁻³ Psoriasis affects the physical and emotional well-being of patients, and its effect on quality of life is similar to that seen with other major medical diseases.³ Significant unmet need remains for safe, highly effective, and convenient treatments. Aberrant type 1 immune responses have been linked to the pathogenesis of psoriasis.⁴⁻⁷ and cytokines that elicit these immune responses (e.g., interleukin-12 and interleukin-23) may represent appropriate therapeutic targets.⁸ Interleukin-12 p40 is overexpressed in psoriasis plaques,⁹ and preclinical studies implicate this cytokine in the pathogenesis of psoriasis.¹⁰,¹¹ Interleukin-23, which also shares the p40 subunit and is overexpressed in psoriasis plaques,¹²,¹³ activates a distinct T-cell lineage that expresses interleukin-17 and that may be pathogenically important.¹⁴,¹⁵

To evaluate the therapeutic effect of blocking interleukin-12 and interleukin-23, a fully human interleukin-12/23 monoclonal antibody (CT10 1275) was developed. This antibody binds with high affinity to the p40 subunit of human interleukin-12 and interleukin-23, neutralizing their bioactivity by blocking interactions with their cognate cell-surface receptors. Early clinical studies showed no dose-limiting toxic effects and suggested potential therapeutic efficacy in certain immune-mediated diseases.¹⁶⁻¹⁸ This phase 2 study evaluated the safety and efficacy of the interleukin-12/23 monoclonal antibody in patients with moderate-to-severe psoriasis.

METHODS

PATIENTS

This study was conducted at 46 sites worldwide after approval of the institutional review boards. The first patient consented to participation on June 25, 2003, and the last patient visit occurred on March 9, 2005. After providing written informed consent, men and women (age, ≥18 years) were eligible if they had a diagnosis of plaque psoriasis for at least 6 months, were candidates for phototherapy or systemic therapy, had a baseline score on the psoriasis area-and-severity index of 12 or higher (on a scale of 0 to 72, with higher scores indicating more severe disease), and had involvement of at least 10% body-surface area.

Patients were ineligible if they had nonplaque forms of psoriasis; had a recent serious systemic or local infection; had active or latent tuberculosis, asthma, or a known malignancy within the previous 5 years (except treated basal-cell skin cancer); had received previous treatment with any agent specifically targeting interleukin-12 or interleukin-23; had received biologic or investigational agents within the previous month or five drug half-lives; had received conventional systemic psoriasis therapy or phototherapy within the previous 4 weeks; or had received topical psoriasis treatment within the previous 2 weeks.

STUDY DESIGN

This placebo-controlled, double-blind, parallel-group, phase 2 study evaluated the safety and efficacy of single and multiple doses of the interleukin-12/23 monoclonal antibody for the treatment of psoriasis. With the use of adaptive treatment allocation based on a biased-coin minimization algorithm,¹⁹ patients were randomly assigned to one 45-mg subcutaneous dose of the interleukin-12/23 monoclonal antibody, one 90-mg dose, four weekly 45-mg doses, four weekly 90-mg doses, or placebo. Randomization was stratified by investigational site and by the weight of the patient relative to 95 kg. At week 16, patients with a physician’s global assessment (termed PGA) of less than excellent (2 or 3 on a scale on which 1 is best and 6 is worst) received one additional injection of their originally assigned dose. At week 20, patients in the placebo group crossed over to receive one 90-mg injection of the interleukin-12/23 monoclonal antibody.

EFFICACY AND SAFETY EVALUATIONS

Efficacy evaluations were performed through week 32. The psoriasis area-and-severity index (termed PASI) combines assessments of the extent of body-surface involvement in four anatomical regions (head, trunk, arms, and legs) and the severity of desquamation, erythema, and plaque induration (thickness) in each region, yielding an overall score of 0 (no psoriasis) to 72 (severe disease).²⁰ The physician’s global assessment rates the patient’s psoriasis overall relative to baseline as 1 (clear), 2 (excellent), 3 (good), 4 (fair), 5 (poor), or 6 (worse) and considers involvement of body-surface area, induration, scaling, and erythema. The 10-item Dermatology Life Quality Index questionnaire, completed by the patient, measures whether psoriasis has an effect on the patient’s quality of life, with scores ranging from 0 (“not at all”) to 30 (“very much”).
We assessed the safety and tolerability of the interleukin-12/23 monoclonal antibody by monitoring adverse events and routine laboratory values through week 36. Serum samples collected through week 52 were tested for antibodies to the interleukin-12/23 monoclonal antibody with the use of an antigen-bridging enzyme immunoassay.16

STATISTICAL ANALYSIS

Efficacy data from all patients who underwent randomization were analyzed according to the treatment group. Missing values at week 12 were replaced with the most recently available values for all efficacy variables; missing data at other time points were not imputed. Patients in the placebo group were included in efficacy summaries after week 20 if they had crossed over to receive the interleukin-12/23 monoclonal antibody. Safety data were summarized according to the actual treatment received.

The data-analysis plan was specified before the treatment assignments were revealed. According to these prespecified rules, the proportions of patients responding to treatment were compared using the Cochran–Mantel–Haenszel chi-square test stratified according to low and high body weight. Continuous response variables were compared with the use of analysis of covariance for van der Waerden normal scores21 with weight as a binary covariate. We used Fisher’s exact test to compare safety-event rates. All statistical tests were two-sided and were performed at an alpha level of 0.05.

The primary efficacy end point — the proportion of patients achieving at least 75% improvement from baseline in the psoriasis area-and-severity index at week 12 — was analyzed on an intention-to-treat basis to include data from all patients who had undergone randomization; no interim analyses were performed. Patients who discontinued study treatment because of lack of efficacy or loss of response or who used prohibited medications were considered to have had no response. An overall type I error rate of 0.05 was maintained by performing four pairwise comparisons, one between each active-treatment group and placebo sequentially, from the group that received the highest dose to the group that received the lowest dose of the interleukin-12/23 monoclonal antibody.22 If a comparison was not significant at an alpha level of 0.05, subsequent sequential comparisons were considered non-significant. Assuming 60 patients per group and response rates of 50% for each treatment group and 10% for the placebo group, this study had more than 99% power at the 5% level of significance to detect at least one pairwise treatment effect.

The steering committee of the CNTO 1275 Psoriasis Study Group and Centocor designed and conducted this study. Centocor analyzed the data, and the steering committee and Centocor jointly interpreted the data and contributed to the manuscript. The academic authors had full access to the data, and all authors vouch for the integrity and completeness of the data and analyses.

RESULTS

CHARACTERISTICS OF THE PATIENTS

Most of the 320 patients randomly assigned to the five trial groups (Fig. 1) were men. Baseline demographics and disease characteristics were similar among the trial groups (Table 1). Among the five groups, the average body-surface area affected by psoriasis was approximately 25% and the average duration of psoriasis ranged from approximately 17 to 20 years.

EFFICACY

Thirty-three of 64 patients (52%) treated with one 45-mg dose of the interleukin-12/23 monoclonal antibody, 38 of 64 (59%) treated with one 90-mg dose, 43 of 64 (67%) treated with four weekly 45-mg doses, and 52 of 64 (81%) treated with four weekly 90-mg doses had at least 75% improvement in the psoriasis area-and-severity index at week 12 (the primary end point), as compared with only 1 of 64 patients (2%) receiving placebo (P<0.001 for each comparison) (Table 2

Figure 1 (facing page). Enrollment and Treatment of Patients in the Study.

At week 16, all patients in the active-treatment groups who had a physician’s global assessment of 3 or higher received the assigned treatment, and those who had an assessment of less than 3 received placebo. The placebo group at week 16 does not include one patient who received a single 90-mg dose of interleukin-12/23 monoclonal antibody. At week 20, all patients in the placebo group were expected to receive a single 90-mg dose of interleukin-12/23 monoclonal antibody, but one patient received placebo in error.
487 Patients screened

320 Underwent randomization

64 Assigned to placebo
- 13 Discontinued study agent
  - 6 Had unsatisfactory therapeutic effect
  - 2 Withdrew consent
  - 1 Was lost to follow-up
  - 4 Had other reasons

64 Assigned to 45 mg interleukin-12/23 monoclonal antibody
- 1 Received no treatment
- 7 Discontinued study agent
  - 2 Had unsatisfactory therapeutic effect
  - 5 Had adverse event

64 Assigned to 90 mg interleukin-12/23 monoclonal antibody
- 3 Discontinued study agent
  - 2 Had adverse event
  - 1 Withdrew consent

64 Assigned to 4 weekly 45-mg doses interleukin-12/23 monoclonal antibody
- 4 Discontinued study agent
  - 1 Had unsatisfactory therapeutic effect
  - 1 Had adverse event
  - 1 Withdrew consent
  - 1 Had other reasons

64 Assigned to 4 weekly 90-mg doses interleukin-12/23 monoclonal antibody
- 2 Discontinued study agent
  - 1 Had adverse event
  - 1 Had other reasons

34 Received placebo
- 3 Discontinued study agent
  - 2 Were lost to follow-up
  - 1 Had other reasons

24 Received placebo
- 2 Discontinued study agent
  - 1 Had adverse event
  - 1 Had other reasons

34 Received placebo
- 2 Discontinued study agent
  - 1 Had unsatisfactory therapeutic effect
  - 1 Had other reasons

43 Received placebo
- 0 Discontinued study agent

49 Received placebo
- 2 Discontinued study agent
  - 1 Had adverse event
  - 1 Had other reasons

50 Assigned to placebo
- 47 Completed study
- 17 Discontinued the study through wk 36
  - 10 Withdrew consent
  - 4 Were lost to follow-up
  - 3 Had other reasons

32 Received assigned treatment
- 13 Discontinued study agent
  - 6 Had unsatisfactory therapeutic effect
  - 2 Withdrew consent
  - 1 Was lost to follow-up
  - 4 Had other reasons

26 Received assigned treatment
- 11 Discontinued the study through wk 36
  - 10 Withdrew consent
  - 1 Was lost to follow-up
  - 3 Had other reasons

18 Received assigned treatment
- 6 Discontinued the study through wk 36
  - 3 Withdrawn consent
  - 3 Had other reasons

11 Received assigned treatment
- 8 Discontinued the study through wk 36
  - 7 Withdrawn consent
  - 1 Was lost to follow-up

47 Completed study
- 47 Crossed over to 90 mg interleukin-12/23 monoclonal antibody

53 Completed study
- 53 Completed study

58 Completed study
- 58 Completed study

52 Completed study
- 52 Completed study

56 Completed study
and Fig. 2A). Similar proportions of patients were assessed as being clear or having an excellent response according to physicians’ global assessments, with all active-treatment groups significantly improved as compared with the placebo group (P<0.001 for each comparison) (Table 2 and Fig. 2B).

Marked clinical responses were observed in significantly more patients in each active-treatment group than in the placebo group, as measured by at least 90% improvement in the psoriasis area-and-severity index at week 12 (P<0.001 for each comparison) (Table 2 and Fig. 2D). Significantly more patients in the three higher-dose groups than in the placebo group had 100% improvement in the psoriasis area-and-severity index or had a physician’s global assessment of clear (P≤0.001 for each comparison with placebo) (Table 2). At least 75% of patients in all active-treatment groups had at least 50% improvement in the psoriasis area-and-severity index score from baseline were significant for all active-treatment groups by week 2 (P<0.001 for each comparison with placebo).

The Dermatology Life Quality Index showed significant improvements in quality of life for all active-treatment groups at week 12 (P<0.001 for each comparison with placebo) (Table 2) and week 24 (Fig. 2E). Significantly more patients in all active-treatment groups reported that their psoriasis had no measurable effect on their quality of life, as indicated by a Dermatology Life Quality Index score of 0 at week 12 (P<0.001 for each comparison with placebo) (Table 2).

Of 237 patients in the interleukin-12/23 monoclonal antibody groups continuing study treatments at week 16, only 87 patients (37%) were eligible to receive one additional dose at week 16 on the basis of a response of less than excellent according to physicians’ global assessments (Fig. 1). Nevertheless, the proportions of patients with at least 50%, 75%, and 90% improvement in the psoriasis area-and-severity index or with an excellent response or better according to physicians’ global assessments remained relatively stable through week 24 (Fig. 2A through 2D). Response rates declined after treatment was discontinued.

### Table 1. Baseline Characteristics of the Patients. *

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Placebo (N=64)</th>
<th>Interleukin-12/23 Monoclonal Antibody (N=256)</th>
<th>P Value †</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45 mg (N=64)</td>
<td>90 mg (N=64)</td>
<td>4× 45 mg (N=64)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>44±14</td>
<td>46±14</td>
<td>46±13</td>
</tr>
<tr>
<td>Male sex (no. %)</td>
<td>46 (72)</td>
<td>38 (59)</td>
<td>47 (73)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>92.8±20.8</td>
<td>94.3±25.5</td>
<td>92.9±19.1</td>
</tr>
<tr>
<td>Duration of psoriasis (yr)</td>
<td>16.9±11.0</td>
<td>19.1±12.3</td>
<td>17.9±11.6</td>
</tr>
<tr>
<td>Involved body-surface area (%)</td>
<td>26.6±18.4</td>
<td>28.5±16.6</td>
<td>26.3±17.6</td>
</tr>
<tr>
<td>Patients with psoriatic arthritis (no. %)</td>
<td>12 (19)</td>
<td>13 (20)</td>
<td>12 (19)</td>
</tr>
<tr>
<td>Psoriasis area-and-severity index score</td>
<td>19.9±8.3</td>
<td>19.0±7.4</td>
<td>18.8±7.3</td>
</tr>
<tr>
<td>Dermatology Life Quality Index score</td>
<td>12.0±7.2</td>
<td>11.9±7.0</td>
<td>13.4±7.3</td>
</tr>
<tr>
<td>Patients treated previously (no. %) ‡</td>
<td>61 (95)</td>
<td>63 (98)</td>
<td>63 (98)</td>
</tr>
<tr>
<td>Topical agent</td>
<td>39 (61)</td>
<td>39 (61)</td>
<td>37 (58)</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD.
† To identify any significant differences in demographic and baseline characteristics among the treatment groups, we calculated P values using the chi-square test for categorical variables and an analysis of variance for van der Waerden normal scores for continuous variables.
‡ No significant differences in sex were found between the placebo group and each of the active-treatment groups. However, significant differences in sex were found between the group receiving 45 mg of interleukin-12/23 monoclonal antibody and the group receiving four weekly 90-mg doses (P=0.007), as well as between the group receiving four weekly 45-mg doses and the group receiving four weekly 90-mg doses (P=0.01).
§ According to the protocol, all topical therapies (except moisturizers and shampoos) had to be discontinued 2 weeks before randomization, and all systemic therapies 4 weeks before randomization.
Efficacy in those patients in the placebo group who received one 90-mg injection of the interleukin-12/23 monoclonal antibody at week 20 mirrored the improvements observed in patients originally assigned to one 90-mg dose at baseline (Fig. 2A through 2D).

**Safety**

Through week 20 (the placebo-controlled portion of the study), 79% of patients who received active treatment had adverse events (mean duration of follow-up, 19.7 weeks) as compared with 72% who received placebo (mean duration of follow-up,
17.9 weeks; \( P = 0.19 \). Adverse events leading to the discontinuation of treatment occurred in 4% of patients who received active treatment as compared with 3% who received placebo (\( P = 1.00 \)). Forty-three percent of patients who received active treatment and 39% of patients who received placebo had infections. We observed no obvious association between a higher dose and an increased rate of adverse events or infections. Serious adverse events, all of which were classified as serious because the patients required hospitalization, were observed in 4% of patients treated with the interleukin-12/23 monoclonal antibody (9 of 252) and 1% of patients in the placebo group (1 of 67, \( P = 0.69 \)). In the groups that received the interleukin-12/23 monoclonal antibody, two patients were hospitalized for infection (one for cellulitis and one for pneumonia), two patients (a 61-year-old man with diabetes and hypertension and a 54-year-old man with diabetes) had myocardial infarctions and required coronary bypass surgery for multivessel coronary artery disease, and one patient (a 59-year-old woman with hypertension and hyperlipidemia) had a stroke and subsequently underwent endarterectomy. In the placebo group, one serious adverse event (aggravated psoriasis) resulted in hospitalization. One patient (1%) in the placebo group had basal-cell skin cancer, and three patients (2%) who received the interleukin-12/23 monoclonal antibody had cancer (one had basal-cell skin cancer, one squamous-cell skin cancer, and one prostate cancer). Injection-site reactions were observed in 2% of injections of the interleukin-12/23 monoclonal antibody (17 of 706) and 2% of injections of placebo (17 of 819) (Table 3).

Through week 20, the proportions of patients who had significant abnormalities in laboratory values for hematologic and blood chemical tests were low and similar between the groups receiving active treatment and the placebo group. A nonsignificantly greater proportion of patients treated with the interleukin-12/23 monoclonal antibody than with placebo had elevated glucose levels while in a nonfasting state (24 of 252 [10%] vs. 3 of 67 [4%, \( P = 0.23 \)). No significant differences between the active-treatment groups and the placebo group were observed in changes in absolute lymphocyte counts or lymphocyte subsets (data not shown).

Adverse events observed in the overall study population after patients in the placebo group crossed over were similar to the pattern of adverse events observed during the placebo-controlled portion of the study. A small number of patients had serious adverse events, including one occurrence each of coronary artery disorder, congestive heart failure, viral syndrome, urinary tract infection, cellulitis in a surgical wound, and elevated liver-enzyme levels. In addition, the patient in the placebo group in whom basal-cell skin cancer was diagnosed had a second basal-cell cancer reported after crossover to active treatment. Of the 293 patients treated with the interleukin-12/23 monoclonal antibody, antibodies to the study agent were detectable in 12 patients (4%) once or more during the 52 weeks of monitoring for these antibodies. Neither the presence of nor the levels of antibodies were associated with injection-site reactions.

**DISCUSSION**

Our findings demonstrate the efficacy and adverse-event profile of an interleukin-12/23 monoclonal antibody in the treatment of psoriasis. They also establish a central role for the interleukin-12/23 p40 cytokines in the pathophysiology of psoria-
sis. The percentage of patients who had at least 75% improvement in the psoriasis area-and-severity index at week 12 increased in a dose-dependent manner. Marked clearance of psoriatic skin lesions was noted in many patients; significant proportions of patients had at least 90% improvement in the psoriasis area-and-severity index, and significant proportions of patients had complete clearance of psoriasis. Almost all patients had clinically meaningful improvement, as measured by 50% improvement in the psoriasis area-and-severity index.24
Table 3. Adverse Events during the Placebo-Controlled Phase (Weeks 0 to 20) and through the End of the Study (Weeks 0 to 36).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Weeks 0 to 20</th>
<th>Placebo Crossover to 90-mg Active Treatment (0–20 wk)</th>
<th>All Four Active-Treatment Groups (0–36 wk)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo (N = 67)</td>
<td>Interleukin-12/23 Monoclonal Antibody (N = 63)</td>
<td>(N = 49)</td>
</tr>
<tr>
<td>Mean duration of follow-up — wk</td>
<td>17.9 19.4 19.3 20.0 20.1 19.7</td>
<td>17.9</td>
<td>16.9</td>
</tr>
<tr>
<td>Patients with at least 1 adverse event — no. (%)†</td>
<td>48 (72) 57 (90) 52 (81) 49 (78) 42 (68) 200 (79) 48 (72) 25 (51) 216 (86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common adverse events — no. of patients (%)‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper respiratory tract infection</td>
<td>14 (21) 16 (25) 20 (31) 9 (14) 11 (18) 56 (22) 14 (21) 6 (12) 78 (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>11 (16) 12 (19) 12 (19) 2 (3) 9 (15) 35 (14) 11 (16) 1 (2) 38 (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>2 (3) 3 (5) 8 (12) 3 (5) 4 (6) 18 (7) 2 (3) 4 (8) 22 (9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinusitis</td>
<td>3 (4) 3 (5) 4 (6) 3 (5) 3 (5) 13 (5) 3 (4) 1 (2) 20 (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injury</td>
<td>4 (6) 2 (3) 3 (5) 1 (2) 2 (3) 8 (3) 4 (6) 1 (2) 19 (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pruritus</td>
<td>2 (3) 6 (10) 3 (5) 5 (8) 3 (5) 17 (7) 2 (3) 0 18 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhinitis</td>
<td>3 (4) 4 (6) 5 (8) 5 (8) 1 (2) 15 (6) 3 (4) 1 (2) 18 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diarrhea</td>
<td>0 3 (5) 4 (6) 4 (6) 2 (3) 13 (5) 0 1 (2) 17 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myalgia</td>
<td>2 (3) 4 (6) 4 (6) 3 (5) 1 (2) 12 (5) 2 (3) 0 15 (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>1 (1) 3 (5) 3 (5) 2 (3) 3 (5) 11 (4) 1 (1) 0 14 (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purpura</td>
<td>2 (3) 4 (6) 5 (8) 1 (2) 3 (5) 13 (5) 2 (3) 0 14 (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>0 3 (5) 1 (2) 2 (3) 2 (3) 8 (3) 0 0 14 (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td>3 (4) 1 (2) 4 (6) 1 (2) 4 (6) 10 (4) 3 (4) 1 (2) 13 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharyngitis</td>
<td>4 (6) 1 (2) 1 (2) 5 (8) 3 (5) 10 (4) 4 (6) 2 (4) 12 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain</td>
<td>0 3 (5) 2 (3) 4 (6) 1 (2) 10 (4) 0 0 10 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggravated psoriasis</td>
<td>4 (6) 2 (3) 0 0 1 (2) 3 (1) 4 (6) 0 7 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggravated arthritis</td>
<td>5 (7) 1 (2) 0 1 (2) 0 2 (1) 5 (7) 0 3 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse events leading to withdrawal of study agent — no. of patients (%)†</td>
<td>2 (3) 8 (13) 1 (2) 1 (2) 1 (2) 11 (4) 2 (3) 0 12 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serious adverse events — no. of patients (%)‡‡</td>
<td>1 (1) 3 (5) 1 (2) 2 (3) 3 (5) 9 (4) 1 (1) 1 (2) 13 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>0 0 1 (2) 0 1 (2) 2 (1) 0 0 2 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychosis secondary to methamphetamine use</td>
<td>0 1 (2) 0 0 0 1 (&lt;1) 0 0 2 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coronary-artery disorder</td>
<td>0 0 0 0 0 0 0 0 1 (&lt;1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse events of special interest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infections — no. of patients (%)†</td>
<td>26 (39) 35 (56) 28 (44) 27 (43) 19 (31) 109 (43) 26 (39) 11 (22) 142 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer — no. of patients (%)†</td>
<td>1 (1) 1 (2) 0 0 1 (2) 2 (1) 1 (1) 1 (2) 2 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutaneous¶</td>
<td>0 1 (2) 0 0 0 1 (&lt;1) 0 0 1 (&lt;1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncutaneous‖</td>
<td>NA 4/94 (4) 3/89 (3) 6/270 (2) 4/253 (2) 17/706 (2) NA 0/50 17/707 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injection-site reactions — no./total no. (%)**</td>
<td>3/303 (1) 6/203 (3) 7/220 (3) 0/43 1/50 (2) 14/516 (3) 3/303 (1) 0/2 14/742 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>Interleukin-12/23 monoclonal antibody</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3/303 (1) 6/203 (3) 7/220 (3) 0/43 1/50 (2) 14/516 (3) 3/303 (1) 0/2 14/742 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NA 4/94 (4) 3/89 (3) 6/270 (2) 4/253 (2) 17/706 (2) NA 0/50 17/707 (2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Safety data include only those patients who received at least one injection of study agent and were summarized according to the actual treatment received. One patient did not undergo randomization but was treated with placebo at week 0, and two patients were never treated.

† Significant differences were not observed in the rates of patients who had at least one adverse event (P = 0.19), adverse events leading to withdrawal of the study agent (P = 1.00), serious adverse events (P = 0.51), or malignant conditions (P = 1.00), although the study was not designed or powered to detect small differences in rates of adverse events. All statistical comparisons for common adverse events resulted in P values exceeding 0.05, with the exception of that for cases of aggravated psoriasis (P = 0.04) and aggravated arthritis (P = 0.005), both of which occurred more commonly in the placebo group.

‡ Common adverse events were those occurring in at least 5% of patients in any treatment group.

¶ A serious adverse event was defined as any adverse event that resulted in any of the following outcomes: death, a life-threatening condition, in-patient hospitalization or prolongation of existing hospitalization, persistent or significant disability or incapacity, or a congenital anomaly or birth defect, regardless of its relationship to the study agent.\(^2\)

‖ Cutaneous cancers include squamous-cell and basal-cell carcinomas.

\(^{11}\) Noncutaneous cancer is prostate cancer.

**There were a total of 819 injections of placebo and 706 injections of interleukin-12/23 monoclonal antibody. Maintaining the double-blind aspect of the study required that patients randomly assigned to treatment with the interleukin-12/23 monoclonal antibody receive placebo at dosing visits when they were not scheduled to receive the study agent. NA denotes not applicable.
Patient-reported outcome measures improved significantly in patients treated with the interleukin-12/23 monoclonal antibody. Significantly more patients treated with the interleukin-12/23 monoclonal antibody than with placebo reported improvement of their overall skin condition at week 12, and this improvement was accompanied by quality-of-life improvements reflected in the Dermatology Life Quality Index. Many patients indicated that psoriasis had no detectable adverse effect on their quality of life after active treatment. Elevated levels of interleukin-12 have been associated with major depression. Whether quality-of-life improvements resulted from a reduction in the levels of interleukin-12, independent of improvements in overall skin condition, warrants further evaluation.

The proportions of patients who had adverse events and serious adverse events were higher among patients treated with the interleukin-12/23 monoclonal antibody than among those treated with placebo; the differences were not statistically significant, but the trial was not large enough to detect differences in uncommon serious adverse events. Rates of injection-site reactions were similar between patients treated with the interleukin-12/23 monoclonal antibody and those who received placebo. Development of antibodies to the interleukin-12/23 monoclonal antibody occurred in 4% of patients. The relationship between treatment with the interleukin-12/23 monoclonal antibody and glucose abnormalities in nonfasting patients, cancer, and the serious adverse events observed (including serious infections and myocardial infarctions) requires further study. Neither tuberculosis nor opportunistic infections developed in any patients, although both have been reported in persons congenitally deficient in interleukin-12 p40 or interleukin-12 receptor β1.

The moderate size of the study precludes meaningful assessment of the effect that blocking interleukin-12 and interleukin-23 can have on the risk of cancer. Preclinical models suggest that the interleukin-12/23 p40 cytokines play a role in tumor immunity, though it remains unclear whether blocking the interleukin-12 and interleukin-23 cytokines will affect the risk of cancer in humans or whether any effects would be detrimental or potentially protective.

Our findings support previous studies showing an immune basis for psoriasis. The high proportion of patients responding to the interleukin-12/23 monoclonal antibody and the high level of response of the patients implicate the p40 subunit, shared by the interleukin-12 and interleukin-23 cytokines, as a key mediator of psoriasis. Produced by antigen-presenting cells, such as Langerhans’ cells in the skin, interleukin-12 activates CD4+ T cells and natural killer cells to induce expression of type 1 cytokines, such as interferon-γ and tumor necrosis factor α, which have been shown to be important in the pathophysiology of psoriasis. Interleukin-23 activates a distinct T-cell lineage that expresses interleukin-17 and increases keratinocyte expression of inducible nitric oxide synthase, which has also been implicated in the pathophysiology of psoriasis. The relative contributions of interleukin-12 and interleukin-23 to the pathophysiology of psoriasis are incompletely understood, though recent evidence supports a prominent role of interleukin-23.

These results indicate that the interleukin-12/23 p40 cytokines may be an important new therapeutic target for patients with psoriasis. Although we recognize the limitations of comparisons across studies, the high level of efficacy observed in this study compares favorably with the efficacy reported for currently available intramuscularly or subcutaneously administered biologic therapeutics for psoriasis. Administration of the interleukin-12/23 monoclonal antibody led to pharmacodynamic effects that were sustained for many weeks and provided significant and prolonged efficacy after only one dose or four weekly doses. This trial was not designed to evaluate the efficacy and safety of long-term use. Additional studies are needed to characterize the safety and efficacy of the interleukin-12/23 monoclonal antibody in patients with psoriasis and to define the dose schedule that will safely maintain the high level of response.

Supported by Centocor, Malvern, PA.

Dr. Krueger reports receiving fees as a consultant or advisory board member for Abbott, Almirall, Alza, Amgen, Astellas, Boehringer Ingelheim, Barrier Therapeutics, Bristol-Myers Squibb, Centocor, Connetics, and Genentech; Dr. Langley, for Centocor, Abbott, and Amgen/Wyeth; Dr. Leonardi, for Abbott, Amgen, Centocor, and Genentech; and Dr. Lebwohl, for Abbott, Amgen, Astellas, Centocor, Connetics, Galdema, Genentech, Novartis, PharmaDerm, and Warner Chilcott. Dr. Krueger reports receiving lecture fees from Abbott, Amgen, Boehringer Ingelheim, Centocor, and Connetics; Dr. Langley, from Abbott and Amgen/Wyeth; Dr. Leonardi, from Abbott, Amgen, Centocor, and Genentech; and Dr. Lebwohl, from Abbott, Astellas, Amgen, Centocor, Connetics, Galdema, Genentech, PharmaDerm, and Warner Chilcott. Dr. Krueger reports receiving stipends for a clinical research fellowship from Amgen and Centocor; Dr. Langley, grant support from Centocor, Abbott, and Amgen/Wyeth;
APPENDIX

Investigators in the CNTO 1275 Psoriasis Study Group include the following: P. Almeyer, St. Josef Hospital, Bochum, Germany; R. Bissoinnette, Innovadwer Research, Montreal; R. Blum and M. Lebwohl, Mt. Sinai School of Medicine, New York; M. Bourcier, Dermatology Clinic, Moncont, NB, Canada; W. Carey, Royal Victoria Hospital, Montreal; M. Chambers, Radiant Research, Columbus, OH; D. Fiorentino (replaced A. Kimball), Stanford University School of Medicine, Stanford, CA; S. Fretzin, Dawes Fretzin Clinical Research Group, Indianapolis; D. Gratton, International Dermatology Research Centre, London, ON, Canada; R. Hamlin, Associates in Research, Fresno, CA; D. Harvey, Jacksonville Center for Clinical Research, Jacksonville, FL; C. Hudson, Research Solutions, Evansville, IN; R. Kaufmann, Johann Wolfgang Goethe Universität, Frankfurt, Germany; L. Kirikc, Physicians Skin Care, Louisville, KY; R. Kirsner, Veterans Affairs Medical Center, Miami; N. Korman, University Hospitals of Cleveland, Cleveland; G. Krueger, University of Utah, Salt Lake City; P. Krusinski and M. Zavod, Fletcher Allen Health Care, Burlington, VT; J. Lambert, Universititsziekenhuis Antwerpen Dermatologie, Edegem, Belgium; R. Langle, Eastern Canada Cutaneous Research Associates, Halifax, NS, Canada; C. Leonard, Central Dermatology, St. Louis; M. Ling, Medaphase, Newan, GA; B. Lubin, Hampton Roads Center for Clinical Research, Norfolk, VA; T. Luger, Universität Münster, Münster, Germany; L. Marot, Université Catholique de Louvain, Hospital Saint-Luc, Brussels; R. Matheson, Oregon Medical Research Center, Portland; A. Menter, Texas Dermatology Research Institute, Dallas; A. Nagay, Sneeze, Whitehouse, NJ; M. Nghiem, et al., Clinical Research Center, Central New Jersey, East Windsor; T. Nigra, Dermatologie Associates, Windsor, DC; B. Packman (replaced A. Mangione), Radiant Research, Philadelphia; K. Papp, Proibty Medical Research, Waterloo, ON, Canada; Y. Poulin, Centre Dermatologique du Québec Metropolitan, Sainte-Foy, QC, Canada; K. Reich, Geh-August-Universität, Göttingen, Germany; P. Rich, Northwest Cutaneous Research Specialists, Portland, OR; L. Rosoph, Proibty Medical Research, North Bay, ON, Canada; J. Sobell, ORA Clinical Research and Development, North Andover, MA; J. Swinehart, Colorado Medical Research Center, Denver; Z. Tomi, New Lab Clinical Research, St. John’s, NL, Canada; D. Toth, Proibty Medical Research, Windsor, ON, Canada; A. Truett III, Pedia Research, Owensboro, KY; E. Tschen, Academic Dermatology Associates, Albuquerque, NM; N. Wesl, Edmonton, AB, Canada; and M. Zanolli, Dermatology Consultants, Nashville.

REFERENCES

23. 21.C.E.R. § 312.32 (revised as of April 1, 2006).

INTERLEUKIN-12/23 MONOClonAL ANTIBODY FOR THE treatment OF PSORIASIS

Dr. Leonard, educational grants from Amgen and Genentech; and Dr. Lebwohl, grants from Abbott, Amgen, Astellas, Centocor, Connetics, Galderma, Genentech, PharmDerm, and Warner Chilcott. Drs. Yawalkar, Karlen, Hunger, and Dooley report being employees of Centocor. Dr. Krueger reports owning stock options from ZARS Pharma; Drs. Yawalkar, Guzzo, and Dooley report holding stock and stock options in Johnson & Johnson; and Dr. Wang reports being a stockholder in Johnson & Johnson. No other potential conflict of interest relevant to this article was reported.

We thank Cynthia Arnold and Mary Whitman for their editorial assistance; Daniel Hofstaedter, Eric Zhang, and Stephen Xu for their programming assistance; and Jacqueline Benson for her critical review of the manuscript.

Interleukin-12/23 Monoclonal Antibody for the Treatment of Psoriasis


Religion, Conscience, and Controversial Clinical Practices

Farr A. Curlin, M.D., Ryan E. Lawrence, M.Div., Marshall H. Chin, M.D., M.P.H., and John D. Lantos, M.D.

ABSTRACT

BACKGROUND
There is a heated debate about whether health professionals may refuse to provide treatments to which they object on moral grounds. It is important to understand how physicians think about their ethical rights and obligations when such conflicts emerge in clinical practice.

METHODS
We conducted a cross-sectional survey of a stratified, random sample of 2000 practicing U.S. physicians from all specialties by mail. The primary criterion variables were physicians’ judgments about their ethical rights and obligations when patients request a legal medical procedure to which the physician objects for religious or moral reasons. These procedures included administering terminal sedation in dying patients, providing abortion for failed contraception, and prescribing birth control to adolescents without parental approval.

RESULTS
A total of 1144 of 1820 physicians (63%) responded to our survey. On the basis of our results, we estimate that most physicians believe that it is ethically permissible for doctors to explain their moral objections to patients (63%). Most also believe that physicians are obligated to present all options (86%) and to refer the patient to another clinician who does not object to the requested procedure (71%). Physicians who were male, those who were religious, and those who had personal objections to morally controversial clinical practices were less likely to report that doctors must disclose information about or refer patients for medical procedures to which the physician objected on moral grounds (multivariate odds ratios, 0.3 to 0.5).

CONCLUSIONS
Many physicians do not consider themselves obligated to disclose information about or refer patients for legal but morally controversial medical procedures. Patients who want information about and access to such procedures may need to inquire proactively to determine whether their physicians would accommodate such requests.
Recent controversies regarding physicians and pharmacists who refuse to prescribe or dispense emergency and other contraceptives have sparked a debate about conscientious objection in health care.\textsuperscript{1-5} On the one hand, most people believe that health professionals should not have to engage in medical practices about which they have moral qualms. On the other hand, most people also believe that patients should have access to legal treatments, even in situations in which their physicians are troubled about the moral implications of those treatments.\textsuperscript{6} Such situations raise a number of questions about the balance of rights and obligations within the doctor–patient relationship. Is it ethical for physicians to describe their objections to patients? Should physicians have the right to refuse to discuss, provide, or refer patients for medical interventions to which they have moral objections?

The medical profession appears to be divided on this issue. Historically, doctors and nurses have not been required to participate in abortions or assist patients in suicide, even where those interventions are legally sanctioned. In recent years, several states have passed laws that shield physicians and other health care providers from adverse consequences for refusing to participate in medical services that would violate their consciences.\textsuperscript{7} For example, the Illinois Health Care Right of Conscience Act protects a health care provider from all liability or discrimination that might result as a consequence of “his or her refusal to perform, assist, counsel, suggest, recommend, refer or participate in any way in any particular form of health care service which is contrary to the conscience of such physician or health care personnel.”\textsuperscript{8} In the wake of recent controversies over emergency contraception, editorials in leading clinical journals have criticized these “conscience clauses” and challenged the idea that physicians may deny legally and medically permitted medical interventions, particularly if their objections are personal and religious. Charo, for example, suggests that the conflict about conscience clauses “represents the latest struggle with regard to religion in America,” and she criticizes those medical professionals who would claim “an unfettered right to personal autonomy while holding monopolistic control over a public good.”\textsuperscript{9} Sauvulescu takes a stronger stance, arguing that “a doctor’s conscience has little place in the delivery of modern medical care” and that “if people are not prepared to offer legally permitted, efficient, and beneficial care to a patient because it conflicts with their values, they should not be doctors.”\textsuperscript{10}

In spite of such debates, there have been few empirical studies of how physicians think about their responsibilities when their own moral convictions conflict with their patients’ requests for legal medical procedures. We examined data from a national survey of U.S. physicians to determine what practicing physicians think their obligations are when a patient requests a legal medical procedure to which the physician has a religious or other moral objection. We quantify the percentage of physicians who might refrain from presenting all treatment options to patients or refuse to refer them to an accommodating provider, and we examine whether particular subgroups of physicians are more likely to do so. We then discuss the implications for ongoing debates concerning the ethics of the doctor–patient relationship.

**Methods**

This study’s methods have been described in detail elsewhere.\textsuperscript{10,11} In 2003, we mailed a confidential, self-administered, 12-page questionnaire (see the Supplementary Appendix, available with the full text of this article at www.nejm.org) to a random sample of 2000 practicing U.S. physicians 65 years of age or younger. The sample was stratified according to specialty. These physicians were chosen from the American Medical Association Physician Masterfile — a database intended to include all physicians in the United States. We included modest oversamples of psychiatrists and physicians who work in several other subspecialties that deal particularly with death and severe suffering, in order to enhance the power of analyses that are not central to this article. Physicians received up to three separate mailings of the questionnaire, and the third mailing offered $20 for participation. The study was approved by the institutional review board of the University of Chicago.

**Questionnaire**

The primary criterion variables were physicians’ responses to the following three questions: “If a patient requests a legal medical procedure, but the patient’s physician objects to the procedure for religious or moral reasons, would it be ethical for the physician to plainly describe to the patient why he or she objects to the requested procedure? Does the physician have an obligation to present all possible options to the patient, including infor-
mation about obtaining the requested procedure? Does the physician have an obligation to refer the patient to someone who does not object to the requested procedure?” Response categories were yes, no, and undecided.

We also assessed physicians’ intrinsic religiosity and religious affiliations. Intrinsic religiosity — the extent to which a person embraces his or her religion as the “master motive” that guides and gives meaning to his or her life — was measured on the basis of agreement or disagreement with two statements: “I try hard to carry my religious beliefs over into all my other dealings in life” and “My whole approach to life is based on my religion.” Both statements are derived from Hoge’s Intrinsic Religious Motivation Scale and have been validated extensively in previous research. Intrinsic religiosity was categorized as being low if physicians disagreed with both statements, moderate if they agreed with one but not the other, and high if they agreed with both.

The religious affiliations of the physicians in the survey were categorized as none (a category that included atheist, agnostic, and none), Protestant, Catholic, Jewish, or other (a category that included Buddhist, Hindu, Mormon, Muslim, Eastern Orthodox, and other). Organizational or participatory religiosity was measured according to the frequency of attendance at religious services (never, once a month or less, or twice a month or more).

To determine whether physicians’ judgments about their ethical obligations are associated with their views on controversial clinical practices, we asked the survey respondents whether they have a religious or moral objection to terminal sedation (administering sedation that leads to unconsciousness in dying patients), abortion for failed contraception, and the prescription of birth control to adolescents without parental approval. Secondary predictors were the demographic characteristics (age, sex, race or ethnic group, and region) of the physicians surveyed and whether they worked in an academic health center or a religiously oriented or faith-based institution. The primary medical specialty was included as a control variable in the multivariate analyses.

**Statistical Analysis**

Weights were assigned and included in the analyses to account for the sampling strategy and the modest differences in response rates according to the respondents’ sex and whether they had graduated from a U.S. or foreign medical school. We first generated overall population estimates for agreement with each of the criterion measures. We then used a Mantel–Haenszel test for trend with one degree of freedom (for ordinal predictors) and the chi-square test (for nonordinal predictors) to examine the associations between each predictor and each criterion measure. Finally, we used multivariate logistic regression to examine whether associations persisted after controlling for other covariates. All reported P values are two-sided and have not been adjusted for multiple statistical testing. All analyses were conducted with Stata SE statistical software (version 9.0).

**Results**

Of the 2000 potential respondents, an estimated 9% could not be contacted because their addresses were incorrect or they had died (see the Supplementary Appendix). Among physicians who could be contacted, the response rate was 63% (1144 of 1820). Graduates of foreign medical schools were less likely to respond than graduates of U.S. medical schools (54% vs. 65%, P<0.001), and men were less likely to respond than women (61% vs. 67%, P=0.03). These differences were accounted for by assigning case weights. The response rates did not differ significantly according to age, region, or board certification. The characteristics of the respondents are listed in Table 1.

On the basis of these results, we estimated that when a patient requests a legal medical procedure to which the doctor objects for religious or moral reasons, most physicians believe it is ethically permissible for the doctor to describe that objection to the patient (63%) and that the doctor is obligated to present all options (86%) and to refer the patient to someone who does not object to the requested procedure (71%) (Table 2).

Physicians who were more religious (as measured by either their attendance at religious services or their intrinsic religiosity) were more likely to report that doctors may describe their objections to patients, and they were less likely to report that physicians must present all options and refer patients to someone who does not object to the requested procedure (Table 3). As compared with those with no religious affiliation, Catholics and Protestants were more likely to report that physicians may describe their religious or moral objections and less likely to report that physicians are obligated to refer patients to
someone who does not object to the requested procedure.

Physicians who objected to abortion for failed contraception and prescription of birth control for adolescents without parental consent were more likely than those who did not oppose these practices to report that doctors may describe their objections to patients (P<0.001 for both comparisons); the association for the objection to terminal sedation was not significant (P=0.11) (Table 4). Physicians who objected to the three controversial medical practices were less likely to report that doctors must present all options and refer patients to other providers (P<0.001 for all comparisons). The associations for religious characteristics and objections to controversial clinical practices persisted after controlling for age, sex, ethnic group, region, and specialty.

After adjustment for religious characteristics and other covariates, region, race or ethnic group, practice in an academic medical center, and practice in a religiously oriented health center were not significantly associated with any of the criterion variables. However, with increasing age, physicians were more likely to report that doctors may describe their objections to patients (odds ratio for each additional year of age, 1.02; 95% confidence interval [CI], 1.00 to 1.04). Men were more likely

---

**Table 1. Characteristics of the 1144 Survey Respondents and Objections to Controversial Clinical Practices.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No./Total No. (%)</th>
<th>Characteristic</th>
<th>No./Total No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female sex</strong></td>
<td>300/1142 (26)</td>
<td><strong>Religious characteristics</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Race or ethnic group†</strong></td>
<td></td>
<td><strong>Intrinsic religiosity</strong></td>
<td></td>
</tr>
<tr>
<td>White, non-Hispanic</td>
<td>869/1121 (78)</td>
<td>Low</td>
<td>407/1098 (37)</td>
</tr>
<tr>
<td>Asian</td>
<td>138/1121 (12)</td>
<td>Moderate</td>
<td>292/1098 (27)</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>57/1121 (5)</td>
<td>High</td>
<td>399/1098 (36)</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>26/1121 (2)</td>
<td>Attendance at religious services</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>31/1121 (3)</td>
<td>Never</td>
<td>114/1128 (10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Once a month or more</td>
<td>499/1128 (44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Twice a month or more</td>
<td>515/1128 (46)</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td><strong>Religious affiliation</strong></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>386/1142 (34)</td>
<td>Protestant</td>
<td>428/1127 (38)</td>
</tr>
<tr>
<td>Midwest</td>
<td>276/1142 (24)</td>
<td>Catholic</td>
<td>244/1127 (22)</td>
</tr>
<tr>
<td>Northeast</td>
<td>264/1142 (23)</td>
<td>Jewish</td>
<td>181/1127 (16)</td>
</tr>
<tr>
<td>West</td>
<td>216/1142 (19)</td>
<td>None</td>
<td>117/1127 (10)</td>
</tr>
<tr>
<td>Practice in academic medical center</td>
<td>353/1115 (32)</td>
<td>Other</td>
<td>157/1127 (14)</td>
</tr>
<tr>
<td>Practice in religiously oriented center</td>
<td>138/1111 (12)</td>
<td><strong>Opinions about controversial clinical practices</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Primary specialty</strong></td>
<td></td>
<td>Terminal sedation</td>
<td></td>
</tr>
<tr>
<td>Medical and subspecialties</td>
<td>231/1142 (20)</td>
<td>Do not object</td>
<td>915/1097 (83)</td>
</tr>
<tr>
<td>Family practice</td>
<td>158/1142 (14)</td>
<td>Object</td>
<td>182/1097 (17)</td>
</tr>
<tr>
<td>Pediatrics and subspecialties</td>
<td>147/1142 (13)</td>
<td>Abortion due to failed contraception</td>
<td></td>
</tr>
<tr>
<td>General internal medicine</td>
<td>129/1142 (11)</td>
<td>Do not object</td>
<td>527/1091 (48)</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>100/1142 (9)</td>
<td>Object</td>
<td>564/1091 (52)</td>
</tr>
<tr>
<td>Surgery and subspecialties</td>
<td>100/1142 (9)</td>
<td>Prescription of birth control without parental consent</td>
<td></td>
</tr>
<tr>
<td>Obstetrics and gynecology</td>
<td>80/1142 (7)</td>
<td>Do not object</td>
<td>647/1108 (58)</td>
</tr>
<tr>
<td>Other</td>
<td>197/1142 (17)</td>
<td>Object</td>
<td>461/1108 (42)</td>
</tr>
</tbody>
</table>

* Numbers do not all sum to 1144 because not all respondents answered all the questions. The mean (±SD) age of respondents was 49.0±8.3 years.
† Race and ethnic group were reported by patients on the survey.
than women to report that physicians may describe their objections (odds ratio, 1.8; 95% CI, 1.3 to 2.5) and less likely to report that physicians are obligated to present all options (odds ratio, 0.5; 95% CI, 0.3 to 0.9) and refer patients to an accommodating provider (odds ratio, 0.5; 95% CI, 0.3 to 0.7).

**DISCUSSION**

Most of the physicians in our survey reported that when a patient requests a legal medical intervention to which the physician objects for religious or moral reasons, it is ethically permissible for the physician to describe the reason for the objection but that the physician must also disclose information about the intervention and refer the patient to someone who will provide it. However, the number of physicians who disagreed with or were undecided about these majority opinions was not trivial. If physicians’ ideas translate into their practices, then 14% of patients — more than 40 million Americans — may be cared for by physicians who do not believe they are obligated to disclose information about medically available treatments they consider objectionable. In addition, 29% of patients — or nearly 100 million Americans — may be cared for by physicians who do not believe they have an obligation to refer the patient to another provider for such treatments. The proportion of physicians who object to certain treatments is substantial. For example, 52% of the physicians in this study reported objections to abortion for failed contraception, and 42% reported objections to contraception for adolescents without parental consent. The findings of this study may be important primarily for patients. They should know that many physicians do not believe they are obligated to disclose information about or provide referrals for legal yet controversial treatments that they consider objectionable. In addition, 29% of patients — or nearly 100 million Americans — may be cared for by physicians who do not believe they have an obligation to refer the patient to another provider for such treatments.

The findings of this study may be important primarily for patients. They should know that many physicians do not believe they are obligated to disclose information about or provide referrals for legal yet controversial treatments. Patients who want full disclosure from their own physicians might inform themselves of possible medical interventions — a task that is not always easy — and might proactively question their physicians about these matters. Patients may not have ready access to information about physicians’ religious characteristics and moral convictions. Thus, if patients are concerned about certain interventions for sexual and reproductive health and end-of-life care, they should ask their doctors ahead of time whether they will discuss such options.

If a patient wants a treatment that the physician will not provide, the patient may choose to consult a different physician.

Physicians’ judgments about their obligations are significantly associated with their own religious characteristics, sex, and beliefs about morally controversial clinical practices. Female physicians are more supportive of full disclosure and referral than are male physicians, perhaps because many controversial issues in medicine (e.g., abortion, contraception, and assisted reproductive technologies) disproportionately involve the sexual and reproductive health of women. Religious physicians are less likely to endorse full disclosure and referral than are nonreligious physicians, perhaps because, as many previous studies have shown, religious physicians are more likely to have personal objections to many controversial medical interventions. Thus, those physicians who are most likely to be asked to act against their consciences are the ones who are most likely to say that physicians should not have to do so.

These conflicts might be understood in the context of perennial debates about medical paternalism and patient autonomy. Strong forms of

---

**Table 2. Opinions about the Ethical Obligations of a Physician Who Objects to a Legal Medical Procedure Requested by a Patient.**

<table>
<thead>
<tr>
<th>Question and Response</th>
<th>No. (%)&lt;sup&gt;*&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would it be ethical for the physician to plainly describe to the patient why he or she objects to the requested procedure?</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>715 (63)</td>
</tr>
<tr>
<td>Undecided</td>
<td>168 (15)</td>
</tr>
<tr>
<td>No</td>
<td>244 (22)</td>
</tr>
<tr>
<td>Does the physician have an obligation to present all possible options to the patient, including information about obtaining the requested procedure?</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>981 (86)</td>
</tr>
<tr>
<td>Undecided</td>
<td>61 (6)</td>
</tr>
<tr>
<td>No</td>
<td>86 (8)</td>
</tr>
<tr>
<td>Does the physician have an obligation to refer the patient to someone who does not object to the requested procedure?</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>820 (71)</td>
</tr>
<tr>
<td>Undecided</td>
<td>114 (11)</td>
</tr>
<tr>
<td>No</td>
<td>194 (18)</td>
</tr>
</tbody>
</table>

<sup>*</sup> Population estimates account for the survey design. Percentages reflect weighted results.
Table 3. Opinions about Physicians’ Ethical Obligations According to the Religious Characteristics of the Respondents.*

<table>
<thead>
<tr>
<th>Religious Characteristic</th>
<th>No. of Respondents (N=1144)</th>
<th>Physicians May Describe Their Moral Objections</th>
<th>Physicians Are Obligated to Disclose All Possible Options</th>
<th>Physicians Are Obligated to Refer the Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% P Value Multivariate Odds Ratio (95% CI)</td>
<td>% P Value Multivariate Odds Ratio (95% CI)</td>
<td>% P Value Multivariate Odds Ratio (95% CI)</td>
</tr>
<tr>
<td>Intrinsic religiosity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low†</td>
<td>405</td>
<td>56</td>
<td>1.0</td>
<td>92</td>
</tr>
<tr>
<td>Moderate</td>
<td>290</td>
<td>62</td>
<td>1.4 (1.0–2.0)</td>
<td>84</td>
</tr>
<tr>
<td>High</td>
<td>397</td>
<td>73</td>
<td>2.5 (1.7–3.5)</td>
<td>81</td>
</tr>
<tr>
<td>Attendance at religious services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never†</td>
<td>111</td>
<td>51</td>
<td>1.0</td>
<td>94</td>
</tr>
<tr>
<td>Once a month or less</td>
<td>496</td>
<td>59</td>
<td>1.5 (0.9–2.4)</td>
<td>89</td>
</tr>
<tr>
<td>Twice a month or more</td>
<td>513</td>
<td>71</td>
<td>2.7 (1.6–4.3)</td>
<td>82</td>
</tr>
<tr>
<td>Religious affiliation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>427</td>
<td>70</td>
<td>2.3 (1.4–3.8)</td>
<td>86</td>
</tr>
<tr>
<td>Catholic</td>
<td>243</td>
<td>63</td>
<td>1.8 (1.1–3.0)</td>
<td>79</td>
</tr>
<tr>
<td>Jewish</td>
<td>179</td>
<td>56</td>
<td>1.1 (0.6–1.9)</td>
<td>93</td>
</tr>
<tr>
<td>None†</td>
<td>116</td>
<td>52</td>
<td>1.0</td>
<td>92</td>
</tr>
<tr>
<td>Other</td>
<td>153</td>
<td>63</td>
<td>1.5 (0.8–2.7)</td>
<td>89</td>
</tr>
</tbody>
</table>

* Population estimates account for the survey design. Percentages reflect weighted results.
† This was the reference category.
paternalism are based on the assumption that physicians know what is best for their patients and may therefore make decisions without informing their patients of all the facts, alternatives, or risks. Paternalism is widely criticized for violating the right of adults to self-determination. The inverse of strong paternalism is a strict emphasis on patient autonomy, which suggests that physicians must simply disclose all options and allow patients to choose among them. Models that emphasize patient autonomy to such an extent have been criticized for diminishing the moral agency and responsibility of physicians by making them mere technicians or vendors of health care goods and services.2,19-23

This study suggests that the balance that most physicians strike between paternalism and autonomy involves both full disclosure and an open dialogue about the options at hand. This balance resembles the interactive models proposed by Emanuel and Emanuel,19 Quill and Brody,20 Siegler,23 and Thomasma.21 These ethicists have all recommended models for the doctor–patient relationship that retain the moral agency of both the physician and the patient by encouraging them to engage in a dialogue and negotiate mutually acceptable accommodations that do not require either of the parties to violate their own convictions. In Emanuel and Emanuel’s terms, these interactive models retain a role for the influence of “the physician's values, the physician's understanding of the patient's values, [and] his or her judgment of the worth of the patient's values.”19 Although these models require physicians to disclose all information relevant to patients’ decisions, they do not require physicians to be value-neutral. Rather, they allow physicians to explain the reasons for their objections to the requested procedures.

The lack of consensus among physicians about whether referrals to other providers who will offer a controversial treatment should be required mirrors the ambivalence about this point within the field of bioethics. Childress and Siegler22 say that physicians “may” have a duty to inform patients about other physicians who would provide what the patient requests, and Quill and Brody20 comment that physicians are “perhaps” obligated to facilitate the transfer of care. This ambivalence stems from a long-standing concern that physicians not be asked to act in ways that “would violate [their] personal sense of responsible con-

<table>
<thead>
<tr>
<th>Respondents (N=1,144)</th>
<th>View on Controversial Clinical Practice</th>
<th>Physicians Are Obligated to Disclose All Possible Options</th>
<th>Physicians Are Obligated to Describe Their Moral Objections</th>
<th>Physicians Are Obligated to Refer the Patient</th>
<th>Multivariate Odds Ratio (95% CI)</th>
<th>% P Value</th>
<th>Multivariate Odds Ratio (95% CI)</th>
<th>% P Value</th>
<th>Multivariate Odds Ratio (95% CI)</th>
<th>% P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminal sedation</td>
<td>Do not object†</td>
<td>911</td>
<td>62</td>
<td>1.0</td>
<td>78</td>
<td>1.0</td>
<td>0.001</td>
<td>95% CI</td>
<td>0.001</td>
<td>0.4 (0.2–0.6)</td>
</tr>
<tr>
<td>Object</td>
<td>182</td>
<td>69</td>
<td>1.4 (0.9–2.0)</td>
<td>58</td>
<td>0.5 (0.3–0.7)</td>
<td>1.4 (0.9–2.0)</td>
<td>58</td>
<td>0.5 (0.3–0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abortion for failed contraception</td>
<td>Do not object†</td>
<td>524</td>
<td>55</td>
<td>1.0</td>
<td>58</td>
<td>0.4 (0.2–0.6)</td>
<td>1.0</td>
<td>0.001</td>
<td>0.3 (0.3–0.5)</td>
<td></td>
</tr>
<tr>
<td>Object</td>
<td>562</td>
<td>70</td>
<td>2.0 (1.5–2.7)</td>
<td>78</td>
<td>0.3 (0.2–0.5)</td>
<td>2.0 (1.5–2.7)</td>
<td>78</td>
<td>0.3 (0.2–0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prescription of birth control to adolescents without parental consent</td>
<td>Do not object†</td>
<td>646</td>
<td>58</td>
<td>1.0</td>
<td>78</td>
<td>1.0</td>
<td>0.001</td>
<td>0.3 (0.2–0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object</td>
<td>479</td>
<td>72</td>
<td>1.6 (1.2–2.2)</td>
<td>83</td>
<td>0.3 (0.2–0.5)</td>
<td>1.6 (1.2–2.2)</td>
<td>83</td>
<td>0.3 (0.2–0.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Population estimates account for the survey design. Percentages reflect weighted results.†This was the reference category.
Unfortunately, at times the only accommodation that is acceptable to both the patient and the physician may be termination of the clinical relationship.\textsuperscript{10,20,22,23}

Our study has several important limitations. Although we did not find substantial evidence of a response bias,\textsuperscript{10,11} unmeasured characteristics may have systematically affected physicians’ willingness to respond in ways that bias our results. In addition, physicians in different specialties face different arrays of morally controversial practices. Because this study included physicians from all specialties, many participants were asked to report moral judgments about medical practices with which they may have had little or no clinical experience. Moreover, physicians’ judgments about their general obligations do not necessarily correspond with their judgments about any particular clinical scenario, and we do not know how their judgments about their obligations translate into their actual practices. Finally, we had three criterion measures and several predictors. Therefore, although hypotheses were theoretically specified and the expected associations were consistently observed, there was the risk of an inflated type 1 error due to multiple comparisons. For all of these reasons, our findings should be considered preliminary, and future studies should use vignettes, patients’ reports, or direct observation to measure more directly the ways in which physicians respond to moral conflict in the clinical encounter.

Notwithstanding these limitations, the results of our study suggest that when patients request morally controversial clinical interventions, male physicians and those who are religious will be most likely to express personal objections and least likely to disclose information about the interventions or to refer patients to more accommodating providers. Ongoing debates about conscientious objections in medicine should take account of the complex relationships among sex, religious commitments, and physicians’ approaches to morally controversial clinical practices. In the meantime, physicians and patients might engage in a respectful dialogue to anticipate areas of moral disagreement and to negotiate acceptable accommodations before crises develop.

Supported by grants from the Greenwall Foundation and the Robert Wood Johnson Clinical Scholars Program (to Drs. Curlin, Chin, and Lantos) and the National Center for Complementary and Alternative Medicine (1 K23 AT002749, to Dr. Curlin).

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

We thank Lainie Ross, Debra Stulberg, and Lynn Vance for their insightful comments on earlier drafts of this manuscript and Ted Karrison, Jack Iwashyna, and Anirban Basu for their assistance with the statistical analysis.

\textbf{REFERENCES}

23. Copyright © 2007 Massachusetts Medical Society.
The Incidentally Discovered Adrenal Mass

William F. Young, Jr., M.D.

The Clinical Problem

An adrenal “incidentaloma” is an adrenal mass, generally 1 cm or more in diameter, that is discovered serendipitously during a radiologic examination performed for indications other than an evaluation for adrenal disease. This definition excludes cases in which a symptomatic adrenal-dependent syndrome is “missed” because of a superficial interview or physical examination. The widespread use of abdominal ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) has resulted in the clinical dilemma of the adrenal incidentaloma. Numerous autopsy studies have examined the frequency of incidental adrenal nodules. In a report on 25 studies, the overall frequency of adrenal adenomas in 87,065 autopsies was 6% (range, 1 to 32). Abdominal CT yields similar findings; a recent study reported a prevalence of adrenal incidentaloma of 4%. The prevalence of adrenal adenomas increases with increasing age: the probability of finding an unsuspected adrenal adenoma on abdominal CT in a patient between 20 and 29 years of age would be approximately 0.2%, as compared with approximately 7% in a patient over 70 years of age.

The majority of adrenal incidentalomas are clinically nonhypersecreting, benign adenocortical adenomas. Other frequently reported diagnoses include cortisol-secreting adenocortical adenoma, pheochromocytoma, adenocortical carcinoma, and metastatic carcinoma.

Strategies and Evidence

The optimal diagnostic approach to a patient who has an adrenal incidentaloma has not been established. However, it is reasonable to start by taking a careful history and performing a physical examination, focusing on the signs and symptoms suggestive of adrenal hyperfunction or malignant disease (Table 1) and hormonal testing (Table 2). Although no specific diagnostic approach has been prospectively validated, an algorithm based on clinical experience and data regarding laboratory and imaging studies is shown in Figure 1.

From the Division of Endocrinology, Diabetes, Metabolism, Nutrition, and Internal Medicine, Mayo Clinic, Rochester, MN. Address reprint requests to Dr. Young at the Division of Endocrinology, Diabetes, Metabolism, Nutrition, and Internal Medicine, Mayo Clinic, 200 First St. SW, Rochester, MN 55905, or at young.william@mayo.edu.

Copyright © 2007 Massachusetts Medical Society.
Table 1. Symptoms and Signs Suggestive of Adrenal Hyperfunction or Malignant Disease.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Symptoms</th>
<th>Signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cushing’s syndrome</td>
<td>Patient may be asymptomatic if disease is subclinical; symptoms may include weight gain with central obesity, facial rounding and plethora, supraclavicular and dorsocervical fat pads, easy bruising, thin skin, poor wound healing, purple striae, proximal muscle weakness, emotional and cognitive changes (e.g., irritability, spontaneous tearfulness, depression, and restlessness), opportunistic and fungal infections, altered reproductive function, acne, and hirsutism</td>
<td>Hypertension, osteopenia, osteoporosis, fasting hyperglycemia, diabetes mellitus, hypokalemia, hyperlipidemia, and leukocytosis with relative lymphopenia</td>
</tr>
<tr>
<td>Pheochromocytoma</td>
<td>Patient may be asymptomatic; episodic symptoms may occur in spells (paroxysms) that can be extremely variable in presentation but typically include forceful heartbeat, pallor, tremor, headache, and diaphoresis; spells may be either spontaneous or precipitated by postural change, anxiety, medications (e.g., metoclopramide, anesthetic agents), and maneuvers that increase intraabdominal pressure (e.g., change in position, lifting, defecation, exercise, colonoscopy, pregnancy, and trauma)</td>
<td>Hypertension (paroxysmal or sustained), orthostatic hypotension, pallor, retinopathy grades 1 to 4, tremor, and fever</td>
</tr>
<tr>
<td>Primary aldosteronism</td>
<td>If hypokalemia is present, nocturia, polyuria, muscle cramps, and palpitations may be present</td>
<td>Hypertension, mild or severe; possibly hypokalemia and mild hypernatremia</td>
</tr>
<tr>
<td>Adrenocortical carcinoma</td>
<td>Symptoms may include mass effect (e.g., abdominal pain) and symptoms related to adrenal hypersecretion of cortisol (Cushing’s syndrome), androgens (hirsutism, acne, amenorrhea or oligomenorrhea, oily skin, and increased libido), estrogen (gynecomastia), or aldosterone (hypokalemia-related symptoms)</td>
<td>Hypertension, osteopenia, osteoporosis, fasting hyperglycemia, diabetes mellitus, hypokalemia, hyperlipidemia, and leukocytosis with relative lymphopenia</td>
</tr>
<tr>
<td>Metastatic cancer</td>
<td>History of an extraadrenal cancer</td>
<td>Cancer-specific signs</td>
</tr>
</tbody>
</table>

**HORMONAL EVALUATION**

**Subclinical Cushing’s Syndrome**

The term “subclinical” Cushing’s syndrome is used to refer to autonomous cortisol secretion in patients who do not have the typical signs and symptoms of hypercortisolism. Although the obvious stigmata of Cushing’s syndrome are absent, these patients may have the adverse effects of continuous, endogenous cortisol secretion, including hypertension, obesity, diabetes mellitus, and osteoporosis.9,10-13

In a report summarizing the results of 13 studies including 2005 patients with adrenal incidentalomas, autonomous cortisol secretion (independent of normal hypothalamic–pituitary control) was found in 5.3% of the patients. Since such patients do not have clinical Cushing’s syndrome and may have normal 24-hour urinary cortisol excretion, a measure of autonomous adrenocortical secretion is the best strategy for testing. Because there is no reliable way to distinguish between low-normal values and suppressed values with most commercially available corticotropin assays, adrenal autonomy is best assessed by an overnight dexamethasone (1 mg) suppression test (Table 2). Although the optimal cutoff value is debated, the use of a cortisol level greater than 5 μg per deciliter (138 nmol per liter) is standard to define abnormal values according to this test,15,16 because this level is considered to be a reasonable criterion for clinically significant glucocorticoid secretory autonomy.15 The specificity of the 1-mg overnight dexamethasone suppression test is 91%;17,18 if the result is abnormal, confirmatory testing should be performed to rule out a false positive result (Table 2).

Data from randomized trials are lacking to guide the optimal management of subclinical Cushing’s syndrome. A reasonable strategy is to consider adrenalectomy for younger patients (below the age of 40 years) and those with disorders that are potentially attributable to autonomous glucocorticoid secretion (e.g., the recent onset or worsening of underlying hypertension, diabetes mellitus, obesity, or osteoporosis). A patient with subclinical Cushing’s syndrome should receive glucocorticoid therapy perioperatively because of the risks of adrenal insufficiency, hemodynamic crisis, and death.19 The need for longer-term replacement and slow tapering of exogenous glucocorticoids should be assessed postoperatively. In limited case series, weight loss, improvement in hypertension or glycemic control or both, and the normalization of markers of bone turnover were reported after unilateral adrenalectomy in patients with subclinical Cushing’s syndrome.9,11,14
**Table 2. Laboratory Evaluation of the Patient with Adrenal Incidentaloma.**

<table>
<thead>
<tr>
<th>Possible Diagnosis</th>
<th>Screening Test</th>
<th>Causes of False Positive Results</th>
<th>Confirmatory Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subclinical Cushing’s syndrome</td>
<td>Overnight dexamethasone (1 mg) suppression test; abnormal result: serum cortisol, &gt;5 μg per deciliter (138 nmol per liter); some clinicians use a higher dose of dexamethasone (e.g., 3 mg instead of the standard 1 mg) to reduce the possibility of a false positive result without a change in sensitivity</td>
<td>Medications that accelerate hepatic metabolism of dexamethasone (e.g., antidepressants); noncompliance with dexamethasone regimen</td>
<td>Consider the following tests: serum corticotropin, cortisol in a blood specimen and 24-hr urine specimen, midnight salivary measurement of cortisol, and a formal 2-day high-dose dexamethasone suppression test (the result is considered abnormal when the cortisol level in the 24-hr urine specimen is greater than the lower limit of the normal range for the local laboratory)</td>
</tr>
<tr>
<td>Pheochromocytoma</td>
<td>Measurement of fractionated metanephrines and catecholamines in a 24-hr urinary specimen; imaging phenotype may also suggest pheochromocytoma</td>
<td>Any situation (e.g., illness requiring hospitalization) or medication (e.g., tricyclic antidepressant) that increases endogenous production of catecholamines⁷</td>
<td>Consider iodine-123 metaiodobenzylguanidine scintigraphy, MRI, subspecialty consultation, and surgery</td>
</tr>
<tr>
<td>Primary aldosteroneism</td>
<td>Morning measurement of the plasma aldosterone concentration and plasma renin activity,⁶ which can be performed while the patient is receiving any antihypertensive drug except spironolactone (Aldactone, Searle), eplerenone (Inspra, Pfizer), or high-dose amiloride (Midamor, Merck); the plasma aldosterone concentration and plasma renin activity ratio of ≥20 and a plasma aldosterone concentration of ≥15 ng per deciliter are positive results (but the cutoff for a positive result is laboratory-dependent)</td>
<td>Assay and biologic variability</td>
<td>To confirm the diagnosis of primary aldosteroneism: aldosterone suppression testing with either a saline infusion test or 24-hour urinary aldosterone excretion test while the patient maintains a high-sodium diet⁸</td>
</tr>
</tbody>
</table>

* In this test, values for the plasma aldosterone concentration are in nanograms per deciliter, and values for plasma renin activity are in nanograms per milliliter per hour.

Long-term prospective studies are needed to provide a better understanding of the natural history of subclinical Cushing’s syndrome and better guidance for decisions regarding surgical intervention.

At least two reports have suggested that cortisol secretion may be normal when the adrenal incidentaloma is discovered but may become autonomous during a subsequent period of 4 years or longer.²⁰,²¹ Until data are available from large prospective studies, these observations suggest that it is reasonable to repeat the hormonal screening annually for 4 years, as suggested by the National Institutes of Health (NIH) state-of-the-science statement.⁵

**Clinically Silent Pheochromocytoma**

Approximately 5% of adrenal incidentalomas have proved to be pheochromocytomas.³ In one study, 19 of 33 adrenal pheochromocytomas (58%) were detected initially as incidental adrenal masses, and only 10 of the 19 patients had hypertension.²² However, even clinically silent pheochromocytomas can be lethal.²³

The characteristics of an adrenal mass on imaging — the imaging phenotype — can be helpful in determining whether it is a pheochromocytoma (Table 3). Findings consistent with (although not diagnostic of) pheochromocytoma include increased attenuation on unenhanced CT, prominent vascularity of the mass (Fig. 2A), delayed washout of contrast medium, and high signal intensity on T₂-weighted MRI.²²

Because not all pheochromocytomas have this phenotype and because the expertise of radiologists and clinicians in identifying this rare neoplasm can vary, biochemical assessment is warranted for all patients. Studies reporting the characteristics of biochemical tests for pheochromocytoma are based on data from both symptomatic and asymptomatic patients. The measurement of fractionated metanephrines and
catecholamines in a 24-hour urine specimen is recommended for all patients with adrenal incidentalomas; the detection of elevated levels of fractionated metanephrines, catecholamines, or both has high sensitivity and specificity for pheochromocytoma (91 to 98% in Mayo Clinic series, for both). The additional measurement of fractionated catecholamines in the 24-hour urinary specimen increases the sensitivity of this approach by 5% and is especially helpful in diagnosing patients with dopamine-secreting neoplasms. If the suspicion of subclinical pheochromocytoma is high on the basis of the imaging phenotype but the results of 24-hour urinary studies are normal, the measurement of fractionated plasma free metanephrines (available at most reference laboratories) may be useful. Although elevated levels of fractionated plasma metanephrines have high sensitivity for pheochromocytoma (96 to 100%), the test has low specificity (85 to 89% overall and 77% in patients older than 60 years). Thus, the measurement of fractionated plasma metanephrines is recommended only when suspicion is high, to minimize the risk of false positive results that might lead to unnecessary surgery.

Primary Aldosteronism
Approximately 1% of adrenal incidentalomas have proved to be aldosterone-producing adenomas. Excessive secretion of aldosterone is associated with an increased risk of cardiovascular disease.
and other disorders, and the normalization of circulating aldosterone levels or mineralocorticoid receptor blockade is warranted in patients with excessive secretion of aldosterone.

Screening for hyperaldosteronism is routinely recommended for hypertensive patients who have an adrenal incidentaloma. Given that patients with aldosterone-producing adenomas may have normal levels of potassium in the blood, the measurement of potassium levels is not reliable in screening. A reasonable screening test is the ratio of the ambulatory morning plasma aldosterone concentration to plasma renin activity (Table 2).

### Table 3. Characteristics of Adrenal Incidentalomas on Imaging (Imaging Phenotype).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adrenocortical Adenoma</th>
<th>Adrenocortical Carcinoma</th>
<th>Pheochromocytoma</th>
<th>Metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Small, usually ≤3 cm in diameter</td>
<td>Large, usually &gt;4 cm in diameter</td>
<td>Large, usually &gt;3 cm in diameter</td>
<td>Variable, frequently &lt;3 cm</td>
</tr>
<tr>
<td>Shape</td>
<td>Round or oval, with smooth margins</td>
<td>Irregular, with unclear margins</td>
<td>Round or oval, with clear margins</td>
<td>Oval or irregular, with unclear margins</td>
</tr>
<tr>
<td>Texture</td>
<td>Homogeneous</td>
<td>Heterogeneous, with mixed densities</td>
<td>Heterogeneous, with cystic areas</td>
<td>Heterogeneous, with mixed densities</td>
</tr>
<tr>
<td>Laterality</td>
<td>Usually solitary, unilateral</td>
<td>Usually solitary, unilateral</td>
<td>Usually solitary, unilateral</td>
<td>Often bilateral</td>
</tr>
<tr>
<td>Attenuation (density)</td>
<td>≤10 Hounsfield units (usually &gt;25)</td>
<td>&gt;10 Hounsfield units (usually &gt;25)</td>
<td>&gt;10 Hounsfield units (usually &gt;25)</td>
<td>&gt;10 Hounsfield units (usually &gt;25)</td>
</tr>
<tr>
<td>Vascularity on contrast-enhanced CT</td>
<td>Not highly vascular</td>
<td>Usually vascular</td>
<td>Usually vascular</td>
<td>Usually vascular</td>
</tr>
<tr>
<td>Rapidity of washout of contrast medium</td>
<td>≥50% at 10 minutes</td>
<td>&lt;50% at 10 minutes</td>
<td>&lt;50% at 10 minutes</td>
<td>&lt;50% at 10 minutes</td>
</tr>
<tr>
<td>Appearance on MRI†</td>
<td>Isointense in relation to liver on T2-weighted image</td>
<td>Hyperintense in relation to liver on T2-weighted image</td>
<td>Markedly hyperintense in relation to liver on T2-weighted image</td>
<td>Hyperintense in relation to liver on T2-weighted image</td>
</tr>
<tr>
<td>Necrosis, hemorrhage, or calcifications</td>
<td>Rare</td>
<td>Common</td>
<td>Hemorrhage and cystic areas common</td>
<td>Occasional hemorrhage and cystic areas</td>
</tr>
<tr>
<td>Growth rate</td>
<td>Usually stable over time or very slow (&lt;1 cm per year)</td>
<td>Usually rapid (&gt;2 cm per year)</td>
<td>Usually slow (0.5 cm to 1.0 cm per year)</td>
<td>Variable, slow to rapid</td>
</tr>
</tbody>
</table>

* Adrenal hemorrhage and myelolipoma are usually easily characterized because of their distinctive imaging characteristics. Myelolipomas are composed of myeloid, erythroid, and adipose tissue. On imaging, they have low attenuation on unenhanced CT, and they are hyperintense on T1-weighted in-phase MRI. The presence of pure fat within an adrenal lesion on CT is consistent with the presence of a myelolipoma. Acute adrenal hemorrhage has increased attenuation on unenhanced CT, and on T1-weighted MRI, there is hyperintensity secondary to methemoglobin. In a chronic adrenal hemorrhage, a dark rim develops along the periphery of the mass on the T2-weighted image because of the hemosiderin-laden macrophages.

† If the imaging characteristics are indeterminate on both unenhanced and enhanced CT, MRI may be considered to clarify the imaging phenotype.

ASSESSMENT OF MALIGNANT POTENTIAL

The possibility of malignant disease is the major concern when an incidental adrenal mass is identified. Among 2005 patients in whom adrenal incidentalomas were detected, adrenocortical carcinoma was found in 4.7% of the patients and...
metastatic cancer in 2.5%. The size of the mass and its appearance on imaging are the two major predictors of malignant disease.

**Size of Adrenal Mass**

In a report involving 887 patients who had adrenal incidentalomas, a diameter greater than 4 cm was shown to have 90% sensitivity for the detection of adrenocortical carcinoma but low specificity; only 24% of lesions greater than 4 cm in diameter were malignant. Size is also important because the smaller an adrenocortical carcinoma is at the time of diagnosis, the lower the tumor stage is and the better the overall prognosis will be. Although most experts would recommend resection of adrenal masses larger than 6 cm in diameter, decisions regarding surgery should also take into account the imaging phenotype of the mass, as well as the patient’s age and any coexisting conditions. For example, a nonfunctioning adrenal incidentaloma that is 6.5 cm in diameter and has a benign imaging phenotype may be reasonably followed in an octogenarian. Because the prevalence of benign adrenal cortical adenomas increases with age, the finding of a nonfunctioning adrenal mass that is 3.2 cm in diameter in a younger patient (e.g., below the age of 30 years) should increase the suspicion of an alternative diagnosis. The size of an adrenal incidentaloma does not affect recommendations regarding biochemical testing.

**Imaging Phenotype**

The CT features used to distinguish adenomas from nonadenomas are the lipid content of the

---

**Figure 2.** Pheochromocytoma (Panel A), Benign Cortical Adenoma (Panel B), and Adrenocortical Carcinoma (Panel C).

A heterogeneous (vascular), contrast-enhanced, right adrenal mass, 4.5 cm in diameter (Panel A, arrow), was incidentally revealed on abdominal CT in a 48-year-old woman who was being evaluated for possible appendicitis. The unenhanced CT attenuation was 40 Hounsfield units, and the contrast-medium washout was less than 50% at 10 minutes. The patient had no symptoms or signs of pheochromocytoma. Both urine and plasma normetanephrine levels were markedly elevated. She was treated with α- and β-adrenergic blockade, and a pheochromocytoma was removed. A right adrenal mass (Panel B, arrow), 3.6 cm by 2.5 cm, was incidentally discovered on abdominal CT (performed because of diffuse abdominal discomfort) in a 62-year-old woman with normal blood pressure. The unenhanced CT density (~10 Hounsfield units) and the contrast-medium washout of more than 50% at 10 minutes were consistent with the presence of a cortical adenoma. Hormonal testing for subclinical Cushing’s syndrome and pheochromocytoma was negative. The patient is being followed with repeated imaging and hormonal testing.

A heterogeneous, contrast-enhanced, left adrenal mass (Panel C, arrow), 7.5 cm by 5.5 cm by 6.5 cm, was detected on abdominal CT after measurement of a minimally elevated level of 24-hour urinary 5-hydroxyindoleacetic acid in a 27-year-old woman who had flushing and loose stools. The unenhanced CT attenuation was greater than 10 Hounsfield units, and the contrast-medium washout at 10 minutes was below 50%. Hormonal testing revealed that the mass was nonfunctioning. A laparotomy was performed to remove the mass; the finding on pathological examination was adrenocortical carcinoma.
adrenal mass and rapidity of the washout of contrast medium (Table 3). The intracytoplasmatic fat in adenomas results in low attenuation on unenhanced CT (Fig. 2B); nonadenomas have higher attenuation on unenhanced CT. On chemical-shift MRI (a form of lipid-sensitive imaging that is routinely used), benign adrenocortical adenomas lose signal on out-of-phase images, as compared with in-phase images. However, up to 30% of adenomas do not contain large amounts of lipid and may be indistinguishable from nonadenomas on both unenhanced CT and chemical-shift MRI.

On delayed contrast-enhanced CT, adenomas typically exhibit rapid washout of contrast medium, whereas other adrenal nonadenomas have delayed washout of contrast material (Table 3). Ten minutes after the administration of the contrast medium, an absolute washout of more than 50% of the contrast medium was reported to be 100% sensitive and specific for adenoma in a comparison between patients with adenomas and those with carcinomas, pheochromocytomas, or metastatic disease. Although the imaging phenotype does not predict hormonal function, it does predict the underlying pathology, and surgical resection should be considered for patients who have adrenal incidentalomas with a suspicious imaging phenotype (Fig. 2C).

**Metastatic Disease**

Metastases are the cause of the adrenal incidentaloma in approximately half of patients who have a history of malignant disease. Tumors that commonly metastasize to the adrenals include carcinomas of the lung, kidney, colon, breast, esophagus, pancreas, liver, and stomach (Fig. 3A and 3B). Metastases to the adrenal glands are frequently bilateral. The primary cancer usually has already been recognized when an adrenal incidentaloma is discovered; metastatic cancer to the adrenal without a known primary cancer is extremely rare.

Positron-emission tomography (PET) with \(^{18}\)F-fluorodeoxyglucose (\(^{18}\)F-FDG) can be helpful in selected patients (those with a history of malignant disease) because of its high sensitivity in detecting malignant diseases. However, 16% of benign adrenal lesions may have greater FDG-PET uptake than the background uptake. The absence of activity on \(^{11}\)C-metomidate (MTO)–PET appears to be specific for tumors of nonadrenocortical origin (e.g., pheochromocytomas and metastatic disease), but this type of imaging is not routinely available. Because of their cost and because there are insufficient data to support their routine use, FDG-PET and MTO-PET are not recommended for the evaluation of a patient with an adrenal incidentaloma who does not have a history of malignant disease.

**Fine-Needle Aspiration Biopsy**
The primary role of fine-needle aspiration biopsy is to differentiate between adrenal tissue and nonadrenal tissues (e.g., metastases or infection). Image-guided fine-needle aspiration biopsy is relatively safe; the complication rate was 2.8% in one series of 277 biopsies. The risks of this procedure include adrenal hematoma, abdominal pain, hematuria, pancreatitis, pneumothorax, formation of an adrenal abscess, and tumor recurrence along the needle track. Also, fine-needle aspiration biopsy of a pheochromocytoma may result in hemorrhage and hypertensive crisis, and the possibility of pheochromocytoma should always be ruled out by biochemical testing before fine-needle aspiration biopsy is undertaken.

**BILATERAL ADRENAL MASSES**

When adrenal masses occur bilaterally (as they do in up to 15% of patients with adrenal incidentaloma), the most likely diagnoses are metastatic disease, congenital adrenal hyperplasia, bilateral cortical adenomas, and infiltrative disease of the adrenal glands. Adrenocortical hypofunction may occur in patients with bilateral adrenal masses, so screening for adrenocortical hypofunction may be prudent in such patients, although the yield is unknown.

### AREAS OF UNCERTAINTY

The optimal frequency and duration of follow-up for patients who have adrenal incidentalomas are uncertain, and prospective data to guide the clinician are scarce. Repeated imaging is commonly recommended at 6, 12, and 24 months; earlier follow-up (at 3 months) has been suggested when the imaging phenotype is suspicious (Fig. 1), with the rationale that many malignant lesions will grow during this 3-month interval (Fig. 3A and 3B), resulting in earlier intervention. However, the yield and cost-effectiveness of repeated imaging at these intervals are uncertain. On the basis of our
unpublished experience with nine patients who underwent serial imaging, the typical rate of growth of benign adrenal pheochromocytoma is approximately 0.5 to 1.0 cm in diameter per year (Fig. 3C through 3F), whereas adrenocortical carcinomas typically have a rapid growth rate (>2 cm per year) (Fig. 3G and 3H). However, most adrenal masses that grow are not malignant. In case series of adrenal incidentalomas followed for an average of 4 years, 5 to 20% increased in size, and 1.3 to 5.2% decreased in size.\textsuperscript{1,21,48} In two of these series, only 1 of 9 patients and none of 11 patients with enlarging adrenal masses who underwent surgery were found to have malignant tumors.\textsuperscript{21,48} Less frequent imaging during follow-up is reasonable for patients who have no history of malignant disease and who have small (<2 cm), uniform, hypodense cortical nodules.

The observation that abnormal adrenal function (secretion of glucocorticoids and catecholamines) that is not present at baseline may be detected during follow-up testing\textsuperscript{21,48,49} has led to the recommendation of repeating hormonal evaluation annually for at least 4 years when the initial evaluation is negative.\textsuperscript{5,48,49} However, the yield and cost-effectiveness of such testing are unknown.

**GUIDELINES**

No comprehensive guidelines have been published by professional societies to guide the evaluation of patients with adrenal incidentalomas. The rec-
Conclusions and Recommendations

For the patient described in the vignette, a thorough history should be obtained and a physical examination performed to assess the evidence of adrenal hormone excess (Table 1). I would perform a 1-mg overnight dexamethasone suppression test, collect a 24-hour urinary specimen for measurement of fractionated metanephrines and catecholamines, and (because she has hypertension) measure the plasma aldosterone concentration and plasma renin activity. If the results of the initial hormonal testing are consistent with autonomous hormone secretion, and if this finding is confirmed by subsequent studies, unilateral laparoscopic adrenalectomy should be considered.

The adrenal imaging should be reviewed with a radiologist. If the imaging phenotype suggests infection or metastatic disease, CT-guided fine-needle aspiration biopsy should be considered (after biochemical testing to rule out pheochromocytoma). If the results of hormonal testing are normal and the imaging features are consistent with benign disease, I would recommend repeating the imaging studies at 6, 12, and 24 months and repeating the hormonal evaluation yearly for 4 years, even though there are no data from large, long-term studies to support these recommendations. Although the data are also scarce to suggest when surgery is necessary, I would recommend consideration of adrenalectomy if the adrenal mass is 4 cm or greater in diameter, if the mass enlarges by 1 cm or more during the period of observation, or if evidence of autonomous hormonal secretion develops.

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

References

A 38-year-old man presented to the emergency department after reportedly ingesting antifreeze. He appeared to be intoxicated and was agitated and combative; chemical sedation was induced. Initial laboratory studies revealed a pH of 7.0, an anion gap of 22 mmol per liter, and an osmolar gap of 79 mOsm. It was noted that the patient’s urine fluoresced under ultraviolet light (in the basin on the left), as compared with a negative control (in the basin on the right), which shows the purple reflection of the ultraviolet light (arrow). The patient received fomepizole, thiamine, folate, pyridoxine, and bicarbonate; he subsequently underwent hemodialysis. Laboratory studies revealed that his ethylene glycol level had been 222 mg per deciliter when the treatment began. His recovery was uneventful.

Fluorescein is a fluorescent dye added to antifreeze preparations to aid in the detection of radiator leaks. In addition to the history and elevated osmolar and anion gaps, the fluorescence of urine under ultraviolet light may aid in the early identification of ethylene glycol poisoning. False negative and false positive results may occur. For example, many containers, such as urine collection bags, may be characterized by native fluorescence.

Copyright © 2007 Massachusetts Medical Society.
Presentation of Case

A 56-year-old woman was admitted to the hospital because of rapidly progressive vertigo and ataxia.

The patient had been well until approximately 10 weeks before admission, when occasional dizziness and nausea occurred, followed during the next several weeks by increasing positional vertigo and severe vomiting. Antiemetic agents were administered, the vomiting resolved, and her dizziness improved. Shortly thereafter, slurred speech, rapidly progressive ataxia, and difficulty with ambulation developed.

Approximately 6 weeks before admission, the patient saw a physician at another facility. Cranial magnetic resonance imaging (MRI) showed an increased T2-weighted signal in the periventricular white matter that was thought to represent either ischemia or demyelination. Computed tomographic (CT) scanning of the brain showed no discrete lesions. Warfarin therapy was initiated but discontinued shortly thereafter, when coagulation studies yielded supratherapeutic results. Fresh-frozen plasma was infused, and the results of coagulation studies returned to normal.

A lupus anticoagulant was not detected. A complete blood count, levels of serum protein, electrolytes, calcium, and phosphorus, and the results of renal- and liver-function tests were normal. A lumbar puncture revealed a normal opening pressure; the results of laboratory tests on the cerebrospinal fluid are shown in Table 1. No organisms were seen on staining or in cultures; a cryptococcal antigen test and a polymerase-chain-reaction test for herpes simplex virus were both negative. Tests for human immunodeficiency virus types 1 and 2, antinuclear antibodies, and rapid plasma reagin were also negative.

One week later, a second MRI study showed a new, small, linear area of increased signal in the left pontine region on fluid-attenuated inversion recovery (FLAIR) sequences. A magnetic resonance angiogram of the carotid arteries revealed an anatomical variant of the left vertebral artery, mild irregularities of the basilar artery without evidence of occlusion, and less than 15% stenosis of the carotid arteries. A 3-day course of high-dose methylprednisolone was administered, followed by a tapered course of prednisone, with no improvement.

Approximately 3 weeks before admission, the patient was evaluated in a neurology clinic at another hospital. On examination, she was seated in a wheelchair and had a to-and-fro head tremor, flat affect, and severely dysarthric speech. She was alert...
and oriented, with some deficiencies in performing multistep, sequenced commands. Visual testing revealed diminished pursuit and saccadic overshoot. Alternating movements of the hands were severely slowed, dexterity was diminished, and there was mild hip-flexor weakness and paratonia; strength and tone in other muscles were normal. Finger-to-nose testing revealed severe dysmetria bilaterally, and heel-knee-shin testing could not be performed. There was truncal titubation, and the patient was unable to walk. Reflexes were normal.

One week before admission, CT scanning of the chest revealed minimal thickening of the pleura of the right lung, strands of subsegmental atelectasis or fibrosis in the right posterior costophrenic sulcus, and oblique fissure of the left lung. Mammography and CT scanning of the abdomen and pelvis showed no abnormalities.

One week later, the patient saw a neurologist at this hospital. Her neurologic symptoms had been stable for the previous 2 weeks. She had hypertension and hypercholesterolemia. Her medical history included an episode of otitis media, 7 years earlier, and a bilateral tubal ligation. She did not use alcohol and had stopped smoking cigarettes 20 years earlier. She was married, with a 35-year-old daughter in good health. Her mother had had breast cancer in her eighth decade, and her father had Parkinson’s disease. There was no family history of multiple sclerosis. Medications included atorvastatin, aspirin, and hydrochlorothiazide, with promethazine as needed for dizziness.

The vital signs were normal. The patient was alert and oriented, with tremors and stuttering speech. The score on the Mini–Mental State Examination was 27 (the highest possible score is 30), with points lost on calculations and on the recall of two out of three objects at 5 minutes. Her gaze had lost smooth pursuit, and saccadic overshoot was present. There was difficulty with pronunciation of the letter K, and the gag reflex was reduced. Cranial nerve II was not tested; cranial nerves II through VII were intact. Motor and sensory examinations were normal; reflexes were 2+ to 3+ throughout. Marked dysmetria and dysdiadochokinesia were present in the arms, and there was difficulty with the amplitude and velocity of fine finger movements and with the control of movement during strength testing. Examination of the legs revealed dysmetria with foot tapping; the heel-knee-shin test could not be performed. The patient was unable to stand unless given full assistance, and she was unable to walk. There was no response to plantar stimulation. The results of the remainder of the examination were normal.

The patient was admitted to the hospital that day. A complete blood count, prothrombin and partial-thromboplatin times, liver- and renal-function tests, and serum levels of electrolytes, immunoglobulins, CA 125, CA 19-9, carcinoembryonic antigen, thyroxine, and thyroid-stimulating hormone were normal; additional results were pending. On the second hospital day, MRI revealed a mass, 12 mm in diameter, in the inferior left breast. CT scanning of the chest and abdomen showed focal fatty change in the liver and a nodule, 11 mm in diameter, in the left adrenal gland that appeared to be an adenoma; no other abnormalities were seen. Ultrasonography of the pelvis showed no abnormalities.

On the fourth hospital day, MRI of the brain showed moderate volume loss in the cerebellar

<table>
<thead>
<tr>
<th>Table 1. Results of Cerebrospinal Fluid Tests.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
</tr>
<tr>
<td><strong>Reference Range in Adults</strong></td>
</tr>
<tr>
<td><strong>6 Weeks before Admission</strong></td>
</tr>
<tr>
<td>Red-cell count (per mm$^3$)</td>
</tr>
<tr>
<td>White-cell count (per mm$^3$)</td>
</tr>
<tr>
<td>Differential count (%)</td>
</tr>
<tr>
<td>Neutrophils</td>
</tr>
<tr>
<td>Lymphocytes</td>
</tr>
<tr>
<td>Monocytes</td>
</tr>
<tr>
<td>Protein (mg/dl)</td>
</tr>
<tr>
<td>Glucose (mg/dl)$†$</td>
</tr>
<tr>
<td>Venereal Disease Research Laboratory test</td>
</tr>
<tr>
<td>IgG (mg/dl)</td>
</tr>
<tr>
<td>IgG synthetic rate (mg/day)</td>
</tr>
<tr>
<td>IgG index</td>
</tr>
<tr>
<td>IgG:albumin ratio</td>
</tr>
<tr>
<td>Albumin (mg/dl)</td>
</tr>
<tr>
<td>Myelin basic protein (ng/ml)</td>
</tr>
<tr>
<td>Angiotensin-converting enzyme (U/liter)</td>
</tr>
<tr>
<td>Antineuronal nuclear antibodies (anti-Hu and anti-Ri antibodies)</td>
</tr>
</tbody>
</table>

* Reference values are affected by many variables, including the patient population and the laboratory methods used. The ranges used at Massachusetts General Hospital are for adults who are not pregnant and do not have medical conditions that could affect the results. They may therefore not be appropriate for all patients.

† To convert the values for glucose to millimoles per liter, multiply by 0.05551.
hemispheres without evidence of abnormal T₂-weighted signals or abnormal diffusion. There were multiple foci of T₂-weighted hyperintensity in the white matter, with subtle gadolinium enhancement in the right corona radiata. Two days later, positron-emission tomographic (PET) scanning showed hypermetabolic foci in both axillae and in a lower thoracic vertebra at approximately T10. However, a review of the chest CT scan obtained earlier in this hospitalization did not reveal any corresponding mass or lymphadenopathy in the axillary regions. Serum protein electrophoresis revealed an abnormal pattern, with one very-low-concentration band identified as IgG kappa M component. On the seventh hospital day, a biopsy of a hyperechoic solid mass in the inferior left breast, performed under ultrasonographic guidance, revealed atrophic breast tissue with no malignant cells. On the next day, an excisional biopsy of the lesion in the left breast revealed a myxoid fibroadenoma, with no evidence of cancer.

A diagnostic test result was received.

**DIFFERENTIAL DIAGNOSIS**

_Dr. Josep Dalmau:_ May we review the imaging studies?

_Dr. R. Gilberto Gonzalez:_ The MRI scan obtained on the patient’s admission to this hospital revealed evidence of cerebellar and brain-stem atrophy and abnormal signals in the cerebral white matter. A sagittal midline T₁-weighted image of the brain (Fig. 1A) shows shrinkage of the cerebellar vermis with prominent fissures. Axial T₂-weighted images through the posterior fossa showed tissue loss in the cerebellum, extensive cerebrospinal fluid in the posterior fossa, and enlargement of the fourth ventricle. The cerebellar hemispheres, vermis, and pons were small. In the cerebrum, there were scattered foci of T₂-weighted hyperintensity in the subcortical white matter. The largest signal abnormality was a focus of about 8 mm in diameter in the right corona radiata that showed subtle enhancement on images obtained after the administration of contrast material.

An ¹⁸F-fluorodeoxyglucose PET scan obtained on the sixth day showed three abnormal foci: one each in the right and left axillae (Fig. 1B) and a third in a thoracic vertebra at approximately T10.

_Dr. Dalmau:_ In this adult with no history of immunodeficiency, the subacute development of a predominant cerebellar syndrome with inflammatory abnormalities of the cerebrospinal fluid and hypermetabolic foci on a body PET scan narrows the differential diagnosis to a few disorders. Diagnoses that could have been considered before the PET scan was obtained are shown in Table 2. The absence of a history of alcohol abuse and nutritional deficiencies, the rapid course of the disease, the inflammatory findings in the cere-
brospinal fluid, the atrophic changes in the symptomatic regions on MRI, and the PET findings, either alone or in combination, readily rule out most of these disorders.

CEREBROVASCULAR DISEASE
The rapid development of dysarthria and ataxia and the presence of risk factors for cerebrovascular disease initially suggested a vertebrobasilar system stroke. Treatment with warfarin was begun, but a supratherapeutic effect was noted and the treatment was discontinued. The presence of a lupus anticoagulant was ruled out. These findings do not rule out the possibility of a subclinical tendency to fibrinolysis, which occurs in some malignant conditions. In this patient, the results of magnetic resonance angiography ruled out vascular occlusion, and the cerebrospinal fluid findings suggested causes other than stroke.

DEMYELINATING DISORDERS
The patient was treated with corticosteroids, probably because a demyelinating disorder was considered. Among patients with multiple sclerosis, the median age at diagnosis is 30 years; this 56-year-old woman had no history of visual or neurologic deficits. The supratentorial findings on $T_2$-weighted MRI are more suggestive of small-vessel ischemic disease. Elevated levels of myelin basic protein in the cerebrospinal fluid can occur in disorders other than multiple sclerosis and may also occur in paraneoplastic disorders. These findings and the absence of a clinical response to corticosteroids do not support the diagnosis of multiple sclerosis.

SARCOIDOSIS
Sarcoidosis should be included in the differential diagnosis. However, the symptoms of sarcoidosis are not as acute as those in this patient, and the cranial nerves, basal meninges, hypothalamus, and pituitary region are often involved, which is not the case in this patient. A CT scan of the chest is abnormal in 60 to 70% of patients with sarcoidosis. In this patient, the absence of such findings and the normal level of angiotensin-converting enzyme in the cerebrospinal fluid argue against the diagnosis of sarcoidosis.

PARANEOPLASTIC SYNDROMES AFFECTING THE CENTRAL NERVOUS SYSTEM
I suspect that this patient has a neurologic complication of cancer. Metastatic involvement of the central nervous system is unlikely because of the absence of contrast-enhancing abnormalities in the posterior fossa. Rare tumors such as intravascular lymphoma and so-called carcinomatous encephalitis caused by miliary metastases may not be characterized by contrast enhancement or mass effect, but the predominant cerebellar syndrome and cerebrospinal fluid abnormalities in this patient do not support these diagnoses.

Paraneoplastic disorders can affect any part of the nervous system and are believed to be immune-mediated. The likelihood that a disorder is paraneoplastic depends on the clinical syndrome (i.e., cerebellar degeneration or limbic encephalitis) and on whether paraneoplastic antibodies and a tumor are detected. In 60% of patients with paraneoplastic disorders, the neurologic symptoms precede the diagnosis of cancer; the absence of a known cancer in this patient is therefore typical. It has been postulated that the antitumor immune response may contribute to the small size of the tumor, making detection difficult. A combination of CT and body PET imaging facilitates detection of these tumors.

Immunologic Features of Paraneoplastic Syndromes
The immune responses associated with paraneoplastic syndromes are shown in Table 3. Antibodies appear to be necessary but not sufficient alone to cause neurologic dysfunction, and cytotoxic T-cell responses are also involved. In this patient, identification of any of the antibodies listed in Table 3 would confirm the diagnosis of a paraneoplastic syndrome; depending on the antibody identified, this finding would also direct the search for the tumor. In approximately 40% of patients, no antibodies are identified.

<table>
<thead>
<tr>
<th>Table 2. Possible Causes of Acquired, Rapidly Progressing Cerebellar Ataxia in Immunocompetent Patients.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
</tr>
<tr>
<td>Sarcoidosis</td>
</tr>
<tr>
<td>Primary or metastatic tumors in the posterior fossa</td>
</tr>
<tr>
<td>Paraneoplastic neurologic disorders</td>
</tr>
<tr>
<td>Exposure to toxins and drugs (lead, anticonvulsants, salicylates, aminoglycosides, sedatives, fluorouracil, cytarabine)</td>
</tr>
<tr>
<td>Miller Fisher syndrome</td>
</tr>
<tr>
<td>Infection (human immunodeficiency virus infection, viral cerebellitis, Creutzfeldt–Jakob disease)</td>
</tr>
<tr>
<td>Alcoholic cerebellar degeneration</td>
</tr>
<tr>
<td>Vitamin deficiency (thiamine)</td>
</tr>
<tr>
<td>Autoimmune disease (systemic lupus erythematosus, Sjögren’s syndrome, Hashimoto’s disease, cerebellar ataxia with anti–glutamic acid decarboxylase antibodies, cerebellar ataxia with antigliadin antibodies)</td>
</tr>
</tbody>
</table>

n engl j med 356;6 www.nejm.org february 8, 2007 615
Clinical Features of Paraneoplastic Cerebellar Degeneration

The development of paraneoplastic cerebellar degeneration is usually rapid, leaving the patient severely disabled in days to weeks.14 This patient’s presentation is typical, beginning with the dizziness, vertigo, and nausea that may suggest a peripheral vestibular problem. These symptoms are followed by ataxia of the trunk and limbs, variable oscillopsia, blurry vision, diplopia, nystagmus, dysarthria, tremor, and sometimes dysphagia. The ocular motor abnormalities can be complex because there is always some degree of brain-stem involvement, as shown in autopsy studies.15 The clinical features in this patient indicate predominant cerebellar dysfunction; the decreased gag reflex with no other abnormalities on cranial-nerve examination suggests mild involvement of the brain stem. Mild memory and cognitive deficits, detected during the examination, occur in about 20% of patients with paraneoplastic cerebellar degeneration.14

The abnormalities in the cerebrospinal fluid suggest an inflammatory or immune-mediated process; similar abnormalities are found in approximately 70 to 80% of patients with paraneoplastic cerebellar degeneration. Paraneoplastic antineuronal nuclear antibodies (anti-Hu and anti-Ri antibodies) were not detected in this patient’s cerebrospinal fluid; we do not know whether tests for other antibodies were performed. The two antibodies that have the highest specificity for cerebellar dysfunction are anti-Yo16,17 and anti-Tr antibodies.18 The remaining antibodies listed in Table 3 either are associated with syndromes that frequently affect other areas of the nervous system in addition to the cerebellum or are infrequently detected in patients with paraneoplastic syndromes (i.e., mGluR1).14

In most patients with paraneoplastic cerebellar degeneration, MRI scans are normal early in the disease whereas subsequent studies reveal cerebellar atrophy, as occurred in this case.13,14,16,19 The clinical significance of the other MRI findings in this patient is unclear. The abnormality in the FLAIR signal in the brain stem could be related to transient inflammation, which occurs in the hippocampus in patients with limbic encephalitis.20 The subtle enhancement in the right corona radiata suggests telangiectasia, and the

<table>
<thead>
<tr>
<th>Antibodies Predominantly associated with PCD</th>
<th>Predominant Syndrome</th>
<th>Associated Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-Yo (PCA-1) antibodies</td>
<td>PCD</td>
<td>Ovarian and breast cancers</td>
</tr>
<tr>
<td>Anti-Tr antibodies</td>
<td>PCD</td>
<td>Hodgkin’s lymphoma</td>
</tr>
<tr>
<td>Anti-mGluR1 antibodies‡</td>
<td>PCD</td>
<td>Hodgkin’s lymphoma</td>
</tr>
<tr>
<td>Anti-Zic4 antibodies‡</td>
<td>PCD</td>
<td>Small-cell lung cancer</td>
</tr>
<tr>
<td>Anti-VGCC antibodies</td>
<td>Eaton–Lambert syndrome, PCD</td>
<td>Small-cell lung cancer and lymphoma</td>
</tr>
<tr>
<td>Anti-Hu (ANNA-1) antibodies</td>
<td>Encephalomyelitis, PCD, sensory neuropathy</td>
<td>Small-cell lung cancer and other cancers</td>
</tr>
<tr>
<td>Anti-Ri (ANNA-2) antibodies</td>
<td>PCD, brain-stem encephalitis, paraneoplastic opoclonus–myoclonus</td>
<td>Breast, gynecologic, and small-cell lung cancers</td>
</tr>
<tr>
<td>Anti-CV2/CRMP5 antibodies</td>
<td>Encephalomyelitis, PCD, chorea, peripheral neuropathy, uveitis</td>
<td>Small-cell lung cancer, thymoma, and other cancers</td>
</tr>
<tr>
<td>Anti-Ma protein antibodies§</td>
<td>Limbic, hypothalamic, brain-stem encephalitis (infrequently PCD)</td>
<td>Testicular, lung, and other cancers</td>
</tr>
<tr>
<td>Anti-amphiphysin antibodies</td>
<td>Stiff-person syndrome, encephalomyelitis, PCD</td>
<td>Breast and small-cell lung cancers</td>
</tr>
</tbody>
</table>

* There is no uniform nomenclature for some of these antibodies; variant names appear in parentheses. mGluR1 denotes metabotropic glutamate receptor 1, Zic4 zing finger of the cerebellum 4, and VGCC voltage-gated calcium channel.‡ Anti-mGluR1 antibodies have been identified in only two patients.‡ Anti-Zic4 antibodies are predominantly associated with PCD only when no other paraneoplastic antibodies are detectable.§ Ma proteins include Ma1 and Ma2. Patients with brain-stem and cerebellar dysfunction usually have antibodies against both Ma1 and Ma2.
abnormalities on T2-weighted images are probably nonspecific findings or manifestations of small-vessel ischemic disease.

Tumors Associated with Paraneoplastic Syndromes of the Central Nervous System

What neoplasm is responsible for this woman’s disorder? At her age, the most common tumors associated with paraneoplastic cerebellar degeneration are cancers of the ovary, breast, and lung and lymphoma. The absence of abdominal symptoms, the negative results of tests for cancer markers, and the distribution of findings on the PET scan do not suggest an ovarian cancer. The patient has a family history of breast cancer, and there is weak evidence that fibroadenoma may be a risk factor for cancer. However, simultaneous bilateral axillary involvement and a metastasis at T10 would be a very unusual presentation for breast cancer. A small single focus of 18F-fluorodeoxyglucose uptake in the axilla on the side of the injection can result from subcutaneous radioactive tracer infiltration and may lead to a false positive result.

Lung cancer, particularly small-cell lung cancer, is the cancer most frequently associated with paraneoplastic syndromes of the central nervous system. These syndromes can be manifested as cerebellar dysfunction that remains isolated or progresses to involve other parts of the central nervous system (encephalomyelitis) and are often associated with anti-Hu antibodies. About 40% of patients who have small-cell lung cancer and cerebellar degeneration without anti-Hu antibodies have voltage-gated calcium-channel antibodies, which are also present in patients with the Eaton–Lambert myasthenic syndrome. Except for mild proximal leg weakness, this patient did not have symptoms suggestive of the Eaton–Lambert myasthenic syndrome. The remote history of smoking, the absence of abnormalities in the lungs and mediastinum on PET scanning, and the absence of anti-Hu antibodies make lung cancer unlikely.

Lymphomas are the next most common tumors associated with paraneoplastic cerebellar degeneration. The distribution of the PET findings and the presence of an IgG kappa M component could suggest non-Hodgkin’s lymphoma, although the latter finding also occurs in 1% of the healthy population. A monoclonal IgG gammopathy is typical of multiple myeloma, but this tumor is rarely associated with cerebellar degeneration and does not involve the axillary nodes. Patients with cerebellar degeneration and Hodgkin’s lymphoma usually have anti-Tr antibodies, whereas those with non-Hodgkin’s lymphoma rarely have any of the antibodies known to be associated with paraneoplastic cerebellar degeneration. In addition to the cerebellar degeneration, this patient may have had a paraneoplastic coagulopathy that contributed to the supratherapeutic effect of warfarin. Coagulopathy may occur in association with diverse types of cancers or B-cell neoplasms. The serum level of lactate dehydrogenase in this patient is not provided, but an elevated level would support the diagnosis of lymphoma.

Summary

I believe that the diagnosis in this case is a paraneoplastic syndrome with predominant cerebellar degeneration. Breast cancer in association with anti-Yo antibodies would be my leading diagnosis. The diagnostic procedure was probably a biopsy of the lesion in the right axilla or the lesion in the T10 vertebral body, detected by PET scanning.

Dr. Nancy Lee Harris (Pathology): Dr. Batchelor, would you summarize your thinking and tell us about additional testing that was done?

Dr. Tracy T. Batchelor (Neuro-oncology): In this patient, a diagnosis of a paraneoplastic cerebellar degeneration was strongly suspected. The anti-Yo antibody test performed at the first hospital was eventually received and was reported to be positive. We repeated the panel of paraneoplastic antibody tests.

Clinical Diagnosis

Paraneoplastic syndrome with predominant cerebellar degeneration.

Dr. Josep Dalmau’s Diagnosis

Paraneoplastic syndrome with paraneoplastic cerebellar degeneration, probably associated with anti-Yo antibodies and a primary breast cancer.

Pathological Discussion

Dr. Batchelor: Tests for MaTa antibodies, CV2/CRMP5 antibodies, antineuronal nuclear (Hu, Ri) antibodies, anti-CAR antibodies, and anti-voltage-
gated calcium-channel antibodies (associated with the Eaton–Lambert myasthenic syndrome) were negative. A test for anti–Purkinje-cell (Yo) antibodies was positive.

A 5-day trial of intravenously administered immunoglobulin did not lead to improvement in the patient’s neurologic condition, and she was discharged to a rehabilitation center on hospital day 12. Because of concern about the possibility of an occult ovarian cancer, laparoscopic hysterectomy and salpingo-oophorectomy were performed 2 months after discharge; no malignant tumor was found. We then repeated the CT and PET scans.

Dr. Gonzalez: A PET scan obtained 4 months after the initial PET scan revealed abnormalities in the right axilla and chest wall (Fig. 1C). There was still an abnormality in the T10 area. The left axilla showed no abnormalities. CT scanning showed the presence of an enlarged lymph node, 1.4 cm in diameter, in the right axilla, which was enhanced after the administration of contrast material.

Dr. Melinda F. Lerwill: Excision of the patient’s enlarged right axillary lymph node revealed a poorly differentiated malignant neoplasm (Fig. 2A); the malignant cells were large, with abundant eosinophilic cytoplasm and marked nuclear pleomorphism (Fig. 2B). The differential diagnosis included carcinoma, lymphoma, and melanoma. Immunohistochemical studies showed that the tumor cells were positive for cytokeratin, supporting a diagnosis of carcinoma and negative for the lymphoid marker leukocyte common antigen and the melanoma marker S-100. The tumor cells were positive for cytokeratin 7 and negative for cytokeratin 20, a profile common to carcinomas from diverse sites, including the breast, lung, and ovary. The cells were diffusely positive for gross cystic disease fluid protein-15 (Fig. 2C), a marker of apocrine differentiation that is expressed in 62 to 77% of breast carcinomas as well as in salivary gland and skin adnexal tumors.26 Gross cystic disease fluid protein-15 is only rarely positive in other cancers and is therefore a fairly specific marker of cancer originating in the breast when salivary gland and skin adnexal tumors are not likely. In conjunction with the clinical picture, the immunohistochemical findings in this case supported a diagnosis of metastatic breast carcinoma. The tumor cells did not express either estrogen or progesterone receptor, but they did show 3+ overexpression of Her-2/neu (Fig. 2D).

Less than 1% of patients with breast cancer present with an axillary metastasis as the first clinical indication of disease.27 Clinical and radiographic evaluation of the breasts often shows no abnormalities in such patients even with the use of MRI.27,28 A primary carcinoma is detected in the breast on pathological examination in 50 to 80% of cases,27 sometimes months to years after the axillary metastases are detected. In nearly two thirds of cases in which axillary metastasis is the first clinical indication of breast cancer, the carcinoma is poorly differentiated, has a diffuse growth pattern, and is composed of large cells with abundant eosinophilic cytoplasm, pleomorphic nuclei, and macronucleoli, all features of this patient’s tumor.29

Dr. Harris: Dr. Schmahmann, would you comment on your evaluation of this patient?

Dr. Jeremy D. Schmahmann (Neurology): I saw this patient during her first admission and was struck by the impairment in mental flexibility combined with poor set shifting, deficient verbal working memory, decreased verbal fluency, poverty of initiation, limited verbal learning and short-term recall, and flattening of affect. Because of the cerebellar involvement, the presence of anti-Yo antibodies, and the absence of demonstrated involvement of the cerebral hemispheres, I believed this patient had the cerebellar cognitive affective syndrome.30 This neurobehavioral disorder is characterized by deficits in executive, visual–spatial, and linguistic performance, with disordered regulation of affect. Treatment with sertraline was instituted, and the patient’s mood and comportment improved. The recognition of the cerebellar cognitive affective syndrome in this case was thus clinically relevant, since it resulted in specific treatment that improved her quality of life.

Dr. David P. Ryan (Medical Oncology): Tumor-marker studies were repeated; the CA 15-3 level was 45 U per milliliter (normal range, 0 to 30), and the carcinoembryonic antigen level was 14.8 ng per milliliter (normal value, <3.4). Repeated MRI showed progressive cerebellar volume loss. The patient was treated with vinorelbine and trastuzumab for 6 months. Follow-up CT scans showed complete regression of the axillary and chest-wall abnormalities, with no new metastatic disease; the levels of tumor markers returned to normal values. However, the patient’s neurologic status was unchanged. Two years after the onset of symptoms, she remained free of cancer; al-
though her neurologic status did not improve, it did not deteriorate further.

A Physician: Do the tumors in such patients express the Yo antigen?

Dr. Dalmau: Yes, the tumors in patients with anti-Yo antibodies express the Yo antigen.31

Dr. Robert H. Brown, Jr. (Neurology): What are the proteins to which the antibodies are directed? Is there any correlation between their distribution in the nervous system and the phenotype of the paraneoplastic syndrome?

Dr. Dalmau: In most of these syndromes, the antigens have been identified, and the genes cloned.32 The Yo antigen is a cytoplasmic protein called CDR2 that interacts with c-Myc. The protein is expressed predominantly in the Purkinje cells of the cerebellum and the large neurons of the brain stem. Studies suggest that CDR2 sequesters c-Myc in the neuronal cytoplasm and downregulates its activity. Disruption of this interaction by anti-Yo antibodies may increase c-Myc activity, leading to apoptosis of the Purkinje cells. Antibodies could therefore play an initial pathogenic role, although it is believed for the most part that the T-cell immune response is the major effector of neuronal degeneration.33

**ANATOMICAL DIAGNOSIS**

Paraneoplastic cerebellar degeneration due to anti-Yo antibodies.

Metastatic poorly differentiated carcinoma, involving a right axillary lymph node, consistent with metastasis from a primary cancer in the breast.
(estrogen receptor negative and progesterone receptor negative, with overexpression of HER2/neu).

Dr. Dalmau reports receiving royalties from a patent held by Memorial Sloan-Kettering Cancer Center, New York, licensed to Athena Diagnostics (Ma-family polypeptides and anti-Ma antibodies) and consulting fees, advisory board payments, and lecture fees from Athena Diagnostics. No other potential conflict of interest relevant to this article was reported.

REFERENCES


Copyright © 2007 Massachusetts Medical Society.
Treating the Polycystic Ovary Syndrome the Old-Fashioned Way

David S. Guzick, M.D., Ph.D.

Of the estimated 6.7 million women with fertility problems in the United States, 35% have received drugs to induce ovulation. The most common cause of anovulation among infertile women is the polycystic ovary syndrome, a condition typically characterized by irregular menses, androgen excess, and polycystic-appearing ovaries. The condition is associated with insulin resistance and obesity.

Although a number of drugs have been used to induce ovulation in women with the polycystic ovary syndrome, clomiphene citrate is a simple, tried-and-true treatment. Clomiphene is an orally active, antiestrogenic substance that promotes the release of follicle-stimulating hormone from the pituitary gland, thus stimulating the development of ovarian follicles and ovulation. Early experience with clomiphene in the treatment of infertile, anovulatory women (about half of whom were likely to have had the polycystic ovary syndrome on the basis of the presence of hyperandrogenemia) yielded a cumulative pregnancy rate of 56% after six cycles of treatment. More recently, a cumulative pregnancy rate of 73% was reported when treatment with clomiphene citrate was repeated for up to nine ovulatory cycles.

Clomiphene has its drawbacks, however. It is associated with a multiple pregnancy rate of 5 to 10%. Also, it does not address the underlying abnormalities in the polycystic ovary syndrome, including hyperinsulinemia and hyperandrogenism. It is well known that interventions that improve insulin resistance and reduce hyperinsulinemia, such as weight loss, in women with this syndrome also reduce hyperandrogenemia and induce ovulation in many cases.

On the basis of these observations, Nestler et al. posited that metformin, which also reduces hyperinsulinemia, might be effective in treating obese, infertile women with the polycystic ovary syndrome. In a clinical trial reported in 1998, 61 such women, who had an average body-mass index (the weight in kilograms divided by the square of the height in meters) of 32, were randomly assigned to receive either metformin or placebo pills. Clomiphene was added in the second cycle if there was no ovulation. Overall, 89% of the women who were treated with metformin ovulated, either spontaneously or in response to clomiphene, as compared with 12% of the women who were treated with placebo, either alone or with clomiphene.

This report sparked the clinical use of metformin in the treatment of infertile women with the polycystic ovary syndrome, as well as a number of clinical trials. A meta-analysis of 13 randomized trials comparing metformin with placebo, or metformin plus clomiphene with clomiphene alone, in women with the polycystic ovary syndrome concluded that metformin increased the ovulation rate by a factor of approximately four. Of note, pregnancy rates did not differ significantly between the metformin groups and the placebo groups, although the pregnancy rates for metformin plus clomiphene were significantly higher than for clomiphene alone. More recently, a clinical trial in Italy, in which 100 infertile, nonobese women with the polycystic ovary syndrome were randomly assigned to receive either metformin or clomiphene, showed similar rates of ovulation in the two groups, although the pregnancy rate in the metformin group was twice that in the clomiphene group. However, in a multicenter, randomized trial that compared clomiphene plus metformin with clomiphene plus placebo in 225 infertile Dutch women with the
polycystic ovary syndrome (mean body-mass index, 28), the addition of metformin did not significantly improve rates of either ovulation or pregnancy.10

In this issue of the Journal, Legro et al. report on the results of a large multicenter, randomized trial comparing the effects of clomiphene, extended-release metformin, and combination therapy with both agents in 626 infertile women with the polycystic ovary syndrome (mean body-mass index, 35).11 At 6 months, the live-birth rate (the primary end point) in the clomiphene group was three times that in the metformin group, and there was no significant benefit of the combination of metformin and clomiphene as compared with clomiphene alone. Although the ovulation rate in the combination-therapy group was significantly higher than that in the other groups, the increase did not translate into a higher live-birth rate. Of the 115 pregnancies that occurred in the two clomiphene groups, there were four pairs of twins and one set of triplets. By contrast, there were no multiple births among the 18 pregnancies that occurred in the metformin group. With this exception, there were no differences in adverse effects among the groups.

How can one reconcile the finding that the live-birth rate in the clomiphene group was three times that in the metformin group with the meta-analysis of several earlier trials suggesting the superiority of metformin?8 Inconsistencies have previously been recognized between results of meta-analyses of small trials and subsequent large, randomized, controlled trials22; such discrepancies are more likely to occur when studies in a meta-analysis have methodologic flaws or when there is heterogeneity in populations, treatments, and outcomes. In the meta-analysis of trials of infertility treatments in women with the polycystic ovary syndrome,8 most of the trials that showed improved rates of ovulation or pregnancy with combination therapy included only women with clomiphene resistance. Moreover, intention-to-treat analyses were seldom used, and rates of ovulation or pregnancy were generally used as end points. Focusing on live-birth rates, the end point used by Legro et al., is preferable in view of potential discrepancies between rates of ovulation and pregnancy and between rates of pregnancy and live births.

Some aspects of the study by Legro et al. warrant comment. The diagnosis of the polycystic ovary syndrome required both hyperandrogenemia and anovulation, which is consistent with the usual research criteria, but the criteria for hyperandrogenemia varied among sites. In addition, inclusion criteria were relatively nonrestrictive; for example, women were eligible for the study if they had only one patent fallopian tube, if their partner had a sperm-count threshold above 20 million per milliliter (regardless of motility and morphology), and if they had previously received medication to induce ovulation. These factors may have led to some heterogeneity in the subject population and perhaps contributed to the relatively low pregnancy rate per treatment cycle. However, the study’s findings are compelling in view of the number of subjects, multicenter recruitment, diverse subject population, consistency of group differences over a broad range of body-mass indexes, use of a live-birth end point, and overall methodologic rigor.

It is increasingly recognized that women with the polycystic ovary syndrome are at increased risk for the metabolic syndrome and associated health risks13 and that metformin may well be important in treating these metabolic disturbances.14 However, the study by Legro et al. provides strong evidence that the metabolic benefits of metformin do not translate into live-birth rates that are as high as those with the old-fashioned standard, clomiphene citrate. Aside from the low but ever-present risk of multiple pregnancy, the use of clomiphene citrate to treat infertility in women with the polycystic ovary syndrome is simple, inexpensive, generally safe, and — as demonstrated by Legro et al. — more efficacious than the use of metformin.

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

From the Department of Obstetrics and Gynecology, University of Rochester School of Medicine and Dentistry, Rochester, NY.


Copyright © 2007 Massachusetts Medical Society.
Regulatory and Judicial Oversight of Nonprofit Hospitals

David M. Studdert, LL.B., Sc.D., Michelle M. Mello, J.D., Ph.D., Christopher M. Jedrey, J.D., Ph.D., and Troyen A. Brennan, M.D., J.D.

The modern hospital bears little resemblance to its ancestors. The charitable institutions of the 19th century mainly tended, rather than treated, the sick, and they served mostly poor patients, whereas the wealthy received care at home. The transformation of hospitals “from places of dreaded impurity and exiled human wreckage into awesome citadels of science and bureaucratic order” occurred during the 20th century, thanks to scientific advances and the maturation of the medical profession and the health insurance industry. Hospitals today are big businesses that derive most of their revenues from paying patients and health care insurers.

Yet one vestige of the ancestral institution remains: nearly two thirds of hospitals are private, nonprofit organizations. In return for exemptions from federal, state, and local taxes, they agree to organize and operate for charitable purposes — stricures that do not apply to their for-profit competitors. In a highly competitive marketplace in which profit margins are narrow, this commitment has generated intense economic, legal, and moral pressures.

Recent litigation illuminates the identity crisis of nonprofit hospitals. More than 100 lawsuits have been filed accusing them of shirking their charitable commitments by charging uninsured patients high fees and then aggressively pursuing these “debts.” The plaintiffs allege that the hospitals have broken their covenant with the community and morphed into profit-seeking businesses. Similar claims have been made by the Internal Revenue Service (IRS) and state attorneys general.

Broader societal expectations of nonprofit hospitals are also growing. As the dominant institutional player in health care delivery, such hospitals naturally are the focus of attention for solutions to the burgeoning problems of runaway costs, unequal access to care, and suboptimal quality. Regulators seek to steer the “nonprofits” toward these policy objectives and to press for loyalty to their charitable mission through a welter of federal, state, and municipal regulations. Meanwhile, nonprofit hospitals must contend with commercial pressures, primarily in the form of sharp competition from for-profit entities.

With these tensions assuming increasing prominence in health policy discussions, it is helpful to examine the legal environment in which nonprofit hospitals operate. What we find is a great deal of regulatory ferment, no clear direction, and few signs that oversight will become more coherent in the foreseeable future.

**The Role of the IRS**

Section 501(c)(3)

The federal government has long overseen nonprofit organizations, including hospitals, through tax laws defining the requirements for nonprofit status. Section 501(c)(3) of the Internal Revenue Code has formed the cornerstone of this regulation since 1954. The provision sets forth a series of requirements for nonprofit status. The organization must pass an “organizational test,” which refers to statements of religious, charitable, scientific, or educational purpose in publicly filed articles of incorporation. Next, an “operational test” requires that the organization actually operate in accordance with the specified purposes. Nonprofits may not distribute earnings to private persons in the way that for-profit firms do to stockholders, and they must pay fair market value for goods and services. Finally, the organization must adhere to reasonable public-policy norms and avoid activities such as political lobbying and engagement in political campaigns.

The operational test applies not only to the hospital but also to its key constituents — the medical staff. The IRS regards physicians as private persons who may benefit inappropriately
from hospital operations. The hospital often sees physicians as potential or actual competitors, particularly for ambulatory services; however, hospitals are generally restricted by IRS policies and antikickback and self-referral laws from entering into joint ventures with their physicians.

The sweeping language of the organizational and operational tests does not provide clear guidance to a nonprofit hospital board seeking to navigate a path between meeting charitable objectives and making effective business decisions. Federal policies aimed at enhancing competition in the hospital sector further blur the regulatory signals. In addition, the board’s task is complicated by ambiguity about which, if any, charitable activities produce the requisite social benefit to justify privileged tax treatment. Indeed, the evidence that nonprofit hospitals return more to the communities in which they operate than do for-profit hospitals is mixed. Increasingly, from the perspective of policymakers, the mere existence of nonprofit status carries no presumption of social benefit.

**INTERPRETING AND ENFORCING SECTION 501(c)(3)**

The IRS has issued several statements explaining the requirements of tax-exempt status for hospitals, focusing especially on what charitable care means. In 1956, the agency promulgated the first formal criteria for hospitals, which required that the hospital be “operated to the extent of its financial ability” to render services for those unable to pay. Practical difficulties in applying this standard, lobbying from the nonprofit hospital sector, and unrealistic hope that the advent of Medicare and Medicaid would reduce the need for charitable care led the IRS to refocus in 1969 on the hospital’s capacity to provide a “community benefit.” New standards specified some practices that demonstrate such a benefit — a board of directors drawn from the local community; an emergency room open to all, including indigent patients; the provision of other services without regard to the ability to pay or the source of insurance; and other activities that promoted health, including teaching, research, and community outreach.

The IRS soon confronted questions, many of them from the hospitals themselves, about the appropriateness of the requirement of an emergency department, particularly as it applied to specialty hospitals. In 1983, the IRS issued a revenue ruling — an official interpretation of how the tax law applies to a specific situation — clarifying that specialty hospitals generally do not need to operate emergency departments and that other hospitals may be excused from this requirement if a state planning agency has deemed such services unnecessary. The ruling created flexibility for some hospitals seeking to obtain or maintain tax-exempt status, although the presence of an emergency department still carries weight in determinations of the tax-exempt status of general hospitals.

These IRS rulings form the main guideposts in federal nonprofit hospital law. Unfortunately, their lack of specificity and their origin in a dramatically different health care environment have made the precise contours of what is permissible behavior today quite uncertain. Joint ventures between nonprofit hospitals and for-profit entities are a particularly gray area.

Historically, regulatory problems also extended to the enforcement tools available. The IRS was limited to two options: it could permit the conduct under scrutiny or revoke the hospital’s tax-exempt status. The severity of the latter option has tended to discourage its use. This problem of enforcement was ameliorated in 1996 by the creation of “intermediate sanctions,” which allow the IRS to impose financial penalties on influential insiders, such as board members and managers, who frustrate the charitable purpose by reaping impermissible private gains or “excess benefits” from the organization. To the best of our knowledge, intermediate sanctions have been applied only once in health care. Nonetheless, they represent a much more readily deployable enforcement mechanism than outright revocation.

**A BLUNT REGULATORY TOOL**

Alternative enforcement mechanisms help, but the IRS still struggles with many questions arising from the increasingly elaborate and innovative arrangements that nonprofit health care systems devise. Consider the 1998 ruling pertaining to a Maryland-based company referred to in the ruling as OMEGA. The company operated several nonprofit subsidiaries, including a hospital, a rehabilitation facility, a medical research institute, fundraising entities, and several long-term care facilities, all of which cared for patients with neurologic disease and brain injury. The company proposed to convert the hospital and nursing
homes into for-profit organizations and to main-
tain the holding company and research institute as nonprofits. Surprisingly, the IRS ruled that the reorganization would not imperil the tax-exempt status of either OMEGA or its remaining non-
profit affiliates.

By contrast, St. David’s Health Care System, a nonprofit based in Texas, formed a joint venture with Healthcare Corporation of America (HCA), a for-profit hospital chain, in which the two organizations were to pool all their hospital assets in the Austin area. The partners agreed that the new hospital conglomerate, including the HCA’s hospitals, would honor the community-benefit standard. Furthermore, the agreement included an escape clause giving St. David’s a unilateral right to terminate the partnership and liquidate its holdings at any time if it determined that the arrangement had begun to jeopardize or compromise the health care system’s charitable duties. After reviewing the joint venture, the IRS revoked St. David’s tax-exempt status. The Fifth Circuit Court of Appeals backed the position of the IRS, questioning both the charitable purposes of the partnership and the protection provided by the escape clause. (However, on remand to a lower court for a decision on the facts, a federal jury found that the escape clause did vest sufficient power in St. David’s to ensure the furtherance of charitable purposes.)

Squaring the outcomes in the St. David’s and OMEGA cases is not easy. Although legal experts may point to subtle distinctions, the broad reality is that the IRS must make fine determinations using a blunt tool — the text of section 501(c)(3) and the modest set of regulations and rulings that apply in the context of hospitals. The confused, sometimes paralyzing nature of federal tax policy and enforcement during the past decade has emboldened state policymakers, who also play a role in the regulation of nonprofit hospitals.

**State Regulatory Initiatives**

Most states regulate nonprofit organizations through statutes that resemble section 501(c)(3). A number of states have adopted the Revised Model Nonprofit Corporation Act of 1987, a comprehensive set of statutes governing the establishment and operation of nonprofit entities. State courts adjudicate disputes between local and state tax authorities and charities, including hospitals. With the deepening of academic concern regarding the extent to which nonprofit hospitals return social benefits commensurate with their preferential tax treatment, state courts have shown increasing interest in scrutinizing such hospitals’ operations, especially the provision of charitable care. The results of the scrutiny are mixed, compounding the uncertainty arising from the federal tax rulings.

In one case, for example, the Utah County tax board declined to exempt a nonprofit hospital from property taxes, ruling that the hospital had failed to deliver sufficient amounts of charitable care. The Utah Supreme Court upheld the board’s decision. Ruling on a similar case several years later, the Vermont Supreme Court ignored the Utah court’s analysis and reached an opposite conclusion. It rejected arguments by the city of Burlington that a local hospital had operated contrary to its charitable responsibilities under state law by remunerating executives inappropriately and providing too little free care. Although these two decisions come from different jurisdictions, they illustrate the difficulty of discerning any pattern in the way state courts have viewed questions about nonprofit hospitals’ compliance with their charitable purposes. Some of the haphazardness can be traced to genuine state-to-state differences in the applicable laws, but much of it reflects variation in fundamental judicial expectations about how nonprofit hospitals should behave.

State attorneys general have also jumped into the fray, using tactics that range from exhortation to litigation. Their involvement appears to have been spurred by several developments: the spectacular collapse of the Allegheny Health, Education, and Research Foundation in 1998, which became the nation’s largest failure by a nonprofit; the explosion in the 1990s of “conversions” (transfers of nonprofit assets to for-profit entities that generate substantial capital and raise questions about insider gains and lost community benefits); and governance failures in the corporate world more broadly, which have led to demands for improved oversight.

Attorneys general and other state agencies have sought to use their authority to bring nonprofit organizations under close supervision. These interventions appear to have been fueled by a perception that the boards of nonprofit orga-
nizations can be lax and insufficiently committed to furthering public-policy goals. However, the involvement of attorneys general varies considerably from state to state, ranging from very active (in New York and Minnesota) to moderately active (in California and Massachusetts) to minimal (in many states in the South and the non-coastal West).

**CORPORATE AND TRUST LAW**

The two main legal bases for state intervention are corporate law and trust law. Corporate fiduciary standards, which have been widely integrated into state nonprofit statutes, demand loyalty (directors must put the corporation’s interests before their own) and reasonable care (to the standard of an ordinarily prudent person). The “business judgment” rule creates a legal presumption that directors have complied with the duty of care in making their decision, unless the plaintiff can show that their actions constitute more than ordinary negligence, which is difficult to do. In general, the application of a corporate legal framework tends to give board members of nonprofit organizations substantial latitude in decision making.

Trust law, an alternative legal framework, focuses on protection of the organization’s assets, rather than directly on the actions of its trustees. State attorneys general have the authority to enforce charitable trusts through the courts as representatives of the beneficiaries, and some have been able to use this approach to achieve more stringent enforcement than corporate law permits. The central assertion in such interventions is that a charity is not being operated in a manner consistent with its original mission.

More than 30 years ago, an influential federal court decision held that the trend was to apply corporate, rather than trust, principles to directors of nonprofit organizations, reasoning that their duties were virtually indistinguishable from those of their corporate cousins. However, commentators have argued that the ability to ensure accountability to the charitable mission suffers when oversight emulates conventional corporate governance. Because nonprofits lack the guiding economic discipline introduced by shareholders and capital markets, corporate-law principles may bestow too much discretion on nonprofit boards.

A recent example of the application of trust-law principles to nonprofit hospitals is the case of Banner Health System. The company, which operated hospitals and nursing homes in a number of Western states, decided to sell 27 of its facilities in seven states and to apply the proceeds of the sale toward expanding operations in the “high-growth markets” of Colorado and Arizona. Anticipating resistance, Banner Health litigated to stop the attorneys general in South Dakota, New Mexico, and North Dakota from blocking transfers of funds out of state. In New Mexico, the case was settled. The North Dakota federal court refused to rule on the matter because the case did not raise issues of federal law. However, the South Dakota Supreme Court ruled against Banner Health, holding that an implied (or “constructive”) trust could be imposed to thwart the company’s plans. The result was that the states successfully used trust doctrine to wring concessions from the charity.

Regardless of the regulatory approach used, Banner and other cases illustrate an emerging theme in state regulation of nonprofit hospitals. State officials will not necessarily defer to nonprofit hospital boards or even give them the benefit of the doubt as the boards chart the future business directions of their organizations. Rather, appealing to the welfare implications of these decisions for the surrounding communities, state officials are using trust-law and corporate-law arguments to interpose their view of what constitutes inappropriate or appropriate action in light of an organization’s charitable purpose. Similar moves have been made to exert control over the behavior of nonprofit health insurance plans.

**CLASS-ACTION LITIGATION**

In the late 1960s, the “war on poverty” spilled over into the hospital arena. Social activists launched a class-action lawsuit on behalf of persons who had been turned away from tax-exempt hospitals because of their inability to pay for care. The plaintiffs’ ultimate target was the community-benefit standard, which the plaintiffs saw as an unacceptable dilution of the obligations imposed on nonprofit hospitals. The litigation went all the way to the U.S. Supreme Court, which ruled in *Eastern Kentucky Welfare Rights Organization v. Simon* that the plaintiffs lacked standing to sue privately to enforce the Internal Revenue Code.
After the defeat in Simon, this type of consumer activism went into hibernation. It has been reawakened during the past few years by a combination of two forces. The first of these forces is the emergence of techniques of mass tort litigation. The mass tort or class-action litigation strategy rests on several procedural devices that permit the consolidation of large numbers of claims alleging similar injuries and raising common questions of law and fact. By rolling dozens to thousands of plaintiffs’ damages into a single award, class-action lawsuits make it cost-effective to bring claims that would otherwise be too small to interest plaintiff’s attorneys.

The second force is a strong push during the past two decades toward the unionization of hospital workers, perhaps most prominently the push by the Service Employees International Union (SEIU).

Hospital unions are becoming politically astute and aggressive. Their leaders and members may gain an advantage from characterizing the management of large nonprofit hospitals as profit-oriented. Generating publicity about a hospital’s alleged failure to care adequately for the uninsured or, worse, its aggressive pursuit of payment from the poor can strengthen the union’s hand in workplace negotiations.

Building on the momentum of these forces, plaintiff’s attorneys have launched scores of lawsuits against nonprofit hospitals. They allege that the hospitals have shirked their community responsibilities, particularly in failing to provide free care to poor and uninsured patients. In 2003, two lawsuits (with roots in the SEIU), one in Illinois and the other in Connecticut, field-tested the strategy. Soon afterward, Richard Scruggs, a prominent Mississippi lawyer and a veteran of successful class-action litigation against the tobacco and asbestos industries, filed “at least 49 federal class action lawsuits charging approximately 370 nonprofit hospitals in 25 states with mistreating uninsured patients by, among other things, failing to provide adequate charity care.”

The New York case of Kolari v. New York–Presbyterian Hospital illustrates the way much of the litigation has unfolded. The four named plaintiffs were treated at New York–Presbyterian Hospital/Weill Cornell Medical Center. Although all of them were uninsured, they were billed higher fees than were insured patients, subjected to aggressive debt-collection efforts, and in the case of Kolari himself, denied some follow-up services.

The affected class consisted of employed poor persons and certain immigrants who did not qualify for Medicaid or for the hospital’s charitable care program yet could not afford the charges. The suit was based on several legal theories, the most prominent of which held that section 501(c)(3) created a contract between the government and the defendants in which the plaintiffs were intended to be third-party beneficiaries.

The court emphatically rejected the claim. Relying on the Simon decision from 30 years earlier, it found that the plaintiffs lacked standing to sue in federal court. The judge’s opinion advised the plaintiffs to “consult a map or a compass or a Constitution because Plaintiffs have come to the judicial branch for relief that may only be granted by the legislative branch.” In addition, the court found nothing in the Internal Revenue Code to support this type of claim, dismissed the other federal-law claims, and rejected various state-law claims brought under fraud and charitable-trust doctrines.

Nearly all federal and state courts that have ruled on class-action lawsuits against nonprofit hospitals to date have reached the same decision as the Kolari court. Although the prospects of success in federal court now appear to be very slim, plaintiffs continue to pursue class-action claims in state courts across the country. More than 100 such claims had been filed by the end of 2005, and there are signs that some of these cases may be more successful than their federal counterparts.

Most notably, Catholic Healthcare West and Sutter Health, two of the largest nonprofit hospital systems in California, have both settled class-action lawsuits, refunding fees charged to hundreds of thousands of uninsured patients. Arguments that nonprofit hospitals engaged in unfair trade practices appear to be faring better than those alleging breaches of tax and contract law.

Although the number and influence of dismissals easily outweigh those of the plaintiffs’ victories at this stage, the situation could change. The history of class-action litigation suggests that even a few early settlements can influence the future course of the litigation. Indeed, negative publicity or perceived legal risk may be enough to prompt hospitals to change their approach to charitable care. It seems likely that many hospitals today are rethinking their discounting and collection policies.
The litigation has clearly renewed attention to the definition of community benefit and the role of charitable care. The American Hospital Association (AHA) recently promulgated new policies on each, providing what it hopes will be a standard definition of community benefit for reporting to the IRS and suggesting that charitable care involve free care for those whose income is below 100% of the poverty line and graduated discounts for those whose income is between 100% and 200% of the poverty line. However, the AHA has also sought legislation to protect hospitals that comply with this definition from class-action lawsuits. The proposed definition of community benefit is broader than that sought by other hospital associations, such as the Catholic Health Association of the United States, and does not seem likely to deter the stepped-up IRS oversight of hospitals on which Congress is now insisting.

**CONCLUSIONS**

The laws that govern the conduct of nonprofit hospitals are complex. Today more than ever, multiple regulators are seeking to enforce standards of charitable purpose, and the rules are constantly evolving. Adding to the complexity is a range of other laws we have not delved into — antitrust, fraud and abuse, and local environmental and land-use regulations.

To some extent, this convoluted evolution is inevitable; it reflects unresolved and contested policy questions that go to the heart of the nonprofit model. How much and what sort of services should the nonprofits deliver to justify their preferential tax treatment? Should they be thrust fully into head-to-head competition with for-profit hospitals or spared that? If the latter, what trade-offs would society confront in the quality, availability, and efficiency of health care services?

However, uncertain the regulatory environment may be, it does not appear to be driving away preferences for this ownership model. The future of nonprofit hospitals seems assured, though formidable challenges lie ahead. These hospitals face cost increases that will outpace reimbursements, particularly from public payers. They also face growing competition from medical staffs interested in tapping into profits from procedure-based care that can now be performed in an outpatient setting. The struggle of nonprofit hospitals to maintain unprofitable services that are needed by the community, such as emergency care, at the same time that more profitable services are lost to their own medical staffs will probably intensify.

These challenges press nonprofit hospitals to develop increasingly aggressive business models. They will seek to avoid bad debt and to attract patients who are well insured. They will also be increasingly interested in joint ventures with medical groups to provide physicians with the profit they might otherwise seek through for-profit ambulatory care and surgical centers. The days when nonprofit hospitals relied on charitable giving and were led by volunteer board members with little understanding of health care markets are gone. Indeed, one might argue that leaders of nonprofit hospitals are failing to act diligently if they do not test the boundaries of legal permissibility by considering every technique that their for-profit competitors use.

As nonprofit hospitals strike out in these directions, federal regulators, state officials, and plaintiffs will police the resultant frictions between the hospitals’ business practices and their charitable obligations. A recent investigation by the U.S. Senate and an ongoing IRS audit of hundreds of nonprofit hospitals provide a taste of things to come. To the hospital board falls the unenviable task of demonstrating and articulating obedience to a charitable purpose in an increasingly harsh commercial environment.

Dr. Brennan is the former president and former chief executive officer of the Brigham and Women’s Physician Organization at the Brigham and Women’s Hospital, a nonprofit hospital in Boston. He is currently chief medical officer at Aetna. Dr. Jedrey and the law firm in which he is a partner provide legal services to Aetna, Hartford, CT (T.A.B.).

From the Harvard School of Public Health (D.M.S., M.M.M.) and McDermott Will & Emory (C.M.J.) — both in Boston; and Aetna, Hartford, CT (T.A.B.).

8. Horwitz JR. Why we need the independent sector: the behav-
42. Littauer Hospital Association v. Spitzer, 287 A.D.2d 202 (3d Dep’t 2001).
Colonoscopy Screening for Detection of Advanced Neoplasia

TO THE EDITOR: In the cross-sectional study of the detection of advanced neoplasia during colorectal-cancer screening, Regula et al. (Nov. 2 issue) observed that Polish men were more than twice as likely as Polish women to have advanced neoplasia. Although they controlled for age and family history, the authors did not examine lifestyle factors known to be associated with colorectal neoplasia.

Clinical studies have consistently shown that current and former smokers have significantly higher rates of colorectal neoplasia than nonsmokers. Results from screening populations have shown that smoking may be at least as potent a risk factor as family history of colorectal cancer.

In Poland, men were nearly twice as likely as women to smoke cigarettes during most of the last decade. In addition, a study of Polish patients with lung cancer showed that men had more pack-years of smoking than women did. We believe that this sex difference in tobacco exposure, coupled with the importance of smoking as a risk factor for colorectal neoplasia, warrants its consideration as a confounding factor that may explain the sex disparity observed in this trial.

Michael W. Latreille, B.S.
Joseph C. Anderson, M.D.
State University of New York at Stony Brook
Stony Brook, NY 11794-8173
jcanerson@notes.cc.sunysb.edu


TO THE EDITOR: Regula et al. found that the number of men 40 to 49 years of age who needed to be screened to detect one advanced colorectal neoplasm (the number needed to screen) was similar to that of women 50 to 54 years of age. This finding is reasonable, since the ratio of the incidence of colorectal cancer for men to that for women in the 40-to-49-year age group in Warsaw is 1.4.

The authors suggest modifying the current screening recommendations for colorectal cancer, perhaps by using their “numbers needed to screen” data. We believe a more comprehensive approach, with additional risk factors included, should be taken. Tobacco and alcohol use are recognized risk factors for colorectal adenomatous polyps.
and cancer\textsuperscript{2,3} and could explain the higher prevalence of neoplasia observed in men. Giovannucci et al. previously suggested lowering the age threshold for colorectal screening among long-term smokers.\textsuperscript{2}

Prognostic models have been developed to assess the individual risk of certain forms of cancer.\textsuperscript{4} A similar risk model for colorectal cancer, which could include information on risk factors such as age, sex, family history, body-mass index, diet, smoking status, alcohol consumption, and previous bowel disease, would be a useful tool for defining the ideal target population for colorectal-cancer screening.

Patrick Maisonneuve, Eng.
European Institute of Oncology
20141 Milan, Italy
patrick.maisonneuve@ieo.it

Albert B. Lowenfels, M.D.
New York Medical College
Valhalla, NY 10595

To the Editor: Regula et al. speculate about why their study yielded a lower-than-expected prevalence of advanced neoplasia. Perhaps the explanation is that only 29.8\% of their patients received intravenous sedation. A colonoscopy in an unsedated patient is not a pleasant experience. An unsedated patient would certainly be less likely to tolerate a complete examination; this is probably reflected in the lower-than-expected cecal intubation rate in their study. In the United States, intravenous sedation is now routinely used, and one should therefore exercise great caution in extrapolating the findings of this study to the United States, as well as to other countries where sedation is the norm.

Stephen M. Picca, M.D.
Nassau University Medical Center
East Meadow, NY 11554
eesp@aol.com

\textbf{The Authors Reply:} Latreille and Anderson suggest that the higher rate of detection of advanced neoplasia in men than in women during colonoscopic screening in our study was due to sex differences in tobacco exposure. This may be true — it has been suggested that smoking is related to colorectal adenomatous polyps and cancer, although the data are not sufficient to call the relationship causal.\textsuperscript{1} Furthermore, smoking may be associated with lifestyle patterns that may be linked to colorectal neoplasia, (i.e., low fiber intake, high alcohol and fat intake, and low levels of physical activity). There may also be other factors that decrease the risk in women, including the presence of sex hormones. Our study was not designed to investigate the influence of smoking or other lifestyle factors, and we can neither reject nor confirm the suggestion of Latreille and Anderson. However, the fact remains that the numbers needed to screen to detect advanced neoplasia were significantly lower for men than for women. A similar finding was reported in studies conducted in the United States.\textsuperscript{2,3}

We agree with Maisonneuve and Lowenfels that there are more complex prognostic models to assess the individual risk of cancer, but in our opinion, they are not ideal to guide practical recommendations for mass screening, because reliable data on factors such as smoking and drinking habits and especially diet are difficult to obtain.

Picca suggests that only colonoscopy with sedation is an acceptable screening tool and links the lower-than-expected prevalence of advanced adenoma in our study to the fact that the majority of screened patients (70.2\%) underwent examination without sedation. However, the cecal intubation rate, an important measure of the quality of colonoscopy, was 91.1\%, which is only slightly lower than the expected rate of 95\% for an expert colonoscopist. Although this might have affected the rate of adenoma detection, the effect could not have been large. On the other hand, screening on a mass scale cannot be limited to expert centers. Other factors have an established effect on the rate of adenoma detection, including the colonoscopy withdrawal time, the quality of bowel preparation, and the adequacy of training.\textsuperscript{4,5}

Sedation is helpful mainly during colonic intubation; its role during withdrawal of the colonoscope, a phase vital for adenoma detection, is certainly much less important. We are not aware of any study indicating that the rate of adenoma...
detection in the colonic segments examined can be influenced by sedation, if the cecal intubation rate remains the same.

Jaroslaw Regula, M.D.
Marcin Polkowski, M.D.
Eugeniusz Butruk, M.D.
Maria Sklodowska-Curie Memorial Cancer Center
02-784 Warsaw, Poland
jregula@coi.waw.pl


TO THE EDITOR: The article by Brennan et al. (Nov. 8 issue), comparing rabbit antithymocyte globulin and basiliximab as induction therapy in patients who received a renal transplant from a deceased donor, states that the patients included in the trial were at high risk for acute rejection. The data provided are insufficient to confirm that the subjects were at high immunologic risk, classically defined as a peak panel-reactive antibody value of more than 50% or a current value of more than 30%, the loss of a previous graft through rejection within 12 months after transplantation, the presence of a donor-specific antibody (a positive T-cell or B-cell crossmatch), or a high degree of HLA mismatching. Recent studies have used such criteria to define high immunologic risk for renal allograft recipients. In the study by Brennan et al., the mean peak panel-reactive antibody value was about 14% in both groups, with a mean value of about 6% at the time of transplantation. No data were provided on the frequency of previous graft loss due to rejection, the frequency of a positive crossmatch, or the degree of HLA mismatching. The provision of such data would enable clinicians to determine the applicability of the results to their own patient populations.

John P. Killen, M.B., B.S.
Steven Chadban, M.B., B.S.
Royal Prince Alfred Hospital
Sydney 2050, Australia
ojpo@hotmail.com

DR. BRENNAN AND COLLEAGUES REPLY: Our study enrolled subjects at risk for delayed graft function or acute rejection. Since delayed graft function is an independent risk factor for acute rejection, recipients at low risk for delayed graft function were eligible if they had at least one recipient risk factor for rejection. All recipients studied had donor risk factors for delayed graft function, placing them at risk for rejection. A historical peak panel-reactive antibody value of more than 20% was present in 18% of recipients, and a current panel-reactive antibody value of more than 50% was present in 5% of recipients, with no significant differences between the two groups. Reasons for the loss of a previous transplant and the presence of donor-specific antibodies were not recorded. Seven patients received a transplant with a positive B-cell crossmatch, and three also had a positive T-cell crossmatch. Three percent had six HLA mismatches. We believe the study enrolled patients who were at high risk for the period when the study was open. Approximately 50% of kidney transplants from deceased donors in the United States would qualify for a study of this design.
today, according to data from the United Network for Organ Sharing.

Daniel C. Brennan, M.D.
Washington University School of Medicine
St. Louis, MO 63110
dbrennan@wustl.edu

Paula Buchanan, M.P.H.
Mark A. Schnitzler, Ph.D.
Saint Louis University Center for Outcomes Research
St. Louis, MO 63104

Dr. Buchanan reports receiving a grant from the American Society of Transplantation. Dr. Schnitzler reports receiving consulting fees from Novartis Pharma, lecture fees from Genzyme, and grant support from Genzyme, Novartis Pharma, Astellas, and TransMedics.


DHEA and Testosterone in the Elderly

TO THE EDITOR: In their report on the effects of dehydroepiandrosterone (DHEA) and testosterone when used as antiaging supplements, Nair et al. (Oct. 19 issue)1 conclude that low-dose testosterone replacement in elderly men has no “physiologically relevant beneficial effects on body composition, physical performance, [or] insulin sensitivity.” However, this conclusion is premature, since the testosterone replacement administered failed to achieve physiologic testosterone levels throughout the study period (Fig. 2 of the article). Moreover, despite the marginal increase in testosterone levels achieved, improvements in fat-free mass, fasting insulin levels, and bone mineral density were observed.

Other studies of testosterone replacement, including those cited to support the authors’ conclusions,2 have shown a decrease in fat mass (12.5%) and an increase in lean mass (4%) when physiologic testosterone levels are achieved in elderly men. Studies of standard doses of testosterone in the treatment of testicular failure3 have shown additional positive effects on muscle strength, physical performance,4 and bone mineral density.5 Large, long-term trials are clearly needed to assess the risks and benefits of testosterone replacement in elderly men, and caution should be exercised regarding the treatment of andropause in men. However, the serum testosterone level achieved should be within the normal range to assess the effect on outcome measures adequately.

Stephanie T. Page, M.D., Ph.D.
University of Washington
Seattle, WA 98195
page@u.washington.edu

TO THE EDITOR: The findings of Nair et al. cannot be generalized, because the study included relatively healthy subjects. To investigate the benefits and risks of androgen-replacement therapy, it is essential to make judicious choices regarding the subjects to be included in the research. In this study, the average baseline scores for the quality of life (on the Health Status Questionnaire [HSQ] and the Medical Outcomes Study 36-item Short-Form General Health Survey [SF-36]) of all the subjects were above 50 for both the physical and mental components. The average score on both instruments in the general U.S. population is 50.1 The high scores of these subjects suggest that the

Alvin M. Matsumoto, M.D.
Veterans Affairs Puget Sound Health Care System
Seattle, WA 98108

William J. Bremner, M.D., Ph.D.
University of Washington
Seattle, WA 98195

study included healthier elderly persons than those who would be representative of the general elderly population.

Moreover, physical exercise is expected to improve and maintain physical functioning in older people. Not only androgen administration but also well-designed physical training is needed to improve the physical performance of elderly persons. The androgen level might be a mediator that could be elevated by exercise training, which would then increase physical performance. The administration of androgen in the absence of exercise may not be enough to improve physical performance among the elderly.

Mitsuko Yasuda, M.D., Ph.D.
Shigeo Horie, M.D.
Teikyo University
Tokyo 173-8605, Japan
shorie@med.teikyo-u.ac.jp


TO THE EDITOR: DHEA was banned in 1985 by the Food and Drug Administration because clinical safety and efficacy data were lacking to support claims of cures for cardiovascular disease and aging. After the passage of the Dietary Supplement Health and Education Act in 1994, DHEA, which had not previously been labeled as a drug, again became available. It is amazing that a previously banned substance can now be sold directly to the public, and it speaks to the lack of oversight and protection afforded by the Dietary Supplement Health and Education Act.

Hormones have long been equated with youth by the public and are thus a favorite type of substance for marketing by the antiaging industry. As one substance falls out of favor, another quickly replaces it: the miracle of melatonin was replaced by the superhormone promise of DHEA. The heir apparent now seems to be growth hormone, which, paradoxically, is illegal to distribute for antiaging uses but constitutes a market estimated at more than $600 million per year in the United States alone.

Thomas T. Perls, M.D.
Boston University Medical Center
Boston, MA 02118
thperls@bu.edu

Drs. Gleicher and Barad are part owners of a pending patent involving the use of DHEA for the improvement of ovarian function in women with diminished ovarian function.


THE AUTHORS REPLY: Our conclusion that “additional long-term studies of testosterone are warranted to determine the risk–benefit ratio of higher doses” is in agreement with the view expressed by Page et al. We showed that low-dose testosterone significantly increased testosterone levels but resulted in no physiologically relevant beneficial effects. However, unlike the investigation of DHEA in the study, the investigation of testosterone did not address the potential beneficial effects of replacement therapy that would increase the plasma levels of testosterone in older people to levels found in young people. Standard testosterone replacement in younger men with testicular failure has a profound effect on body composition, but the response to testosterone therapy in older men with low testosterone levels remains uncertain. Biweekly intramuscular injection of testosterone (at a dose of 200 mg) in older people has been shown to increase peak testosterone levels to values above the normal range in young people and improves physical performance. However, because of adverse events, the investigators had to reduce the dose in some subjects. Transdermal administration of testosterone maintained testosterone levels in older people to levels within the normal range for young people for a period of 36 months but had no effect on physical performance, despite a significant increase in fat-free mass. We agree with Page et al. that there are tantalizing data on the effect of testosterone on bone density. The findings in the studies cited here and in other studies highlight the importance of conducting long-term studies to document adverse events and long-term benefits of restoring testosterone levels in older people to levels seen in young people.

In response to Yasuda and Horie, the HSQ is a measure of perceived but not actual health. On the basis of the HSQ, we cannot state whether the persons in our sample were more or less healthy than the general population. Scores for both the physical and the mental components of the SF-36 questionnaire derived from the HSQ are only minimally different from a score of 50, and our interquartile range includes 50. We agree with Yasuda and Horie that the interaction between testosterone and exercise training in older people remains to be determined and warrants further investigation.

We agree with the comments by Perls about DHEA and the unfortunate rush to hormone therapies to maintain youth, even though no scientific data conclusively support such approaches. In response to Gleicher and Barad, our study specifically addressed whether the long-term administration of DHEA has any beneficial effects in women older than 60 years who have low levels of DHEA. The main outcome measures were objective physiological measurements and responses to the standard HSQ. We did not evaluate the effect of DHEA in women who have postmenopausal symptoms for which estrogen therapy is likely to be more effective than DHEA.

K. Sreekumaran Nair, M.D., Ph.D.
Glenn Smith, Ph.D.
Mayo Clinic
Rochester, MN 55905
nair.sree@mayo.edu


Spinal Epidural Abscess

TO THE EDITOR: In his review of spinal epidural abscess (Nov. 9 issue), Darouiche does not mention brucella as an agent. In Spain, Italy, and the Near East, brucella is encountered frequently, and its treatment poses specific problems.

J. Martinez L. de Letona, M.D.
Universidad Autónoma de Madrid
28035 Madrid, Spain
juan.martinez@uam.es


TO THE EDITOR: In Darouiche’s review of spinal epidural abscess, the algorithm on management (Fig. 4 of the article) does not include immediate magnetic resonance imaging (MRI) or computed tomographic (CT) myelography. Immediate imaging with timely subsequent surgical decompression probably offers the optimal long-term outcome. Although improvement has been reported when decompression is performed within 72 hours, in a single case series, there was no improvement with surgical management performed as soon as 6 to 12 hours after the onset of symptoms. Hence, as soon as there is a clinical suspicion of a neuraxial abscess, imaging studies and a surgical consultation should be obtained to minimize the neurologic deficit.

Neal Gerstein, M.D.
University of New Mexico
Albuquerque, NM 87106
ngerstein@salud.unm.edu


TO THE EDITOR: We do not agree with Darouiche that surgical drainage together with systemic antibiotics constitutes the treatment of choice for spinal epidural abscess in all patients except for those with a poor medical condition or complete paralysis for more than 36 hours. In fact, medical management with vigilant monitoring can be considered in patients without neurologic involvement. In 1999, we reported 75 cases of spinal epidural abscess, of which 22 (29%) were successfully treated conservatively with antibiotics alone and close clinical monitoring. Since then, similar results have been reported with conservative management in patients with no neurologic deficit. Recently, Savage and colleagues reported 29 cases of medically treated spinal epidural abscess, and in 24 cases (83%), the outcome was excellent or good. Spinal epidural abscess without neurologic involvement can be safely and efficiently managed with medical treatment as long as the patient is closely monitored. In the case of neurologic deterioration, prompt surgical decompression is necessary.

Daniele Rigamonti, M.D.
Philippe Metellus, M.D.
Johns Hopkins Hospital
Baltimore, MD 21287
dr@jhmi.edu


THE AUTHOR REPLIES: The three letters address valid issues regarding the cause, diagnosis, and management of spinal epidural abscess. Although it is relatively rare in North America, brucellosis, a zoonosis that is highly prevalent in the Mediterranean region, Arabian Peninsula, Near East, and Central and South America, can cause spinal infections, including spinal epidural abscess, osteomyelitis, and diskitis. As de Letona points out, the often elusive diagnosis of brucellosis and the need to treat it with a combination of antimicrobial agents underscore the unique challenges in managing a spinal epidural abscess that is caused by brucella.

Gerstein accurately articulates the importance of immediate MRI or CT myelography in patients with suspected spinal epidural abscess and the need for urgent subsequent surgical decompres-
sion. The algorithm in Figure 4 of my article, however, does not address imaging studies because it focuses on the management of spinal epidural abscesses.

Although Rigamonti and Metellus argue reasonably that spinal epidural abscess can be safely treated with antibiotic therapy alone in neurologically intact patients as long as they are closely monitored, a critical analysis of the literature provides support for this management principle in only some but not all such patients. Decompressive laminectomy together with systemic antibiotics is preferable to medical treatment alone because the rate of the evolution of neurologic impairment and progression to paralysis is unpredictable, with some patients becoming paralyzed within hours after the onset of the neurologic deficit. Moreover, patients treated with antibiotics alone can still become neurologically impaired, and the neurologic deficits may be irreversible; the preoperative neurologic status is the most important predictor of the final neurologic outcome.2-4

Rabih O. Darouiche, M.D.
Michael E. DeBakey Veterans Affairs Medical Center
Houston, TX 77030
rdarouiche@aol.com


Medical Education — Professionalism

TO THE EDITOR: It is gratifying to see the subject of medical professionalism considered in the review article by Stern and Papadakis (Oct. 26 issue). However, the article fails to meaningfully address the reality that physicians are increasingly employed by or dependent on organizations with a business ethic that is indifferent and occasionally hostile to the values and behaviors of professionalism. Medical practitioners are expected to placate profit-driven employers and insurance carriers, for example, while remaining loyal to the highest standards of medical professionalism. The educational imperative should be not only to teach the values of medical professionalism but also to provide practical instruction for their implementation.

Timothy Howland, M.D.
Lourdes Hospital
Binghamton, NY 13905
thowland@lourdes.com


TO THE EDITOR: Stern and Papadakis underscore the importance of instilling a sense of professionalism again in physicians and students, and they warn us about what medicine stands to lose if we do not do so. However, they leave out something crucially important that was identified by the sociologist Talcott Parsons more than 50 years ago: “The ‘ideology’ of the profession lays great emphasis on the obligation of the physician to put the ‘welfare of the patient’ above his personal interests, and regards ‘commercialism’ as the most serious and insidious evil with which it has to contend.” In 1995, George Lundberg, then editor-in-chief of the Journal of the American Medical Association, repeated this admonishment: “The fundamental purpose of a business is to make money. . . . On the other hand, the fundamental purpose of a profession is to provide a service that reflects commitment to a worthy cause that transcends self-interest.” Specialty hospitals, boutique care at a price, and a range of other practices threaten the core of trust on which our profession stands. We need to teach our students — and model for them in our own practices — that commercialism has no place in the profession of medicine.

Donald A. Barr, M.D., Ph.D.
Stanford University
Stanford, CA 94305
barr@stanford.edu

TO THE EDITOR: The article by Stern and Papadakis on becoming a medical professional rightly stresses the humanistic aspects of professionalism and ethical obligations. The outcome measures in this model of professionalism assess behaviors that reflect appropriate value systems. However, the definition of a professional also includes asymmetries of knowledge between the client and the professional.1 To satisfy this criterion of professionalism, the physician must deliberately reflect on potential knowledge gaps vis-à-vis the changing evidence base for practice. At the core of a “seasoned clinician” is the knowledge that will provide the best available options for the patient and a commitment to maintain this edge. A comprehensive and balanced program of professional development should encourage physicians to seek out this knowledge. Such a program would ensure that the professional will be in a position to practice safely and to advise patients about options for the best clinical outcomes on the basis of current knowledge. In this model of professionalism, outcome measures reflect the view that the maintenance of current clinical and scientific knowledge is paramount, but these measures are also integrated with previous experience and sustained by a value system that is in line with societal expectations.

John I. Balla, F.R.A.C.P.
Melbourne University Geelong 3220, Australia


TO THE EDITOR: In their otherwise informative article on professionalism, Stern and Papadakis suggest that “it is now possible to reliably predict interpersonal and communication skills with the use of multiple, brief standardized interpersonal interactions.” This could be exciting news for the admissions committees of medical schools. However, the authors cite only one reference regarding this topic.1 A careful review of this reference suggests that readers should be cautious in drawing definitive conclusions because of the lack of predictive validity data. Preliminary results of a single study are not a sufficient reason to incorporate the multiple medical interview into such an important decision as whether to admit a candidate to medical school.

Robert Knopp, M.D.
Regions Hospital
St. Paul, MN 55105
knopp003@umn.edu


THE AUTHORS REPLY: Howland and Barr reflect on the challenges that doctors face when confronted with the practice settings, health care systems, and commercial interests that test our professional resolve. It is here that we must not only aspire to the principles of professionalism but also wisely apply them. Medical practice cannot be insulated from an otherwise commercial world, and financial solvency is a necessity. Rather than condemn all forms of commercial interest, we encourage our students to engage in the development of regulatory policies and systems that benefit patients and are concordant with our values. Advocacy at the local, national, and international levels is critical to such engagement.

Professionalism is demonstrated through a foundation of clinical competence, communication skills, and ethical understanding. The aspiration to and wise application of the principles of professionalism — excellence, humanism, accountability, and altruism — are built on this foundation.2 Although our article focuses on these latter principles of professionalism, Balla rightly identifies the foundational element of knowledge without which a doctor would be a communicative, compassionate charlatan.

Selecting students with the aspiration for and ability to wisely apply the principles of professionalism is a long-standing challenge for medical educators. If, before they entered medical school, we could identify those most likely to behave professionally in most situations, we would be doing both the public and the profession a great service. As Knopp suggests, the multiple medical interview is a potentially groundbreaking innovation that deserves further investigation. A growing amount of literature on the multiple medical interview, including data on its validity,3 provides support for our contention that this assessment
method should have a role in the medical school–admissions process.

David T. Stern, M.D., Ph.D.
University of Michigan Medical School
Ann Arbor, MI 48109
dstern@umich.edu

Maxine Papadakis, M.D.
University of California, San Francisco
San Francisco, CA 94143


TO THE EDITOR: Thomsen et al. (Oct. 12 issue) state in their video that thoracentesis must be performed with “extreme care” in mechanically ventilated patients because of a theoretically increased risk of tension pneumothorax. They also state that chest radiographs should be routinely performed if “the patient is critically ill or receiving mechanical ventilation.”

Although care is appropriate in any invasive procedure, pneumothorax is a rare complication as long as ultrasonography is used. In a series of 232 mechanically ventilated patients, 3 patients (1.3%) had pneumothorax. None of these pneumothoraces were under tension, although a chest tube was inserted in all three cases.

This low rate of pneumothorax may indicate that postprocedure radiography is not routinely indicated after thoracentesis guided by ultrasonography. In addition, if pneumothorax is to be ruled out, ultrasound documentation of lung sliding may be superior to chest radiography.

Lewis A. Eisen, M.D.
Beth Israel Medical Center
New York, NY 10003
leisen@gmail.com


TO THE EDITOR: The video about thoracentesis by Thomsen et al. was well prepared and is educational for medical trainees unfamiliar with the procedure. However, it perpetuates a misconceived distinction between diagnostic and therapeutic thoracentesis.

After a patient has been subjected to the risks and discomfort of catheter insertion, there is no benefit of leaving fluid within the chest. The risks of bleeding or pneumothorax, once the needle has been removed, should be unchanged, regardless of the quantity of fluid aspirated. Withdrawing “diagnostic” quantities and leaving large residual effusions subjects the patient to continued dyspnea and future additional procedures. It is important to limit the risk of re-expansion pulmonary edema (by aspiration of <1.5 liters, according to the authors). Within that constraint, all accessible fluid should be removed.

This illogical and detrimental differentiation between diagnostic and therapeutic thoracentesis should be abolished. All thoracenteses should be both diagnostic and therapeutic.

Roy T. Temes, M.D., M.B.A.
Cleveland Clinic
Cleveland, OH 44109
temes@ccf.org

THE AUTHORS REPLY: We are hesitant to conclude that post-thoracentesis chest radiography should be omitted in critically ill or mechanically ventilated patients, even if ultrasonography is used. Although Mayo et al. reported a pneumothorax rate of 1.3%,1 Barnes et al. reported 15 pneumothoraces in 305 ultrasound-guided procedures (4.9%),2 and Gervais et al. reported 6 pneumothoraces in 90 ultrasound-guided procedures (6.7%) in mechanically ventilated patients.3 Pneumothorax may be considered rare, but it is prevalent enough to warrant consideration.

In the retrospective review of thoracentesis by Aleman et al., clinical symptoms such as cough, dyspnea, and pleuritic chest pain were the main indicators of the presence of pneumothorax.4 These findings may be missed (or absent) in criti-
Elderly Survivors with Homozygous Sickle Cell Disease

TO THE EDITOR: Prolonged survival in homozygous sickle cell disease is more common than previously thought. A Jamaican study in 1968 described 60 patients who were 30 years of age or older, and the Cooperative Study of Sickle Cell Disease in the United States estimated a median survival of 42 to 48 years.

The Sickle Cell Clinic at the University of the West Indies has treated 102 patients (64.7% women) who survived beyond their 60th birthday; of these patients, 58 are now dead, 4 have emigrated, and 40 are still alive. The diagnosis of homozygous sickle cell disease was based on criteria that excluded sickle cell–β0-thalassemia or sickle cell–hereditary persistence of fetal hemoglobin. None of the patients received hydroxyurea, and only two patients with renal impairment received regular transfusions. The ages of the patients ranged from 60.2 to 85.6 years, with a marginal excess of women (67.5%, P=0.01). Sickle cell–α+ thalassemia was homozygous in 12% of patients and heterogeneous in 39%, with a normal alpha-gene number in 46%; the trend failed to reach statistical significance (χ²=2.4, P=0.12). The Benin beta-globin haplotype was homozygous in 76% of the patients, and none had the Bantu, Senegal, or Asian haplotype.

Fetal hemoglobin levels were on average 4.9% higher among these patients than values extrapolated from the Jamaican Cohort Study of Sickle Cell Disease (P<0.001) and exceeded 10% in 16 women (24%) and in 5 men (14%), which suggests that higher fetal hemoglobin levels probably conferred protection in childhood. Family studies showed that 13 of 30 siblings of patients with homozygous sickle cell disease (43%) were over the age of 60 years, but their median survival (56.9 years) did not significantly exceed the median survival of 44.6 years in the Sickle Cell Clinic, as calculated by a standard statistical adjustment (P=0.07), or 56.3 years, as calculated by an adjustment for death rate (P=0.34).³

Hemoglobin values declined in both men and women, but more steeply in men (mean decrease, 2.6 g per deciliter) than in women (2.3 g per deciliter), and fell by 0.02 g per deciliter for every 10-unit increase in the creatinine level (P=0.001), or by 0.6 g per deciliter for every 1-unit increase.
in the natural-log–transformed creatinine level (Fig. 1). Renal impairment increased with age and affected 34 of 40 patients (85%), on the basis of a sensitive definition. An incidental presentation occurred in 16 patients (40%). Seven patients reported having had no bone-pain crises (18%); bone pain decreased with age in 30 of the remaining 33 patients; 58% of the patients had a history of leg ulcers. Pregnancy had a benign outcome, with 78 of 96 pregnancies (81%) resulting in live births. In these elderly survivors, high fetal hemoglobin levels and possible familial clustering were consistent with genetic factors. Some features, such as an incidental presentation and successful pregnancy outcomes, suggested an intrinsically mild course, whereas other features, such as severe bone pain, were ameliorated with age. In this group of patients, α-thalassemia did not promote survival, and the group was too homogeneous to show any effect of the beta-globin haplotype. The major clinical problems emerging with age were renal impairment and decreased levels of hemoglobin.

Graham R. Serjeant, M.D., F.R.C.P.
Sickle Cell Trust (Jamaica)
Kingston 6, Jamaica

Douglas R. Higgs, M.D., F.R.C.P.
Weatherall Institute of Molecular Medicine
Oxford OX3 9DS, England

Ian R. Hambleton, Ph.D.
University of the West Indies
Barbados, West Indies


Correspondence Copyright © 2007 Massachusetts Medical Society.

Figure 1. Most Recent Hemoglobin Level and Serum Creatinine Level (Natural-Log–Transformed) in 97 Elderly Patients.

The levels of hemoglobin are depicted as changes from steady-state values, as indicated by the dashed line.
MARC HAUSER’S GROUNDBREAKING BOOK advances a new theory of moral judgment, synthesizing a great deal of work in neuroscience, psychology, and ethology, as well as the author’s own recent experimental work. Hauser aims to demonstrate that morality is innate in the way that language is innate: not in its precise content, but in its form. Just as, according to Noam Chomsky, each child comes into the world with a brain wired for language acquisition, so is each of us born ready to acquire a moral system, Hauser argues. Just as the innate workings of the human mind tightly constrain the grammar of any possible human language, so do they constrain — without determining — the content of any possible moral system.

In Hauser’s view, the moral faculty is triggered by the perception of an action or the omission of an action and automatically produces a judgment that is sensitive to the action’s (or omission’s) causes and consequences and to whether the consequences were intended or foreseen. Actions are perceived to have more moral weight than omissions, and intended harms are seen to be morally worse than foreseen harms. Using his ongoing Web-based surveys, Hauser has amassed an enormous amount of evidence demonstrating that these distinctions are made in the same way in all cultures, across all educational levels, and by both sexes.

But subjects are typically unable to articulate adequate justifications for their judgments. Hauser takes this inability to be evidence for his view that we have a moral faculty that operates below the level of conscious awareness. Once again he points out the parallels to linguistic competence. Just as it is easy for us to judge whether a sentence is grammatical but difficult for us to justify or explain that judgment, so we are able to judge the permissibility of an action easily but find ourselves unable to explain that judgment.

Since our ability to make moral judgments outpaces our ability to justify them, moral judgment does not seem to be the product of rational reflection. Might it be the product of an emotional system instead? Hauser grants that moral judgments are typically accompanied by emotions, but he suggests that these emotions are the result of the moral judgment rather than the cause. Part of his evidence for this claim comes from studies of patients with damage to the ventromedial prefrontal cortex. These patients appear to have reduced emotional responses to moral harms, yet their judgments appear to be identical to those of normal subjects when confronted with most moral dilemmas. Hauser nonetheless believes that emotions powerfully influence moral performance — by influencing our motivation to act morally — but that the moral faculty itself is independent of the emotional system. Moral competence is the product of an innate moral faculty whose optional parameters and exceptions are determined by the culture into which each of us is born.

There is little doubt that this book, written for a general audience, is the most important attempt to date to explain the psychological mechanisms of moral judgments. However, Hauser has made the unusual decision to publish it well before all the experimental data are in. For this reason, the book is sometimes frustratingly vague on key questions. Is the moral faculty cognitively penetrable — that is, can a person gradually alter its parameters through reflection — or is it more like the visual system, which remains subject to visual illusions even when we know full well that they are illusions? What accounts for the differences in moral judgments among people who grow up in the same culture, a difference that has no obvious parallel in the linguistic sphere? It would be churlish to criticize Hauser for the lack of clarity and detail on such matters. Moral Minds
offers us the most important scientific contribution to moral psychology in many decades, and it is certain to inspire and inform debate across many fields for years to come.

Neil Levy, Ph.D.
University of Melbourne
3010 Parkville, Australia
nlevy@unimelb.edu.au

SURGICALLY SHAPING CHILDREN: TECHNOLOGY, ETHICS, AND THE PURSUIT OF NORMALITY


M ost people do not question the benefits of surgeries undertaken to give children a more normal appearance. Some recognize the tension inherent in making such decisions but nevertheless surrender to the social pressures of conforming to normalcy. They feel the need to do something, even when doing so may not be clearly advantageous for the child. They tend to yield to the dubious authority of the technological imperative. For these reasons, this compilation of essays edited by Erik Parens is vitally important.

In his introduction, Parens describes the book as an exploration of the tension between the desire to have surgery performed to spare children the pain and suffering of being different and the desire to spare children the pain and suffering of being subjected to surgery. But the book does much more. It explains the philosophical, psychological, and medical reasons why this tension exists, and it challenges the assumptions that embroil us in that tension. It provides an amazing wealth of practical advice that will help readers understand, confront, and manage the various forces that create the tension. Furthermore, it should give readers both the courage to resist seductive influences and the inspiration to arrive at decisions less likely to lead to remorse, disruption of family ties, or disappointment with unmet expectations.

Two aspects of the book in particular contribute to its success in presenting a balanced perspective on the general issue of “surgically shaping children.” First, the book focuses on three conditions for which normalizing surgery is widely practiced, each giving rise to a different level of controversy. From most controversial to least, these conditions are ambiguous genitalia, dwarfism (achondroplasia), and cleft lip and palate. Through analysis of the justifications for surgery to repair a cleft lip and palate and the moral failings attending most operations on infant genitalia (and associated practices), the contributors develop useful guidelines for decision making. The other aspect of the book that most contributes to its usefulness is its inclusion of many different voices — those of scholars and professionals from different disciplines (including bioethics, medicine, social work, philosophy, psychology, and law), of people with opposing viewpoints, and of people affected by one of the three conditions (some who have been surgically treated and some who have not).

All the chapters are well written and engaging. The positioning of the personal accounts at the front of the book serves not only to engage the reader but also to bring to the fore hitherto neglected issues: How have our surgical practices affected those who have been subjected to them? Is conformity to social ideas about normality of more value than living one’s life differently but authentically? What might we be sacrificing as a result of our desires to protect children from the reactions of those who are intolerant of differences? If parents by nature want to protect, and surgeons by nature want to fix, we need to hear from those with first-hand experience of what is to be gained and lost, because they are the truest authorities.

Parents facing grueling decisions about surgical interventions for their children will find great solace in this book. Its purpose is not to preach but to encourage reflection and facilitate informed consent. It illuminates subtle dangers: the danger that the decision to obtain surgery can send the message that the child falls short of parents’ expectations, the danger that decisions made without the participation of children may alienate them from their parents, and the danger that expectations about the ultimate success or helpfulness of the surgery may be overestimated. Parents can learn to ask important questions, to better prepare their children for the world, to master
their own vulnerabilities, and to develop the strength to oppose unexamined assumptions. In the end, they will be reassured that the most important determinant of their child’s happiness and well-being lies well within their own powers of unconditional love.

Sharon E. Sytsma, Ph.D.
Northern Illinois University
DeKalb, IL 60115
ssytsma@niu.edu

MELANCHOLIA: THE DIAGNOSIS, PATHOPHYSIOLOGY, AND TREATMENT OF DEPRESSIVE ILLNESS


Depression is hardly a neglected subject. Not only is it recognized by the World Health Organization as the leading cause of disability worldwide, but library shelves groan with books on the subject. What distinguishes this book is its emphasis on an extreme form of depression, for which the authors, Michael Alan Taylor and Max Fink, use the old-fashioned term “melancholia.” They believe that patients with melancholia are not receiving optimal treatment because the field has shifted its attention to milder mood disorders. Their aim is to put melancholia back on the map.

They begin by reconsidering the classification of mood disorders in psychiatry’s diagnostic guide, The Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), which, in the absence of sufficient knowledge about cause or pathophysiology, classifies most mental disorders as patterns of behavior. The category called mood disorders has a number of entries, such as major depressive disorder, dysthymic disorder, and bipolar disorder, each with a menu of characteristic symptoms as well as a series of “specifiers.” For major depression, the specifiers include severity (mild, moderate, or severe), the presence or absence of psychotic features (such as the delusion of deserving punishment), and “melancholic features” (such as loss of pleasure in all activities, anorexia, early-morning awakening, and psychomotor retardation or agitation).

Taylor and Fink propose a different way of slicing up the pie. Their objection to the official diagnostic scheme is that it fails to recognize melancholia as a “unique stage” of depression that “represents a qualitative change in underlying pathophysiology,” usually accompanied by impaired feedback regulation of the adrenal cortex. Because they believe that melancholia is a qualitatively different type of depression, they argue that the critical diagnostic distinction between various mood disorders should be that between melancholia and nonmelancholia.

Should a diagnosis of melancholia be made, the authors provide several therapeutic recommendations. The first is to consider initial treatment with a tricyclic antidepressant such as nortriptyline or desipramine, which the authors find more effective for melancholia than the selective serotonin-reuptake inhibitors (SSRIs) that are used widely for milder depression. Taylor and Fink are particularly skeptical of the practice of putting...
patients through lengthy trials with a series of ineffective SSRIs, thereby condemning the patients to months of excruciating mental distress. And they deplore the reluctance to use electroconvulsive therapy, which can quickly relieve melancholia in the vast majority of cases and which they think should be considered a first-line treatment for patients with melancholia accompanied by prominent psychotic symptoms.

For psychiatrists who specialize in the treatment of people with severe depression, there is not a lot that is new in this book. But Taylor and Fink are right to call attention to the current neglect of melancholic depression. For example, the elaborate and expensive study of treatments for major depression, Sequenced Treatment Alternatives to Relieve Depression (STAR*D),1,2 specifically excluded patients with depression with psychotic features, and many of these patients would probably meet Taylor and Fink's criteria for melancholia. The study also omitted electroconvulsive therapy as a treatment alternative for patients who did not respond to multiple drug trials, even though about 20% of the STAR*D patients had depression with melancholic features, as observed by Khan and colleagues.3

In considering the reasons why patients with melancholy no longer occupy center stage, Taylor and Fink speculate that “intrusive actions of the pharmaceutical industry encouraged a weakening of criteria to justify the use of antidepressant drugs in the largest number of persons.” Whereas the expansion of diagnostic criteria for depressive disorders, whatever its causes, has also had benefits, the authors’ emphasis on the special nature of melancholia serves as a welcome reminder that we should pay more attention to this blackest form of mood disorder.

Samuel H. Barondes, M.D.
University of California, San Francisco
San Francisco, CA 94143-0984
barondes@cgl.ucsf.edu


NOTICES

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.

UNIVERSITY OF MINNESOTA
A conference entitled “Creating Stem Cells by Research Cloning: Scientific, Ethical, Legal, & Policy Challenges” will be held in Minneapolis on Feb. 26.

Contact the Consortium on Law and Values in Health, Environment, & the Life Sciences, Mondale Hall, Suite N140, University of Minnesota Law School, 229 19th Ave. S., Minneapolis, MN 55455; or call (612) 625-0055; or fax (612) 624-0143; or see http://www.lifesci.consortium.umn.edu/conferences/scnt.php; or e-mail lawvalue@umn.edu.

INTERNATIONAL ACADEMY OF CARDIOLOGY ANNUAL SCIENTIFIC SESSIONS 2007
The “13th World Congress on Heart Disease” will be held in Vancouver, B.C., Canada, July 28–31. Deadline for submission of abstracts is Feb. 28.

Contact Dr. Asher Kimchi, International Academy of Cardiology, P.O. Box 17659, Beverly Hills, CA 90209; or call (310) 657-8777; or fax (310) 659-4781; or e-mail klimedco@ucla.edu; or see http://www.cardiologyonline.com.

MAYO CLINIC
The following courses will be offered: “Mayo Clinic Menopausal Medicine: Care for the Mature Female” (San Diego, CA, March 1–3); “An Overview of Perioperative Medicine 2007” (Rochester, MN, March 22–24); “2007 Mayo Clinic Headache Symposium” (Chicago, April 13–15); “Indications for Blood & Marrow Transplantation in the Era of Targeted Therapies” (Minneapolis, April 14); and “Mayo Echocardiography Review Course” (Rochester, MN, May 19–22).

Contact Mayo School of CME, 200 First St. SW, Rochester, MN 55905; or call (800) 323-2688 (national) or (507) 280-2509 (MN); or fax (507) 284-0532; or see http://www.mayo.edu/cme; or e-mail cme@mayo.edu.

CONTROVERSIES IN ORTHOGNATHIC SURGERY
The 7th Annual Symposium will be held in Washington, DC, on March 24.

Contact Suzanne Eggers, Medical Education Division, ImproMed, 1116 W. Centre Ave., Suite 2, Portage, MI 49024-5318; or call (877) 665-8326 (national) or (269) 329-0651; or fax (269) 329-0505; or e-mail seggers@impromed.org; or see http://www.impromed.org/cme/acmf/march.

UNIVERSITY OF CALIFORNIA GENETICS OF ABSOLUTE PITCH STUDY
The University of California is recruiting volunteer participants in a study to identify genes responsible for absolute pitch ability. See http://perfectpitch.ucsf.edu; or call (888) TUNEDIN (886-3346); or e-mail ppitch@itsa.ucsf.edu.

SOCIETY OF THORACIC RADIOLOGY
The annual meeting, entitled “Thoracic Imaging 2007,” will be held in Las Vegas, March 25–29.

Contact Society of Thoracic Radiology, c/o Matrix Meetings, P.O. Box 7169, Rochester, MN 55903-7169; or call (507) 288-5620; or see http://www.thoracicrad.org.
A 45-year-old healthy man was involved in demolishing an industrial plant in which glass had been etched. He was exposed to a reservoir of 70% hydrofluoric acid while repairing a pipeline. He was admitted to the intensive care unit for second-degree and third-degree burns from hydrofluoric acid affecting 30% of his body-surface area, including both hands, both forearms, the chest, back, scalp, and neck. After penetrating tissue, hydrofluoric acid dissociates into hydrogen and fluoride ions, of which particularly fluoride is toxic. Since fluoride ions are inactivated by means of precipitation with calcium and magnesium, the infusion of calcium and magnesium is considered a therapy in patients with hydrofluoric acid burns. In this patient, magnesium was infused intravenously, and calcium was infused intravenously and intraarterially (through the brachial artery) and was applied topically to the burned skin. The blood magnesium level was always within the normal range during substitution therapy. Blood levels of ionized calcium were initially elevated to up to 1.75 mmol per liter but were within the normal range after 36 to 48 hours. As a result of this intense calcium and magnesium therapy, cutaneous calcification developed on the fingertips by 36 to 48 hours, as well as on the dorsal and palmar aspects of the hand (Panels A and B, respectively). Three months later, the patient had regained an almost full range of motion, was free of symptoms, and had a good aesthetic result.

Copyright © 2007 Massachusetts Medical Society.