The NEW ENGLAND JOURNAL of MEDICINE

Perspective

JULY 21, 2005

The Controversy over Guidant's Implantable Defibrillators

Robert Steinbrook, M.D.

On October 4, 2001, Joshua Oukrop, a Minnesota teenager with hypertrophic cardiomyopathy and a high risk of sudden death from ventricular fibrillation, received an implantable cardioverter–defibrillator (ICD). The device was a Ventak Prizm 2 DR Model 1861 manufactured by Guidant (Indianapolis). After it was implanted, Oukrop's physicians at the Minneapolis Heart Institute Foundation checked it every three months (most recently on January 31, 2005) and found no problems.[1]

On March 14, 2005, Oukrop, then a 21-year-old college student, collapsed and died in a remote area of southeastern Utah during a spring-break bicycling trip with his girlfriend.[1,2] An autopsy revealed no clinically significant pathology beyond his massive left ventricular hypertrophy. His physicians were stunned by his death. ICDs have been shown to be almost invariably successful in preventing sudden death in young patients with hypertrophic cardiomyopathy, as long as they do not have end-stage heart failure — which Oukrop did not.[3] When the manufacturer analyzed his ICD, it determined that the device had short-circuited internally while trying to deliver high-voltage therapy and had been permanently disabled (see diagram). Moreover, its memory had been destroyed, making the time of failure impossible to pinpoint.[1,4]

Oukrop's physicians at the Minneapolis Heart Institute Foundation included Dr. Barry Maron, the director of the Hypertrophic Cardiomyopathy Center; Dr. Robert Hauser, a senior consulting cardiologist at the institute; and Dr. Charles Gornick, who had implanted the ICD. After a company official told Maron what had happened, Maron called Hauser. Hauser searched the Manufacturer and User Facility Device Experience (MAUDE) database maintained by the Food and Drug Administration (FDA), which contains reports of adverse events involving medical devices. There, he found other reports from Guidant of instances in which the Prizm 2 DR Model 1861 had short-circuited in exactly the same way as Oukrop's had done. Forty-seven other patients with this device are followed at the institute, including Oukrop's father and 10 others who have hypertrophic cardiomyopathy.

In June, Maron recalled: "We became very concerned. We were keeping a secret not just from our patients and their physicians, but also from all the patients with the device and their physicians.

Infection-Control Report Cards — Securing Patient Safety

Robert A. Weinstein, M.D., Jane D. Siegel, M.D., and P.J. Brennan, M.D.

For many of us, the specter of report cards conjures up anxiety dreams. Nevertheless, public report cards have infiltrated many industries — airlines and banking, for instance — and various levels of government, and health care appears to be next. The belief that hospitals reporting lower infection rates are safer and that informed consumers will obtain safer care has driven many U.S. states to consider legislation requiring report cards on nosocomial infections.

Advocates of public reporting have been spurred on by the occurrence of nosocomial infections in 5 to 10 percent of hospitalized patients; increasing rates of antibiotic resistance; press coverage of cases of devastating nosocomial infection; and the view that many infections should be preventable. In little more than a year, 39 states introduced legislation and 6 states passed laws requiring disclosure of such infections to the state and, in most cases, to the public (see map). Although the movement is consumer-driven, health care providers share the goals of reducing infection rates and improving patient safety.[1]

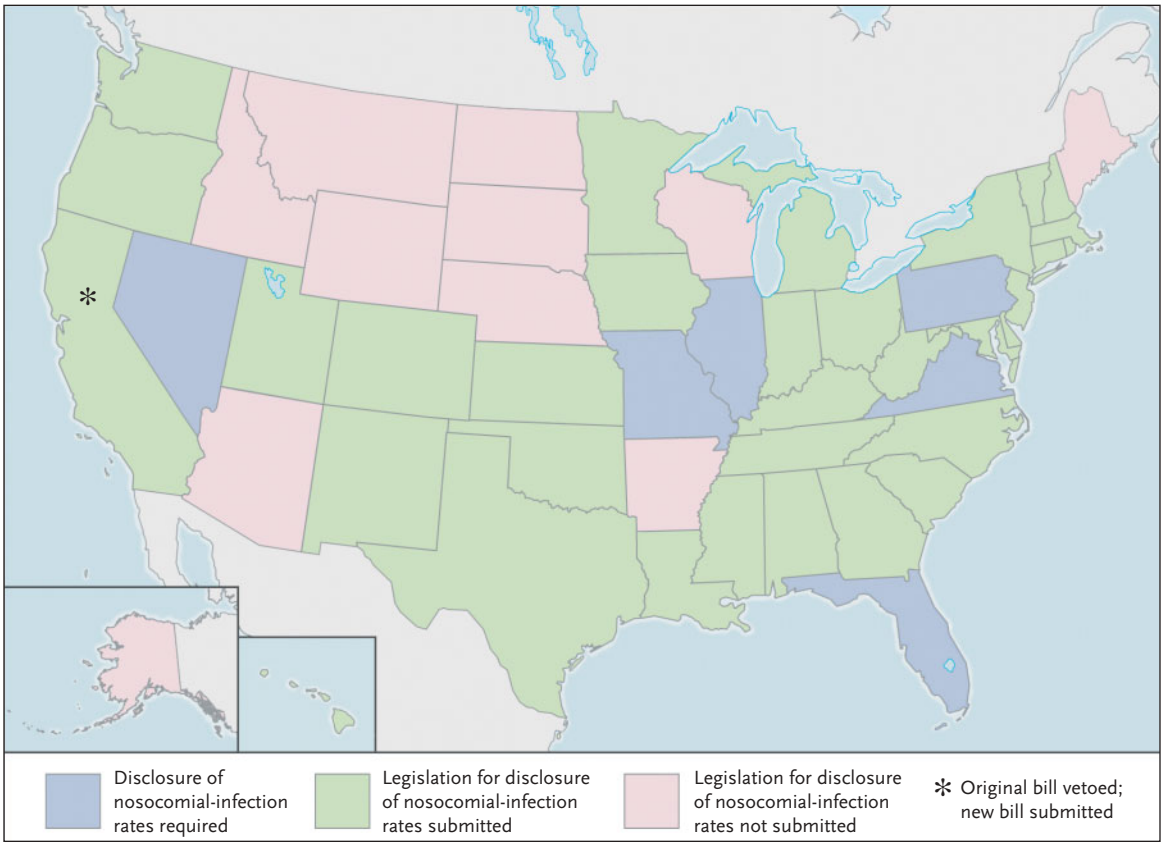
The public reporting of hospitals' performance is not new, although older systems paid little attention to nosocomial infections. For decades, several states have required that hospitals report death rates associated with cardiac surgery and other conditions. Although the effectiveness of these experiments has been mixed, reductions in the rate of death associated with coronary-artery bypass grafting in New York and Pennsylvania have been attributed, in part, to public reporting.

A key lesson from past reporting is the importance of "risk adjusting" the outcome data to account for essential differences in populations of patients. For example, surgical risk among 80-year-old patients with diabetes is greater than that among previously healthy 45-year-olds. We have also learned that we must select denominators carefully in order to avoid artificial inflation or deflation of rates; that sophisticated information technology is required; and that it can be difficult to define useful benchmarks, especially for small hospitals, so that reporting a trend for a particular hospital may provide more useful information than does comparing hospitals.

The Joint Commission on Accreditation of Healthcare Organizations and the Center for Medicare and Medicaid Services (CMS) have established Web-based public reports for participating hospitals. Both groups initially focused on process measures rather than outcome measures, a strategy that avoided the quagmire of risk adjustment. Neither group has yet addressed nosocomial infections, although the CMS is developing a Surgical Care Improvement Project that will probably report process and outcome measures. The Centers for Disease Control and Prevention (CDC), through its National Healthcare Safety Network, provides a mechanism for the confidential disclosure of nosocomial infections, allowing hospitals to compare their performance with that of others. Several states are considering statewide participation in the network as a reporting solution.

Current report cards focus on three types of common infections that are associated with high morbidity and mortality and that are likely to be controllable: infections associated with central venous catheters, surgical-site infections, and ventilator-associated pneumonia. We know that informing surgeons of their wound-infection rates can lead to reductions in those rates, presumably by reinforcing the use of sensible interventions (e.g., limiting the amount of movement in and out of operating rooms in order to lower bacterial loads). The most rigorous study of the impact of surveillance — the CDC's Study on the Efficacy of Nosocomial Infection Control, conducted in the 1970s — demonstrated a 32 percent decrease in infection rates in hospitals that implemented standardized surveillance methods and ongoing control measures and that had adequate infection-control staffing and expert physician epidemiologists.

Studies of infections related to devices, particularly vascular and urinary catheters and ventilators, have demonstrated the usefulness of key performance measures — such as site preparation and care and operator expertise for the control of venous catheter–related infections — and of "bundling" evidence-based prevention mea-



Status of Legislation Requiring Public Disclosure of Rates of Nosocomial Infection, by State.

sures into comprehensive control programs.[2] Studies have also highlighted certain difficulties involved in measuring some outcomes — for instance, the lack of an easily applied clinical definition of ventilator-associated pneumonia, the difficulty of tracking surgical-site infections in the community (now that the average postoperative stay is shorter than the incubation period for most wound infections), and the large confidence intervals around reported infection rates in smaller hospitals and for uncommon procedures.

In the light of these difficulties, will this type of public reporting result in the sort of improvements achieved by reporting wound-infection rates to surgeons? The answer is uncertain. Many experts recommend further study before states initiate costly, labor-intensive reporting programs, and a few states have passed laws that require such studies (see map). But more states have already embarked on the path of public reporting, and their legislators need advice urgently.

Recent recommendations from the CDC suggest that states focus on a combination of linked process and outcome measures.[1] We support the study of reporting when possible. For states that have passed laws requiring public reporting, we suggest measuring rates that can be compared meaningfully, that should be tracked anyway, and whose reporting is most likely to lead to improved care. Such process measures include assessments of the timely administration of perioperative antibiotic prophylaxis, vascular-catheter insertion practices, and hand hygiene. Outcome measures include the rate of infections in the intensive care unit associated with central vascular catheters and the rate of reoperation or rehospitalization for surgical-site infections. Other measures for special settings could include the rates of nosocomial influenza, respiratory syncytial virus, rotavirus infection, and cases of diarrhea associated with *Clostridium difficile*. Infections caused by multidrug-resistant pathogens, such as methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci, are also important, but because of laboratory logistics and the difficulty of verifying an infection's nosocomial origin, meaningful reporting is not yet possible.

States that have involved ex-

perts in health care epidemiology early in the development of laws have produced the most useful legislation. In addition, we favor phasing in reporting requirements incorporating the process measures that are the most readily obtained and compared, in order to allow hospitals and health departments to develop, refine, and validate data-collection systems. States must also consider the cost of these programs and work with hospital associations to develop realistic plans for support and funding.

To understand other relevant concerns, states should review the reasons given by Governor Arnold Schwarzenegger for his recent veto of California's legislation, which included problems with auditing and validating data, the need to redirect resources from successful programs, and extant mandates from national organizations that already scrutinize infection control. Most important, states must work with experts in health care communications and consumer reporting to define the sorts of rates that will tell patients what they need to know.

Report cards assessing nosocomial infections are a reality. Their success will require interdisciplinary collaboration, a greater commitment of resources to infection-prevention practices, and conspicuous inclusion of these efforts in patient-safety programs. Research is needed to identify the most meaningful metrics, determine the best way to report them, and assess whether such reporting improves patient safety. These challenges present unprecedented opportunities to improve patient care, if we can only put our anxieties to rest and move forward.

Dr. Weinstein is the chair of infectious diseases at John H. Stroger (Cook County) Hospital and a professor of medicine at Rush University Medical Center, Chicago; Dr. Siegel is a professor of pediatrics at the University of Texas Southwestern Medical Center and chair of the Infection Control Committee at Children's Medical Center Dallas; and Dr. Brennan is a professor of medicine and chief medical officer at the University of Pennsylvania School of Medicine and Health System, Philadelphia.

1. McKibben L, Horan TC, Tokars JI, et al. Guidance on public reporting of healthcare-associated infections: recommendations of the Healthcare Infection Control Practices Advisory Committee. Am J Infect Control 2005;33:217-26.

2. Institute for Healthcare Improvement: 100k Lives Campaign. (Accessed June 30, 2005, at http://www.ihi.org/IHI/Programs/Campaign.)

CORRECTION

Infection-Control Report Cards — Securing Patient Safety

Infection-Control Report Cards — Securing Patient Safety . On page 226, in the map showing the status of legislation requiring public disclosure of rates of nosocomial infection, Indiana, Louisiana, Texas, and Utah should have been highlighted as states in which legislation for the study of the reporting of nosocomial infection has been passed. (Updated information about the status of such legislation is available at www.apic.org.) Also, on page 227, the volume and page numbers in reference 1 should have read ``26:580-7,´´ rather than ``33:217-26,´´ as printed. We regret the error.

N Engl J Med 2005;353:1869-a

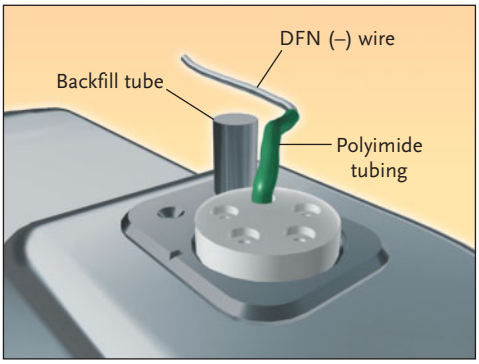


Diagram Showing Components of the Ventak Prizm 2 Device before Modification in April 2002.

A failure involving a deterioration in a wire insulator can result in an electrical short between the DFN wire and the backfill tube. Diagram courtesy of Guidant.

On May 12, four Guidant officials came to my office and gave a very educational presentation. I asked, 'What are we going to do about this? We are in an untenable situation ethically and morally with our patients. How are we going to get the word out?' They said, 'Well, we are not. We don't think we need to. And we don't think it's advisable.' The officials expressed doubt that the patients would be able to understand the medical issues involved in determining whether or not to replace the devices. I said, 'I think this is the biggest mistake you will ever make.' They said they didn't agree."

It was subsequently disclosed that Guidant, the second-largest manufacturer of implantable defibrillators, had identified the electrical flaws in the Prizm 2 DR in February 2002 and had made manufacturing changes, on April 16 and November 13 of that year, in an effort to prevent this rare but unpredictable and catastrophic type of failure.[4] To date, there have been no reports of failures of such devices built after the April 2002 change. Guidant, however, continued to sell devices that had been manufactured before that change was made and issued no public statements about the problem or the corrections. The company's first announcement came on May 23, 2005 — more than three years after Guidant had become aware of the problem and hours before the *New York Times* published an article about Oukrop's death.[2] Between May 23 and June 24, Guidant informed physicians and patients about problems with many of its ICD devices and came under intense regulatory and public scrutiny. After the company provided physicians at the Minneapolis Heart Institute Foundation with a list of the 47 patients with the potentially defective ICDs, all were contacted. As of July 5, nine had had their devices replaced and four had replacements scheduled.

Guidant is still working its way through the ICD issue. On June 17, 2005, the company noted in a statement that "after making the manufacturing changes, Guidant sold products manufactured before the April 2002 change. At that time data did not show an unusual failure rate and Guidant believed the device to be reliable."[4] Since November 2003, Dr. Beverly Lorell, a professor of medicine at Harvard Medical School, has been vice president and chief medical and technology officer at Guidant. In a written statement, Lorell said that the company's "first priority is patient safety and this is the foundation of trust by physicians and patients. We also understand that physicians and individual patients may make different choices regarding treatment options even when the risks of an adverse clinical event related to a life-sustaining device . . . are very low in frequency. For this reason, Guidant's recent voluntary physician communications and recalls for separate and distinct device safety issues reflect a very stringent approach by the firm in the communication of potential safety issues related to marketed devices."

The fallout from the potentially preventable death of Joshua Oukrop has triggered a broad discussion about the propriety of Guidant's actions and the safety of ICDs and medical devices in general. It has also led to debate about the appropriate standards for informing physicians and patients about safety issues and the responsibilities of industry, the FDA, and the medical community. The matter is particularly urgent because the number of people with ICDs is increasing rapidly. In 2003, the Center for Medicare and Medicaid Services paid for 52,500 ICD implantations; in 2004, it paid for 65,000. With expanded coverage, more than 500,000 Medicare beneficiaries may become eligible for an ICD. Hauser has estimated that worldwide more than 200,000 ICDs will be implanted or replaced this year.

On July 1, the FDA classified as class I recalls Guidant's recent notifications with regard to Prizm 2 DRs that were manufactured on or before April 16, 2002, and two other models of implantable defibrillators, the Contak Renewal Model H135 and the Contak Renewal 2 Model H155, that were manufactured on or before August 26, 2004. A class I recall indicates the highest level of risk, because "there is a reasonable probability that if a particular device is malfunctioning, the malfunctioning device will cause serious adverse health consequences or death."[5] Of 42,000 affected devices of the three recalled types

worldwide, 20,600 (including about 13,900 Prizm 2 DRs) are still implanted in patients.[4,5] The devices are subject to the development of an internal short circuit when they are attempting to deliver an electrical shock to the heart — thereby preventing the treatment of arrhythmias. According to the FDA, "the problem is caused by deterioration of electrical insulation in the device and can only be detected after the device has already malfunctioned. The device does not give any sign of impending failure and there is no test that predicts whether the device will fail."[5]

As of June 17, 2005, Guidant and the FDA were aware of 43 reports of device failures, including 28 involving Prizm 2 DR devices.[4] They include in this count the death of Oukrop and the death on May 30, 2005, of a patient with a Contak Renewal Model H135 device that was manufactured in December 1999. Some failures of Prizm 2 DRs have been recognized while the device was being implanted or after a spontaneous defibrillator shock while the patient was still in the hospital.[1]

In the 15 Contak Renewal failures, the short circuit diverted energy away from the heart, so that only about 20 percent of the intended shock energy was delivered. According to the report of the patient death that Guidant filed with the FDA, the device delivered multiple shorted shocks, but they did not convert the patient's ventricular tachycardia. In all cases, the device must be replaced if this short circuit occurs. The death remains under investigation.

Guidant has acknowledged that the actual rate of failure may be higher than the reported rate and that the number of associated deaths may be underreported, since ICDs are not routinely evaluated after death.[4] The FDA and Guidant have advised physicians and patients to make individual decisions about whether to remove and replace the affected defibrillators on the basis of the specific medical situation of the patient; Guidant will provide replacement devices at no charge.[4]

The FDA classified Guidant's notifications about other ICDs as class II recalls, because "the malfunctioning product may cause temporary or medically reversible adverse health consequences, however the probability of serious adverse health consequences is remote."[5] Three types of devices — the Ventak Prizm AVT, Vitality AVT, and Renewal AVT — are subject to a memory error that may limit their ability to provide therapy but that can be corrected with reprogramming. Among the 21,000 such devices that have been implanted worldwide, two incidents have been confirmed. Neither resulted in death or injury.

Five devices — the Contak Renewal 3 and 4, Renewal 3 and 4 AVT, and Renewal RF — are subject to a rare type of component failure; four such failures have been confirmed among the approximately 46,000 devices implanted, and a fifth occurrence is suspected but has not been confirmed. This failure occurs when a magnetic switch sticks in the closed position. This malfunction may inhibit the device's ability to treat ventricular and atrial tachyarrhythmias and may accelerate the depletion of the battery. If this problem develops, the device emits an audible tone. As of June, there had been no injuries to patients beyond the need to replace the device in four instances; the other instance occurred before implantation. Guidant has recommended that physicians stop implanting these devices until further notice. In the case of devices that have already been implanted, it is recommended that physicians make a programming change to ensure that appropriate therapy can be delivered.

Since February 2005, there have been "an above-average number" of notifications to physicians about problems with ICDs, according to Dr. William H. Maisel, a cardiologist at Beth Israel Deaconess Medical Center in Boston and the chair of the FDA's Circulatory System Devices Panel. The notifications involve all the leading manufacturers. In February, Medtronic (Minneapolis), the largest of these manufacturers, advised physicians about the potential for premature battery failure in some of its implantable defibrillators that could worsen over time and eventually affect 0.2 to 1.5 percent of the devices. Of 87,000 potentially affected units, about 13,000 had been replaced as of the end of May. No injuries or deaths have been reported. In June 2005, St. Jude Medical (St. Paul, Minnesota), the third-largest manufacturer, advised physicians about two anomalies that could affect 30,000 implanted ICDs in the United States and that can be corrected with a software download. No clinical complications have been reported.

In the most serious situations, physicians are having difficult conversations with patients about device replacement. All ICDs must be replaced when their batteries wear out — after an average of five years. The surgical procedure, however, carries risks — principally, the risk of infection. Even

if the statistical risk that a patient's implanted device will not save his or her life may be lower than the risks associated with replacing it, the uncertainty can weigh heavily. As Maisel said in a recent interview, "Evidence-based decision making cannot factor in the emotional consequences to patients of living every day knowing that their ICD might not work if it is needed. For some patients, this is a very valid reason to replace a device."

The recalls have also focused attention on Guidant. In June 2003, Endovascular Technologies, a subsidiary of Guidant, pleaded guilty to 10 felony counts and agreed to pay $92.4 million in civil and criminal penalties related to its Ancure Endograft system, a stent–graft device inserted by means of a catheter for the treatment of abdominal aortic aneurysms. The company admitted that it had lied to the government and hidden thousands of serious health problems, including 12 deaths. The company stopped selling the system in March 2001, made changes, and reintroduced it to the market in August 2001.

In 2004, Guidant sold $1.8 billion of implantable defibrillators, nearly half of its $3.8 billion in total sales. On December 15, 2004, it was announced that Johnson and Johnson would acquire Guidant for $25.4 billion. According to securities-fraud lawsuits that have been filed beginning in June 2005, Guidant and its executives have allegedly covered up the problems with the ICDs.

Since the merger agreement was announced, Guidant executives have sold millions of dollars of company stock, according to filings with the Securities and Exchange Commission. For example, on May 17, 2005, Lorell, the chief medical and technology officer, sold 23,300 shares for $1.71 million. On May 23, 2005, the day before the problems with the Prizm 2 DR were the subject of a front-page article in the *New York Times*, she sold 22,667 more shares for $1.68 million. When asked to explain the transactions, Lorell did not respond.

The FDA's investigation of Guidant's actions is continuing, and further information from both the company and the government is likely to be available soon. Later this year, the Heart Rhythm Society, a professional association of arrhythmia specialists, plans to develop guidelines regarding ICD recalls, manufacturer-notification standards, and when to replace devices. The society, however, receives 25 percent of its $8.5 -million annual budget from corporate support — including funds from Guidant, Medtronic, and St. Jude Medical — arousing concern about potential conflicts of interest. Many of its members also have consulting and other financial ties to device manufacturers. In addition, Guidant is establishing its own panel of experts to recommend guidelines for disseminating information.

In the weeks ahead, the FDA will need to reexamine its regulations and procedures for device surveillance. It should consider making changes to better inform physicians and patients and to more thoroughly ensure the quality of medical devices. According to Hauser of the Minneapolis Heart Institute Foundation, "it is very important that the FDA get this done. It is long overdue. It is not just this one incident. It is broader than that." In addition, the entire industry will have to enhance the safety and reliability of these lifesaving devices. For more than three years, Guidant kept quiet about the serious malfunctions of some of its ICDs and continued to sell defective devices after it made manufacturing changes to fix the defects. The company will have to regain the trust of physicians and patients.

Dr. Steinbrook is a national correspondent for the *Journal*.

1. Gornick CC, Hauser RG, Almquist AK, Maron BJ. Unpredictable implantable cardioverter-defibrillator pulse generator failure due to electrical overstress causing sudden death in a young high-risk patient with hypertrophic cardiomyopathy. Heart Rhythm 2005;2:681-3.

2. Meier B. Maker of heart device kept flaw from doctors. New York Times. May 24,2005: A1.

3. Maron BJ, Shen W-K, Link MS, et al. Efficacy of implantable cardioverter–defibrillators for the prevention of sudden death in patients with hypertrophic cardiomyopathy. N Engl J Med 2000;342:365-73.

4. Gorsett A. Urgent medical device safety information & corrective action: VENTAK PRIZM 2 DR, model 1861. Indianapolis: Guidant Cardiac Rhythm Management, June 17, 2005. (Accessed July 6, 2005, at http://guidant.vanosteen.com/news/prizm2_dr.pdf.)

5. FDA updates consumers on Guidant Corporation's implantable defibrillators. News release of the Food and Drug Administraton, Rockville, Md., July 1, 2005. (Accessed July 6, 2005, at http://www.fda.gov/bbs/topics/news/2005/new01198.html.)

THIS WEEK in the JOURNAL

ORIGINAL ARTICLE

Benign Breast Disease and the Risk of Breast Cancer

More than 9000 women were followed for a median of 15 years after a diagnosis of benign breast disease. As compared with women in a SEER database, they had an increased risk of subsequent breast cancer, especially if the benign lesion showed signs of atypia. A family history of breast cancer and younger age at diagnosis also increased the risk. Cancers developed in either breast, but an excess number occurred in the same breast.

SEE P. 229; EDITORIAL, P. 297; CME, P. 326

ORIGINAL ARTICLE

Statins in Type 2 Diabetes and Hemodialysis

Statins reduce cardiovascular events in persons with type 2 diabetes mellitus. This study randomly assigned patients with type 2 diabetes receiving hemodialysis to receive 20 mg of atorvastatin per day or matching placebo. Atorvastatin reduced all cardiac events combined but not all cerebrovascular events combined or total mortality.

SEE P. 238

BRIEF REPORT

Hyperinsulinemic Hypoglycemia with Nesidioblastosis after Gastric-Bypass Surgery

This report describes six patients who had postprandial symptoms of neuroglycopenia from endogenous hyperinsulinemic hypoglycemia after Roux-en-Y gastric bypass surgery. Hypoglycemic symptoms diminished in all patients after partial pancreatectomy. The authors speculate that hyperfunction of pancreatic islets did not lead to obesity but rather that beta-cell trophic factors may have increased as a result of gastric bypass.

SEE P. 249; EDITORIAL, P. 300

SPECIAL ARTICLE

Quality of Care in U.S. Hospitals as Reflected by Standardized Measures, 2002–2004

In 2002, the JCAHO began requiring hospitals to report their performance on standardized measures of the quality of health care and began providing hospitals with feedback on their performance. This study documented substantial improvement between 2002 and 2004 in hospitals' performance for myocardial infarction, congestive heart failure, and pneumonia. Hospitals with a low level of performance at baseline had the largest improvements.

SEE P. 255; EDITORIAL, P. 302

SPECIAL ARTICLE

Care in U.S. Hospitals — The Hospital Quality Alliance Program

This study examined hospitals' performance on standardized measures of the quality of care for acute myocardial infarction, congestive heart failure, and pneumonia. The quality of care varied among hospitals. Hospitals' performance on the measures for myocardial infarction and congestive heart failure was not closely correlated with performance on the pneumonia measure, suggesting that efforts to monitor the quality of care may need to include a wide range of medical conditions.

SEE P. 265; EDITORIAL, P. 302; CME, P. 327

CURRENT CONCEPTS

Benign Breast Disorders

About 50 percent of women have histologic evidence of some degree of fibrocystic changes in the breast. This review article describes the spectrum of benign breast disorders, including those that increase the risk of malignant disease. The authors explain the evaluation and management of breast pain, breast discharge, and focal lesions, as well as the options that can prevent progression to breast cancer.

SEE P. 275; CME, P. 325

CASE RECORDS OF THE MASSACHUSETTS GENERAL HOSPITAL

A Man with Cough, Fever, and Altered Mental Status

An 81-year-old man was admitted to the hospital because he had had cough and fever for two weeks and had recently become confused. He had had chronic lymphocytic leukemia for six years. His mental status continued to deteriorate, with coma and respiratory arrest, and he died on the 12th hospital day.

SEE P. 287

CLINICAL IMPLICATIONS OF BASIC RESEARCH

A New Strategy to Counter Allergy

A chimeric fusion protein inhibits the allergic reaction induced by cat allergen in a mouse model.

SEE P. 310

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 JULY 21, 2005 VOL. 353 NO. 3

Benign Breast Disease and the Risk of Breast Cancer

Lynn C. Hartmann, M.D., Thomas A. Sellers, Ph.D., Marlene H. Frost, Ph.D., Wilma L. Lingle, Ph.D., Amy C. Degnim, M.D., Karthik Ghosh, M.D., Robert A. Vierkant, M.A.S., Shaun D. Maloney, B.A., V. Shane Pankratz, Ph.D., David W. Hillman, M.S., Vera J. Suman, Ph.D., Jo Johnson, R.N., Cassann Blake, M.D., Thea Tlsty, Ph.D., Celine M. Vachon, Ph.D., L. Joseph Melton III, M.D., and Daniel W. Visscher, M.D.

ABSTRACT

BACKGROUND

Benign breast disease is an important risk factor for breast cancer. We studied a large group of women with benign breast disease to obtain reliable estimates of this risk.

METHODS

We identified all women who received a diagnosis of benign breast disease at the Mayo Clinic between 1967 and 1991. Breast-cancer events were obtained from medical records and questionnaires. To estimate relative risks, we compared the number of observed breast cancers with the number expected on the basis of the rates of breast cancer in the Iowa Surveillance, Epidemiology, and End Results registry.

RESULTS

We followed 9087 women for a median of 15 years. The histologic findings were nonproliferative lesions in 67 percent of women, proliferative lesions without atypia in 30 percent, and atypical hyperplasia in 4 percent. To date, 707 breast cancers have developed. The relative risk of breast cancer for the cohort was 1.56 (95 percent confidence interval, 1.45 to 1.68), and this increased risk persisted for at least 25 years after biopsy. The relative risk associated with atypia was 4.24 (95 percent confidence interval, 3.26 to 5.41), as compared with a relative risk of 1.88 (95 percent confidence interval, 1.66 to 2.12) for proliferative changes without atypia and of 1.27 (95 percent confidence interval, 1.15 to 1.41) for nonproliferative lesions. The strength of the family history of breast cancer, available for 4808 women, was a risk factor that was independent of histologic findings. No increased risk was found among women with no family history and nonproliferative findings. In the first 10 years after the initial biopsy, an excess of cancers occurred in the same breast, especially in women with atypia.

CONCLUSIONS

Risk factors for breast cancer after the diagnosis of benign breast disease include the histologic classification of a benign breast lesion and a family history of breast cancer.

From the Divisions of Medical Oncology (L.C.H., M.H.F., J.J.), Experimental Pathology (W.L.L.), General Surgery (A.C.D.), General Internal Medicine (K.G.), Biostatistics (R.A.V., S.D.M., V.S.P., D.W.H., V.J.S.), Epidemiology (C.M.V., L.J.M.), and Anatomic Pathology (D.W.V.), Mayo Clinic College of Medicine, Rochester, Minn.; H. Lee Moffitt Cancer Center and Research Institute, Tampa, Fla. (T.A.S.); Wayne State University, Detroit (C.B.); and the University of California, San Francisco, San Francisco (T.T.). Address reprint requests to Dr. Hartmann at Mayo Clinic College of Medicine, Rochester, MN 55905.

N Engl J Med 2005;353:229-37.

Benign breast disease is an important risk factor for a later breast cancer, which can develop in either breast.[1] It encompasses a spectrum of histologic entities, usually subdivided into nonproliferative lesions, proliferative lesions without atypia, and atypical hyperplasias, with an increased risk of breast cancer associated with proliferative or atypical lesions.[2-4] The identification of benign breast disease has become more common as the use of mammography has increased, and thus, having accurate risk estimates for women who receive this diagnosis is imperative.

Important questions remain, however, about the degree of risk associated with the common nonproliferative benign entities and the extent to which family history influences the risk of breast cancer in women with proliferative or atypical lesions. Dupont and Page found that women with nonproliferative disease did not have an increased risk of a later breast cancer.[2] By contrast, a companion study to the National Surgical Adjuvant Breast and Bowel Project (NSABP) Breast Cancer Prevention Trial (P1) found a relative risk of 1.6 for women who received a diagnosis of a "lower category" of benign breast disease.[5] A limitation of the NSABP study, however, was the lack of central pathological review.

Another major question concerns the possible interplay between atypia and a family history of breast cancer. The Dupont and Page study found that women with atypia and a family history had 11 times the risk of those with nonproliferative lesions and no family history.[2] However, two other major studies of benign breast disease[6,7] did not find a significant interaction between atypia and family history. The duration of increased risk after a finding of benign disease on biopsy is also uncertain.[2,4,8]

Studies of benign breast disease can also clarify whether there is a continuum of breast alterations that culminates in breast cancer. However, it remains unclear which of the benign entities are actual precursors and which reflect a background of increased risk involving all breast tissue in a woman. Determining the extent of agreement between the side (right or left) of the benign lesion and the subsequent breast cancer is one means of assessing these issues.

To investigate these questions, we studied 9087 women with benign breast disease for whom we had follow-up data on breast-cancer events. This cohort has been followed for a median of 15 years, and 707 breast cancers have developed, making this, to our knowledge, one of the largest such studies of its kind. We report on the risk of breast cancer according to histologic findings, the age at diagnosis of benign breast disease, and the strength of the family history. We also recorded the side of the cancer (ipsilateral or contralateral) and the time to the diagnosis of cancer.

METHODS

STUDY POPULATION

We accessed data from the Mayo Clinic Surgical Index and Pathology Index to identify all women 18 to 85 years of age who had undergone surgical excision of a benign breast lesion during the 25-year period from January 1, 1967, through December 31, 1991. For women who had more than one biopsy during this period, we used the first sample. The original list contained 12,132 women, but we excluded 1,047 women for any of the following: a diagnosis of breast cancer or lobular carcinoma in situ at, before, or within six months after the biopsy of the benign lesion; mastectomy (unilateral or bilateral) or breast reduction at or before biopsy; or refusal to allow use of their medical records for research.[9] This left 11,085 women. Of these, 1053 (9.5 percent) had no follow-up information after the biopsy. Thus, a total of 10,032 women met our criteria for study entry and had follow-up information. Of these, 945 women had unusable or unavailable biopsy specimens of the benign lesion. The remaining group of 9087 women constitutes our study cohort. The relative risks of breast cancer (described below) did not differ significantly between the 10,032 women who met our criteria and the 9087 women who made up the study cohort (1.59 and 1.56, respectively).

FAMILY HISTORY AND FOLLOW-UP

A questionnaire designed for this study was used to obtain information about family history and other possible risk factors for breast cancer. Thus, our family-history data were obtained at the time of follow-up contact. We categorized family history as none, weak, or strong. The criteria for a strong family history were as follows: at least one first-degree relative with breast cancer before the age of 50 years or two or more relatives with breast cancer, with at least one being a first-degree relative. Any lesser degree of family history of breast cancer was categorized as weak. The questionnaire also asked about breast-cancer occurrences. Follow-up for breast-cancer events was also obtained through the comprehensive (inpatient and outpatient) Mayo medical

record. Questionnaire information was available for 5619 women (61.8 percent). Of the questionnaires, 604 (10.7 percent) were completed by proxy (the next of kin of a deceased patient). As of August 1, 2004, 7260 (79.9 percent) members of the cohort were still alive. All protocol procedures and patient-contact materials were reviewed and approved by the institutional review board of the Mayo Clinic; returning the contact materials was considered implied consent.

HISTOLOGY

Stored hematoxylin-and-eosin–stained sections from each participant were evaluated by a breast pathologist who was unaware of the initial histologic diagnoses and patient outcomes. Biopsy findings were classified according to the criteria of Page et al.[2,10] into the following categories: nonproliferative fibrocystic changes, proliferative fibrocystic changes without atypia, and proliferative fibrocystic changes with atypia (atypical ductal hyperplasia, atypical lobular hyperplasia, or both) (Fig. 1).[2,10] Biopsy specimens were designated as having proliferative fibrocystic changes if they contained any of the following: ductal hyperplasia (greater than mild), papilloma, radial scar, or sclerosing adenosis. Cysts, fibroadenoma, or columnar changes were considered nonproliferative unless they also contained one of the lesions denoted above.

STATISTICAL ANALYSIS

The duration of follow-up was calculated as the number of days from biopsy of the benign lesion to the date of the diagnosis of breast cancer, death, or last contact. We estimated relative risks on the basis of standardized incidence ratios (SIRs), dividing the observed numbers of incident breast cancers by population-based expected counts. We calculated these expected counts by apportioning each woman's follow-up into five-year age and calendar-

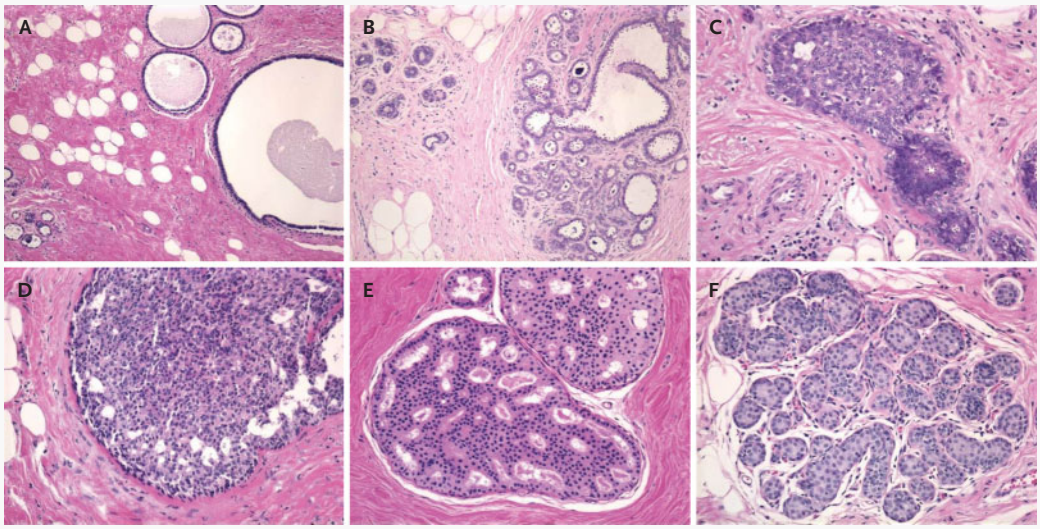


Figure 1. Histopathological Appearance of Benign Breast Disease (Hematoxylin and Eosin).

Panel A shows nonproliferative fibrocystic changes: the architecture of the terminal-duct lobular unit is distorted by the formation of microcysts, associated with interlobular fibrosis. Panel B shows proliferative hyperplasia without atypia. This is adenosis, a distinctive form of hyperplasia characterized by the proliferation of lobular acini, forming crowded gland-like structures. For comparison, a normal lobule is on the left side. Panel C also shows proliferative hyperplasia without atypia. This is moderate ductal hyperplasia, which is characterized by a duct that is partially distended by hyperplastic epithelium within the lumen. Panel D again shows proliferative hyperplasia without atypia, but this is florid ductal hyperplasia: the involved duct is greatly expanded by a crowded, jumbled-appearing epithelial proliferation. Panel E shows atypical ductal hyperplasia: these proliferations are characterized by a combination of architectural complexity with partially formed secondary lumens and mild nuclear hyperchromasia in the epithelial-cell population. Panel F shows atypical lobular hyperplasia: monotonous cells fill the lumens of partially distended acini in this terminal-duct lobular unit.

period categories, thereby accounting for differences associated with these variables. We used the Iowa Surveillance, Epidemiology, and End Results (SEER) registry as the reference population because of its demographic similarities to the Mayo Clinic population (80 percent of cohort members reside in the upper Midwest). Over 95 percent of our cohort was white, equivalent to that reported in Iowa census data during the study period.[11] In the SIR analyses, we considered the time since the original biopsy as a time-dependent variable and all other factors as fixed.

Associations between the risk of breast cancer and histologic findings, the age at diagnosis of benign breast disease, and the strength of the family history of cancer, as well as pairwise combinations of these variables, were examined with the use of Cox proportional-hazards regression analysis. The main effects for each categorized variable and the corresponding interaction terms were included in each model, and the statistical significance of each interaction was evaluated with the use of a multiple-degree-of-freedom likelihood-ratio test.

We studied ipsilateral and contralateral breast cancer as a function of the time since biopsy by estimating the relative risk of cancer in the same as compared with the opposite breast for five-year intervals. When calculating the incidence of ipsilat-

Table 1. Characteristics of the Women According to the Histologic Category of Benign Breast Disease.*

Characteristic	All Women (N=9087)	Nonproliferative Disease (N=6061)	Proliferative Disease without Atypia (N=2690)	Atypical Hyperplasia (N=336)
Percentage of total	100.0	66.7	29.6	3.7
Age at biopsy — no. of women (%)				
<40 yr	1841 (20.3)	1500 (24.7)	323 (12.0)	18 (5.4)
40–49 yr	2474 (27.2)	1621 (26.7)	770 (28.6)	83 (24.7)
50–59 yr	2145 (23.6)	1297 (21.4)	759 (28.2)	89 (26.5)
60–69 yr	1639 (18.0)	1034 (17.1)	522 (19.4)	83 (24.7)
≥70 yr	988 (10.9)	609 (10.0)	316 (11.7)	63 (18.8)
Mean age at biopsy — yr	51.4±14.3	49.9±14.8	53.9±12.6	57.8±12.3
Menopausal status at biopsy — no. of women (%)†				
Premenopausal (<45 yr)	2948 (32.4)	2246 (37.1)	652 (24.2)	50 (14.9)
Perimenopausal (45–55 yr)	2583 (28.4)	1610 (26.6)	871 (32.4)	102 (30.4)
Postmenopausal (>55 yr)	3556 (39.1)	2205 (36.4)	1167 (43.4)	184 (54.8)
Family history of breast cancer — no. of women (%)				
Unknown	4279 (47.1)	2970 (49.0)	1170 (43.5)	139 (41.4)
Known	4808 (52.9)	3091 (51.0)	1520 (56.5)	197 (58.6)
None	2668 (55.5)	1735 (56.1)	831 (54.7)	102 (51.8)
Weak	1174 (24.4)	756 (24.5)	378 (24.9)	40 (20.3)
Strong	966 (20.1)	600 (19.4)	311 (20.5)	55 (27.9)
Breast-cancer status as of August 2004 — no. of women (%)				
Negative	8380 (92.2)	5682 (93.7)	2426 (90.2)	272 (81.0)
Positive	707 (7.8)	379 (6.3)	264 (9.8)	64 (19.0)
Vital status — no. of women (%)				
Deceased	1827 (20.1)	1172 (19.3)	566 (21.0)	89 (26.5)
Alive	7260 (79.9)	4889 (80.7)	2124 (79.0)	247 (73.5)

* Plus–minus values are means ±SD.
† Menopausal status was categorized according to the age at breast biopsy.

eral cancer, we censored follow-up on women with contralateral cancer after the date of diagnosis. Similarly, when calculating the incidence of contralateral cancer, we censored follow-up on women with ipsilateral cancer after the date of diagnosis. Data on women missing information on the side of the cancer or women who had bilateral biopsies or cancer were not included in these analyses. This approach yields identical numbers of person-years for each type of event. As a result, the length of follow-up is no longer a factor in the analysis and the relative risks are equivalent to simple ratios of event counts. We therefore used properties of the binomial distribution to obtain exact P values and 95 percent confidence intervals for these relative risks.[12] Statistical tests were two-sided, and analyses were conducted with the use of SAS (SAS) and Splus (Insightful) software.

RESULTS

CHARACTERISTICS OF PATIENTS AND PATHOLOGICAL SPECIMENS

The final cohort consisted of 9087 women with benign breast disease as determined by open surgical biopsy. Table 1 shows the age at the time of the biopsy, likely menopausal status on the basis of age, and the strength of the family history of breast cancer according to the histologic findings for the benign lesion. The broad histologic classifications included nonproliferative disease in 6061 (66.7 percent), proliferative disease without atypia in 2690 (29.6 percent), and atypical hyperplasia in 336 (3.7 percent). Figure 1 shows examples of these lesions. The mean age was 51.4 years, but women with nonproliferative findings were slightly younger, whereas those with atypia tended to be older (mean age,

Table 2. Risk Factors for Breast Cancer after the Diagnosis of Benign Breast Disease.*

Characteristic	No. of Women	Person-Years	No. of Observed Events	No. of Expected Events	Relative Risk (95% CI)†
Overall	9087	144,881	707	453.0	1.56 (1.45–1.68)
Age at diagnosis of benign breast disease					
<30 yr	726	13,593	21	11.5	1.83 (1.13–2.80)
30–39 yr	1115	20,169	71	38.3	1.85 (1.45–2.34)
40–49 yr	2474	45,780	212	136.3	1.56 (1.35–1.78)
50–59 yr	2145	34,100	196	125.9	1.56 (1.35–1.79)
60–69 yr	1639	21,364	142	94.5	1.50 (1.27–1.77)
≥70 yr	988	9,874	65	46.6	1.40 (1.08–1.78)
Menopausal status‡					
Premenopausal (age <45 yr)	2948	54,419	169	106.1	1.59 (1.36–1.85)
Perimenopausal (age 45–55 yr)	2583	45,872	245	153.4	1.60 (1.40–1.81)
Postmenopausal (age >55 yr)	3556	44,590	293	193.6	1.51 (1.35–1.70)
Histologic findings					
Nonproliferative disease	6061	99,109	379	297.7	1.27 (1.15–1.41)
Proliferative disease without atypia	2690	41,610	264	140.2	1.88 (1.66–2.12)
Atypical hyperplasia	336	4,161	64	15.1	4.24 (3.26–5.41)
Family history of breast cancer§					
None	2668	44,974	171	145.4	1.18 (1.01–1.37)
Weak	1174	21,472	94	65.9	1.43 (1.15–1.75)
Strong	966	18,087	110	57.0	1.93 (1.58–2.32)

* Numbers of women, person-years, and events may not sum to overall totals because of rounding.
† The relative risk reflects the observed number of events as compared with the number expected on the basis of Iowa SEER data. All analyses account for the effects of age and calendar period. CI denotes confidence interval.
‡ Menopausal status was categorized according to the age at breast biopsy.
§ Information on family history was available for 4808 of the 9087 women.

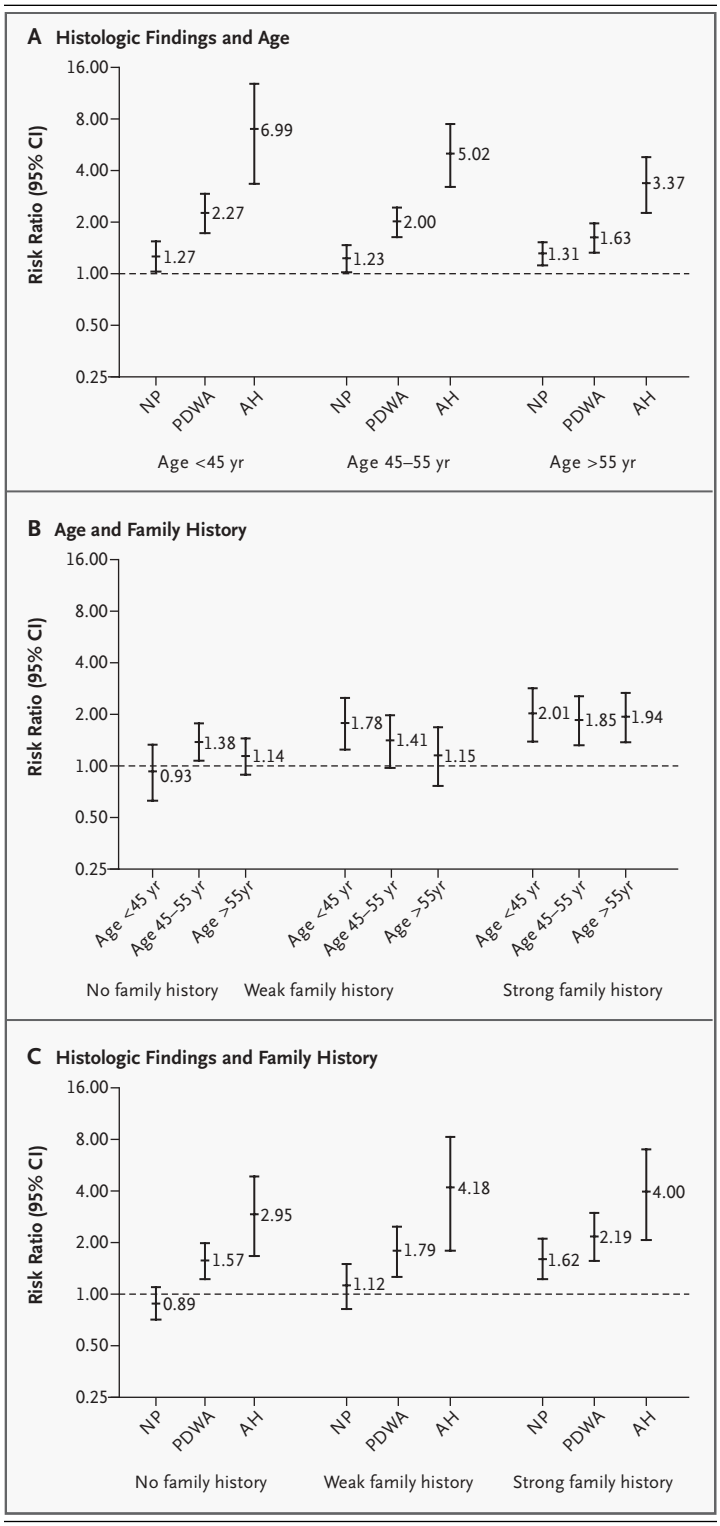


Figure 2. Risk-Factor Interaction Profiles for Benign Breast Disease, Comparing the Number of Events Observed with the Number Expected.

Expected events account for age and calendar period and are calculated with the use of Iowa SEER rates. CI denotes confidence interval, NP nonproliferative disease, PDWA proliferative disease without atypia, and AH atypical hyperplasia.

49.9 and 57.8 years, respectively; P<0.001). Information on family history was available for 4808 women and was negative in 2668 (55.5 percent), weakly positive in 1174 (24.4 percent), and strongly positive in 966 (20.1 percent). More women with atypia than without atypia had a strong family history of breast cancer (27.9 percent vs. 19.8 percent, P=0.06). The risk of cancer was highest in the group with atypia: breast cancer developed in 64 of the 336 women (19.0 percent).

FEATURES OF BENIGN BREAST DISEASE AND SUBSEQUENT RISK OF BREAST CANCER

Patients in the cohort were followed for a median of 15 years. A total of 1827 women (20.1 percent) had died and 7260 (79.9 percent) were alive as of August 2004. We have documented 707 breast cancers to date. The median time from the original biopsy to the diagnosis of breast cancer was 10.7 years. Table 2 shows the estimated relative risks of breast cancer associated with the age at the initial biopsy, the strength of the family history, menopausal status, and histologic findings of the biopsy, as compared with expected population-based incidence. The estimated relative risk of breast cancer in the cohort was 1.56 (95 percent confidence interval, 1.45 to 1.68). The risk was inversely associated with the age at biopsy, with younger women having a greater risk than older women. The type of benign breast disease identified at biopsy was a major predictor of risk. Atypical hyperplasia had a relative risk of 4.24 (95 percent confidence interval, 3.26 to 5.41), proliferative disease without atypia had a relative risk of 1.88 (95 percent confidence interval, 1.66 to 2.12), and nonproliferative lesions had a relative risk of 1.27 (95 percent confidence interval, 1.15 to 1.41). Family history was an independent risk factor. For women with no known family history of breast cancer, the relative risk was only 1.18 (95 percent confidence interval, 1.01 to 1.37), as compared with 1.43 (95 percent confidence interval, 1.15 to 1.75) for women with a weak family history and 1.93 (95 percent confidence interval, 1.58 to 2.32) for those with a strong family history.

Figure 2 shows possible interactions between pairs of the major risk factors of age, histologic findings, and family history. No significant interactions were observed between age and family history or between histologic findings and family history, including atypia and family history. However, there was a significant interaction between age and histologic findings (P=0.05): the risk of breast cancer was 6.99 times the expected risk among women who received a diagnosis of atypia before the age of 45 years; the risk was 5.02 times the expected risk when the atypia was diagnosed between the ages of 45 and 55 years and 3.37 times the expected risk when it was diagnosed after the age of 55 years. An important finding was that for women with nonproliferative disease and no family history or a weak family history, there was no increase in the risk of breast cancer.

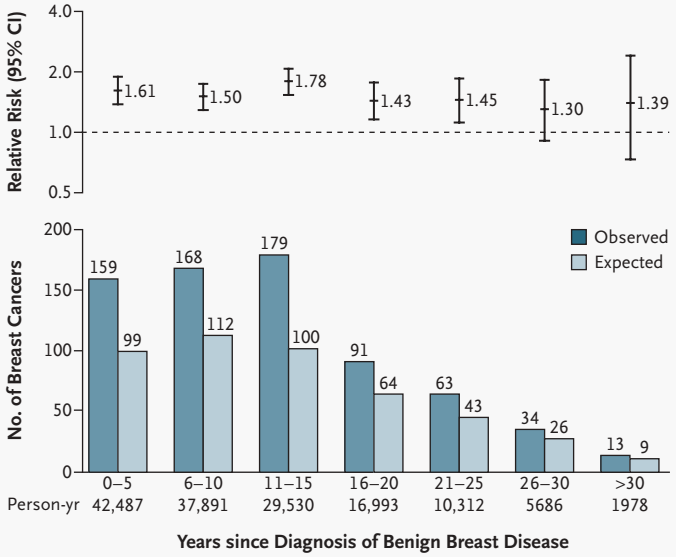


Figure 3. The Number of Breast Cancers Observed as Compared with the Number Expected over Time.

Expected events account for age and calendar period and are calculated with the use of Iowa SEER rates. CI denotes confidence interval.

TIME COURSE AND SIDE OF BREAST CANCER AFTER BENIGN BREAST DISEASE

Figure 3 shows the observed and expected numbers of cancers at five-year intervals. The excess risk persisted for at least 25 years after the initial biopsy and perhaps for 30 years or more, but accuracy was low after 25 years. Figure 4 shows a further breakdown of breast cancers into ipsilateral or contralateral according to the histologic findings in the benign lesion. Of the 616 unilateral cancers, 342 (55.5 percent) developed in the same breast as the initial biopsy and 274 (44.5 percent) developed in the contralateral breast. In the remaining 91 cases, there were bilateral events, either benign or malignant, or information on the side of the cancer was missing. During the first 10 years, there was an excess of ipsilateral cancers, with relative risks of ipsilateral as compared with contralateral cancer of 1.88 (95 percent confidence interval, 1.33 to 2.64) for years 0 through 5 and 1.34 (95 percent confidence interval, 0.96 to 1.85) for years 6 through 10. The 35 women with atypia in whom breast cancer developed within 10 years after the initial biopsy were 2.5 times as likely (P=0.02) to have the cancer in the same breast as in the opposite breast.

DISCUSSION

Retrospective and prospective studies have shown a relative risk of breast cancer of 1.5 to 1.6 for women with benign breast disease as compared with women in the general population.[2,5-7,13-21] The histologic appearance of the benign lesion is a major determinant of risk, yet not all large studies have had access to tissue for re-review. Our investigation was based on a single-institution resource with long-term and complete follow-up for cancer events. All samples containing the benign lesion were read by a breast pathologist who applied current histologic classifications. More than 700 breast cancers developed in this cohort, giving our study good statistical power. The relative risk of breast cancer for our cohort overall was 1.56 (95 percent confidence interval, 1.45 to 1.68), and this increased risk persisted for at least 25 years after the initial biopsy.

The histologic appearance of the benign lesion is strongly associated with the risk of breast cancer. For biopsies with nonproliferative findings, the relative risk was 1.27 (95 percent confidence interval, 1.15 to 1.41), as compared with a relative risk of 1.88 (95 percent confidence interval, 1.66 to 2.12) for findings of proliferative changes but no atypia and of 4.24 (95 percent confidence interval, 3.26 to 5.41) for a finding of atypical hyperplasia. When the family history is known, risk profiles can be refined. For women with nonproliferative findings and no family history or a weak family history of breast cancer, we observed no increased risk. This finding is important, because a sizable proportion

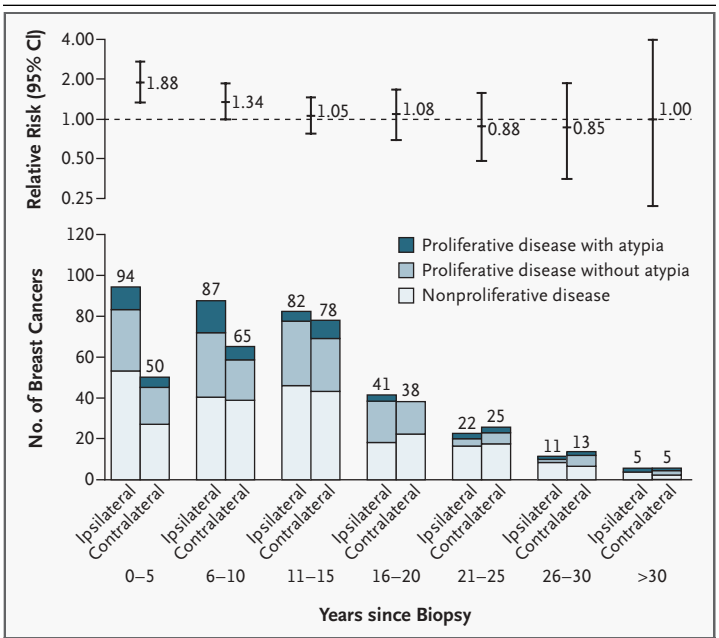


Figure 4. Comparison of the Number of Ipsilateral Breast Cancers with the Number of Contralateral Breast Cancers over Time, According to the Histologic Appearance of Benign Breast Disease.

Results are shown for 616 cancers (342 ipsilateral and 274 contralateral cancers). The remaining 91 cases include women with bilateral benign or malignant lesions or for whom the side of the benign or malignant lesion was unknown. CI denotes confidence interval.

of women with benign breast disease are in this group (52 percent of our cohort with a known family-history status). Dupont and Page made a similar observation in their 1985 report.[2] However, a recent NSABP study found a significantly increased risk of breast cancer among women with lower-category benign breast disease, including nonproliferative disease.[5] In the NSABP P1 trial, which included more than 13,000 women, 1376 had a breast biopsy with benign findings over a mean follow-up period of 79 months. Breast cancer developed in 47 of these women. On the basis of pathology reports from contributing centers, the investigators reported a relative risk of 1.6 among women with lower category findings on breast biopsy as compared with P1 participants who did not undergo a breast biopsy.[5]

In our study, the degree of family history was an independent risk factor. In women with a strong family history of breast cancer, even nonproliferative findings were associated with a risk ratio of 1.62. This subgroup may parallel the high-risk NSABP cohort.[5] Women with atypia are at significantly increased risk, but a family history did not significantly modify the atypia-associated risk (Fig. 2). The risk was four times the expected risk among women with atypia and a family history of breast cancer, regardless of the degree of their family history; among women with atypia without a family history of breast cancer, the risk ratio was 2.95 (95 percent confidence interval, 1.65 to 4.87).

The age at the diagnosis of benign breast disease appears to modify the risks related to the histologic appearance of benign breast disease. The presence of atypia in women under 45 years of age conveyed twice the risk observed among women over 55 years of age (6.99 and 3.37, respectively), which might relate, in part, to menopausal status. The Breast Cancer Detection and Demonstration Project showed that the risk of breast cancer among premenopausal women with atypia was elevated by a factor of 12.0 (95 percent confidence interval, 2.0 to 68.0), as compared with 3.3 among postmenopausal women with atypia (95 percent confidence interval, 1.1 to 10.0), but the numbers of patients in the study were small.[22] The Nurses Health Study also showed an increased risk of breast cancer among premenopausal women with atypia.[7] However, in the NSABP study of women with lower categories of benign breast disease, the risk of breast cancer was greatest among postmenopausal women.[5]

Understanding the risk associated with benign breast disease is important because the increasing use of mammography has increased the frequency of breast biopsies, most of which yield benign findings. In a retrospective study of women undergoing annual mammographic screening, Elmore et al. found that 18.6 percent of women underwent a biopsy after 10 screening mammograms.[23] The use of hormone therapy may also affect the frequency of breast biopsies. Chlebowski et al., reporting for the Women's Health Initiative investigators, found that relatively short-term therapy with estrogen plus progestin increased the percentage of women with abnormal mammograms, a major indicator for breast biopsy.[24]

Regarding the possibility of malignant precursors within benign breast disease, we have information on the side and the time to breast cancer for 616 unilateral events. An excess of breast cancers occurred in the same breast during the first years of follow-up, especially in women with atypia (Fig. 4). This finding suggests that precursors to breast cancer exist in benign breast disease. Work in model systems of early steps in mammary carcinogenesis

has identified alterations in key regulatory indicators that can be studied in selected benign breast lesions.[25,26]

In summary, our study shows that histologic features, the age at biopsy, and the degree of family history are major determinants of the risk of breast cancer after the diagnosis of benign breast disease. We found no increased risk among women with nonproliferative lesions, unless a strong family history was present. No significant interaction between atypia and family history was apparent. The excess risk of cancer in the ipsilateral breast in the first 10 years after the diagnosis of benign breast disease, especially in women with atypia, points to the presence of precursors in some women.

Supported by a Department of Defense Center of Excellence Grant (FEDDAMD17-02-1-0473-1), a grant (R01 CA46332) from the National Institutes of Health, a grant (BCTR99-3152) from the Susan G. Komen Breast Cancer Foundation, the Breast Cancer Research Foundation, and the Andersen Foundation.

We are indebted to Joel Worra and Dr. Piet de Groen for database development; to Teresa Allers, Mary Amundsen, Mary Campion, Lois Penheiter, and Romayne Thompson for data collection; and to Ann Harris and the Survey Research Center for patient follow-up.

REFERENCES

1. Connolly JL, Schnitt SJ. Benign breast disease: resolved and unresolved issues. Cancer 1993;71:1187-9.
2. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. N Engl J Med 1985;312:146-51.
3. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879-86.
4. Fitzgibbons PL, Henson DE, Hutter RV. Benign breast changes and the risk for subsequent breast cancer: an update of the 1985 consensus statement, Cancer Committee of the College of American Pathologists. Arch Pathol Lab Med 1998;122:1053-5.
5. Wang J, Costantino JP, Tan-Chiu E, Wickerham DL, Paik S, Wolmark N. Lower-category benign breast disease and the risk of invasive breast cancer. J Natl Cancer Inst 2004;96:616-20.
6. Carter CL, Corle DK, Micozzi MS, Schatzkin A, Taylor PR. A prospective study of the development of breast cancer in 16,692 women with benign breast disease. Am J Epidemiol 1988;128:467-77.
7. London SJ, Connolly JL, Schnitt SJ, Colditz GA. A prospective study of benign breast disease and the risk of breast cancer. JAMA 1992;267:941-4. [Erratum, JAMA 1992;267:1780.]
8. Schnitt SJ. Benign breast disease and breast cancer risk: potential role for antiestrogens. Clin Cancer Res. 2001;7:Suppl:4411s-4422s.
9. Melton LJ III. The threat to medical-records research. N Engl J Med 1997;337:1466-70.
10. Page DL, Dupont WD, Rogers LW, Rados MS. Atypical hyperplastic lesions of the female breast: a long-term follow-up study. Cancer 1985;55:2698-708.
11. Surveillance, Epidemiology, and End Results (SEER) Program. SEER statistics database. (Accessed June 27, 2005, at http://www.seer.cancer.gov.)
12. Bain LJ, Englehardt M. Introduction to probability and mathematical statistics. 2nd ed. Boston: PWS-Kent Publishing, 1992:369-77.
13. Page DL, Schuyler PA, Dupont WD, Jensen RA, Plummer WD Jr, Simpson JF. Atypical lobular hyperplasia as a unilateral predictor of breast cancer risk: a retrospective cohort study. Lancet 2003;361:125-9. [Erratum, Lancet 2003;361:1994.]
14. Dupont WD, Page DL. Breast cancer risk associated with proliferative disease, age at first birth, and a family history of breast cancer. Am J Epidemiol 1987;125:769-79.
15. Jensen RA, Page DL, Dupont WD, Rogers LW. Invasive breast cancer risk in women with sclerosing adenosis. Cancer 1989;64:1977-83.
16. Dupont WD, Page DL, Parl FF, et al. Long-term risk of breast cancer in women with fibroadenoma. N Engl J Med 1994;331:10-5.
17. Marshall LM, Hunter DJ, Connolly JL, et al. Risk of breast cancer associated with atypical hyperplasia of lobular and ductal types. Cancer Epidemiol Biomarkers Prev 1997;6:297-301.
18. Rohan TE, Hartwick W, Miller AB, Kandel RA. Immunohistochemical detection of c-erbB-2 and p53 in benign breast disease and breast cancer risk. J Natl Cancer Inst 1998;90:1262-9.
19. Kandel R, Li SQ, Ozcelik H, Rohan T. p53 Protein accumulation and mutations in normal and benign breast tissue. Int J Cancer 2000;87:73-8.
20. Kandel R, Zhu XL, Li SQ, Rohan T. Cyclin D1 protein overexpression and gene amplification in benign breast tissue and breast cancer risk. Eur J Cancer Prev 2001;10:43-51.
21. Krieger N, Hiatt RA. Risk of breast cancer after benign breast diseases: variation by histologic type, degree of atypia, age at biopsy, and length of follow-up. Am J Epidemiol 1992;135:619-31.
22. Dupont WD, Parl FF, Hartmann WH, et al. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 1993;71:1258-65.
23. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.
24. Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial. JAMA 2003;289:3243-53.
25. Crawford YG, Gauthier ML, Joubel A, et al. Histologically normal human mammary epithelia with silenced p16(INK4a) overexpress COX-2, promoting a premalignant program. Cancer Cell 2004;5:263-73.
26. Li JJ, Weroha SJ, Lingle WL, Papa D, Salisbury JL, Li SA. Estrogen mediates Aurora-A overexpression, centrosome amplification, chromosomal instability, and breast cancer in female ACI rats. Proc Natl Acad Sci U S A 2004;101:18123-8.

ORIGINAL ARTICLE

Atorvastatin in Patients with Type 2 Diabetes Mellitus Undergoing Hemodialysis

Christoph Wanner, M.D., Vera Krane, M.D., Winfried März, M.D., Manfred Olschewski, M.Sc., Johannes F.E. Mann, M.D., Günther Ruf, M.D., and Eberhard Ritz, M.D., for the German Diabetes and Dialysis Study Investigators*

From the Department of Medicine, Division of Nephrology, University of Würzburg, Würzburg, Germany (C.W., V.K.); the Clinical Institute of Medical and Chemical Laboratory Diagnostics, University General Hospital, Graz, Austria (W.M.); the Department of Medical Biometrics and Statistics, University Hospital of Freiburg, Freiburg, Germany (M.O.); Schwabing General Hospital, Munich, Germany (J.F.E.M.); Clinical Research Department, Pfizer, Karlsruhe, Germany (G.R.); and the Department of Medicine, University of Heidelberg, Heidelberg, Germany (E.R.). Address reprint requests to Dr. Wanner at the Department of Medicine, Division of Nephrology, University Hospital, Josef-Schneider-Str. 2, D-97080 Würzburg, Germany, or at wanner_c@medizin.uni-wuerzburg.de.

*Investigators and research coordinators participating in this study are listed in the Appendix.

N Engl J Med 2005;353:238-48.

ABSTRACT

BACKGROUND

Statins reduce the incidence of cardiovascular events in persons with type 2 diabetes mellitus. However, the benefit of statins in such patients receiving hemodialysis, who are at high risk for cardiovascular disease and death, has not been examined.

METHODS

We conducted a multicenter, randomized, double-blind, prospective study of 1255 subjects with type 2 diabetes mellitus receiving maintenance hemodialysis who were randomly assigned to receive 20 mg of atorvastatin per day or matching placebo. The primary end point was a composite of death from cardiac causes, nonfatal myocardial infarction, and stroke. Secondary end points included death from all causes and all cardiac and cerebrovascular events combined.

RESULTS

After four weeks of treatment, the median level of low-density lipoprotein cholesterol was reduced by 42 percent among patients receiving atorvastatin, and among those receiving placebo it was reduced by 1.3 percent. During a median follow-up period of four years, 469 patients (37 percent) reached the primary end point, of whom 226 were assigned to atorvastatin and 243 to placebo (relative risk, 0.92; 95 percent confidence interval, 0.77 to 1.10; P=0.37). Atorvastatin had no significant effect on the individual components of the primary end point, except that the relative risk of fatal stroke among those receiving the drug was 2.03 (95 percent confidence interval, 1.05 to 3.93; P=0.04). Atorvastatin reduced the rate of all cardiac events combined (relative risk, 0.82; 95 percent confidence interval, 0.68 to 0.99; P=0.03, nominally significant) but not all cerebrovascular events combined (relative risk, 1.12; 95 percent confidence interval, 0.81 to 1.55; P=0.49) or total mortality (relative risk, 0.93; 95 percent confidence interval, 0.79 to 1.08; P=0.33).

CONCLUSIONS

Atorvastatin had no statistically significant effect on the composite primary end point of cardiovascular death, nonfatal myocardial infarction, and stroke in patients with diabetes receiving hemodialysis.

Primary and secondary prevention trials, including those involving persons with diabetes mellitus, have documented substantial cardiovascular benefit from the administration of statins.[1,2] The recent Collaborative Atorvastatin Diabetes Study (CARDS) reported a decrease in deaths from cardiovascular causes among persons with type 2 diabetes mellitus in the absence of marked renal insufficiency.[3] There are no prospective data on the effects of statins in patients with end-stage renal disease with type 2 diabetes mellitus who are receiving hemodialysis, although type 2 diabetes is the most common diagnosis among patients at excessive risk of cardiovascular events[4] whose condition requires hemodialysis in both Germany[5] and the United States.[6] Abnormalities in serum lipid levels that are associated with renal disease rank high among the factors implicated in accelerated atherosclerosis.[7] However, not all the observational data on patients receiving hemodialysis link dyslipidemia with reduced rates of survival; indeed, opposite trends have been noted.[8] An observational retrospective analysis of patients receiving hemodialysis, the U.S. Renal Data System Morbidity and Mortality Study, Wave 2,[9] reported that the risk of death from cardiovascular causes was decreased by 36 percent among patients receiving statins, as compared with those who did not receive statins. There has been concern about the side effects of statins in patients receiving hemodialysis,[10] but data from small cohorts appeared to be reassuring.[11] The present investigator-initiated, prospective, randomized, placebo-controlled study of patients with type 2 diabetes mellitus receiving hemodialysis was designed to answer these questions.

METHODS

STUDY DESIGN

Subjects with type 2 diabetes mellitus 18 to 80 years of age who had been receiving maintenance hemodialysis for less than two years were enrolled at 178 centers in Germany. Exclusion criteria included levels of fasting serum low-density lipoprotein (LDL) cholesterol of less than 80 mg per deciliter (2.1 mmol per liter) or more than 190 mg per deciliter (4.9 mmol per liter), triglyceride levels greater than 1000 mg per deciliter (11.3 mmol per liter); liver-function values more than three times the upper limit of normal or equal to those in patients with symptomatic hepatobiliary cholestatic disease; hematopoietic disease or systemic disease unrelated to end-stage renal disease; vascular intervention, congestive heart failure, or myocardial infarction within the three months preceding the period of enrollment; unsuccessful kidney transplantation; and hypertension resistant to therapy (i.e., systolic blood pressure continuously greater than 200 mm Hg or diastolic blood pressure greater than 110 mm Hg). On enrollment, lipid-lowering medications were discontinued, and patients received placebo during the four-week run-in phase of the study. Thereafter, eligible patients were randomly assigned to double-blind treatment with either atorvastatin at a dose of 20 mg once daily or matching placebo. Data were recorded at four weeks and then every six months. The protocol was approved by the ethics committee at the coordinating center and the 29 regional institutional review boards. Specifically, the ethical implications of the inclusion of a placebo group — that is, of not providing lipid-lowering medications to those randomly assigned to the control group — were taken into account and considered acceptable. Written informed consent was obtained from all patients.

Academic investigators led, managed, and coordinated the study. The principal investigators wrote the protocol and prepared the manuscript. The data were monitored and collected by two contract research organizations supported by Pfizer, one of which (Datamap) holds the data. A university-based, independent statistician performed the statistical analyses. The plan for the statistical analysis was completed before the database was locked and unblinded.

A computer-generated randomization code was prepared by a central Pfizer unit that was independent of local study personnel. Medication was prepackaged on the basis of a block size of four subjects at each center. Each consecutive subject was given the next consecutive randomization number, and eligible patients were assigned in a 1:1 ratio to receive the study drug or placebo. Lipid levels measured after randomization were not released to the clinical sites. If LDL cholesterol levels fell below 50 mg per deciliter (1.3 mmol per liter), the dose of atorvastatin was reduced to 10 mg per day. To maintain blinding, a randomly selected subject from the placebo group received an identical dose reduction. One person in the central laboratory who had access to the randomization code controlled the changes in dose. After a patient reached a primary end point, the study drug could be replaced by treatment with

an active statin. Details of the study design have been described previously.[12,13]

END POINTS

The study end points and serious adverse events were continuously monitored and reported to the contract research organization. Every end point was adjudicated by three members of the end-point committee, on the basis of predefined criteria that are part of the study protocol. All analyses of primary and secondary end points were based on the classification by the end-point committee that was agreed on by consensus or majority vote. All committee members were blinded to the treatment assignments until August 13, 2004. The primary end point was a composite of death from cardiac causes, fatal stroke, nonfatal myocardial infarction, or nonfatal stroke, whichever occurred first. Only one event per subject was included in the analysis. Myocardial infarction was diagnosed when two of the following three criteria were met: typical symptoms; elevated levels of cardiac enzymes (i.e., a level of creatine kinase MB above 5 percent of the total level of creatine kinase, a level of lactic dehydrogenase 1.5 times the upper limit of normal, or a level of troponin T greater than 2 ng per milliliter); or diagnostic changes on the electrocardiogram. A resting electrocardiogram was recorded every six months and evaluated by independent cardiologists from the electrocardiographic monitoring board, according to the Minnesota classification system for the electrocardiogram (codes 1-1-1 through 9-2 for QRS-complex, ST-segment, or T-wave changes). An electrocardiogram that documented silent myocardial infarction was considered evidence of a primary end point.

Stroke was defined as a neurologic deficit lasting longer than 24 hours. Computed tomographic or magnetic resonance imaging of the brain was recommended and available in all but 16 cases. Death from cardiac causes comprised fatal myocardial infarction (death within 28 days after a myocardial infarction), sudden death, death due to congestive heart failure, death due to coronary heart disease during or within 28 days after an intervention, and all other deaths ascribed to coronary heart disease. Patients who died unexpectedly and did not present with a potassium level greater than 7.5 mmol per liter before the start of the three most recent sessions of hemodialysis were considered to have had sudden death from cardiac causes.

Secondary end points included death from all causes, all cardiac events combined, and all cerebrovascular events combined. Death from any cause other than cardiac disease or cerebrovascular disease was treated as a competing risk.

A central laboratory performed all the analyses. LDL cholesterol was measured directly by agarose-gel electrophoresis with subsequent enzymatic staining for cholesterol with the use of the rapid electrophoresis system (Helena Diagnostika). This method produces more accurate measurements of LDL cholesterol than ultracentrifugation and precipitation combined in samples with elevated triglyceride concentrations.[14]

STATISTICAL ANALYSIS

The study was designed to have 90 percent power to detect a 27 percent reduction in the incidence of the composite primary end point at an alpha level of 0.05 in a two-sided test, adjusted for one preplanned interim analysis according to an alpha-spending function based on the O'Brien–Fleming method, yielding a nominal level of significance for the final analysis of 0.045.[15] The alpha-spending function would have allowed for additional interim analyses, if necessary. For the study to have this level of power, at least 424 primary end points had to occur (event-driven analysis), requiring the randomization of at least 1200 patients. This calculation was based on observational studies.[16,17] The results were assessed in an intention-to-treat analysis. The primary end points were evaluated according to time-to-event analysis. Death from other causes was treated as a competing event, and for patients who died from other causes, follow-up was censored as of the date of death.[18] Times to an event for patients without a primary end point or competing event were treated as censored and were calculated as the time from randomization to the date of the last contact.

Cumulative incidence and Kaplan–Meier curves were used only to show the survival curves within the treatment groups and to calculate the corresponding survival probabilities. The Cox proportional-hazards model was used to estimate the multivariate relative risks of the primary and secondary end points with corresponding 95 percent confidence intervals. Adjustments were made for sex, age, and baseline status with respect to coronary heart disease. Unless otherwise stated, the baseline lipid and safety laboratory value was defined as the last value measured during the run-in period. The baseline

data were analyzed with the use of standard descriptive statistics.

RESULTS

PATIENTS

A total of 1255 subjects were randomly assigned to double-blind treatment with either atorvastatin (619) or placebo (636) between March 1998 and October 2002 and were followed until their final visit in March 2004 (Fig. 1). The two groups of patients were well matched with respect to baseline characteristics and concomitant therapy (Table 1). Nineteen percent of the patients had taken statins before entering the study. The mean length of follow-up was 3.96 years in the atorvastatin group and 3.91 years in the placebo group (median, 4.0 and 4.08 years, respectively).

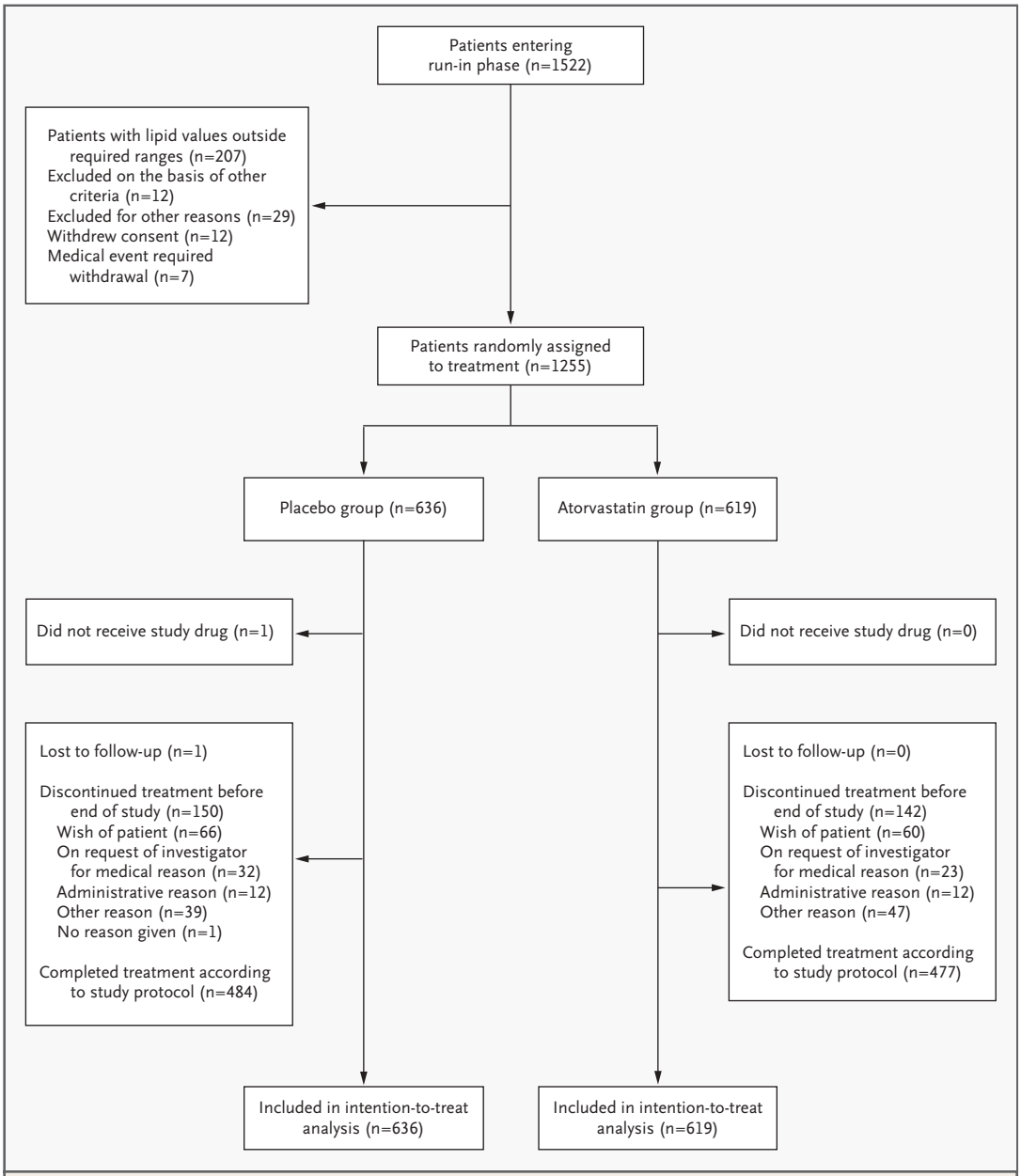


Figure 1. Numbers of Patients Who Entered the Study, Were Assigned to a Study Group, and Completed the Protocol.

Table 1. Baseline Characteristics of Patients in the Placebo and Atorvastatin Groups.*

Characteristic	Placebo Group (N=636)	Atorvastatin Group (N=619)
Age — yr	65.7±8.3	65.7±8.3
Female sex — no. (%)	292 (45.9)	286 (46.2)
Known duration of diabetes — yr	18.7±8.8	17.5±8.7
Time receiving dialysis — mo	8.4±6.9	8.2±6.9
Blood pressure — mm Hg		
Systolic	145±22	146±22
Diastolic	76±11	76±11
Current smoker — no. (%)	58 (9.1)	50 (8.1)
Former smoker — no. (%)	188 (29.6)	211 (34.1)
History of cardiovascular disease and intervention (%)†		
Myocardial infarction	17.3	17.9
Myocardial infarction, either CABG or PTCA, or coronary heart disease‡	28.1	30.7
Myocardial infarction or either CABG or PTCA	22.5	23.7
CABG or PTCA	11.8	14.2
Congestive heart failure§	34.9	35.9
Cardiac-valve disorder	7.7	7.3
Peripheral vascular disease	43.6	45.7
Stroke or TIA	18.2	17.4
Body-mass index¶	27.5±5.0	27.6±4.6
Hemoglobin — g/dl	10.9±1.4	10.9±1.3
Glycosylated hemoglobin — %	6.8±1.3	6.7±1.2
Albumin — g/liter	3.8±0.3	3.8±0.3
Calcium — mg/dl	9.2±0.8	9.2±0.8
Phosphate — mg/dl	6.1±1.6	6.0±1.6

LIPID LEVELS

At randomization, the median level of LDL cholesterol was 121 mg per deciliter (3.13 mmol per liter) in the atorvastatin group and 125 mg per deciliter (3.23 mmol per liter) in the placebo group. After four weeks, in the atorvastatin group, the median level of LDL cholesterol was 72 mg per deciliter (1.86 mmol per liter; median change from baseline, −42 percent). In the placebo group, the level of LDL cholesterol remained essentially unchanged (120 mg per deciliter [3.10 mmol per liter]; median change from baseline, −1.3 percent) (Fig. 2).

PRIMARY OUTCOMES

The cumulative incidence of the primary end point was 12.6 percent at one year and 31.9 percent at three years in the atorvastatin group, as compared with 11.2 percent and 30.5 percent, respectively, in the placebo group (Fig. 3). The relative risk reduction afforded by active treatment, as compared with placebo, was 8 percent (hazard ratio, 0.92; 95 percent confidence interval, 0.77 to 1.10; P=0.37). A similar number of patients died from cardiac causes in the two groups (20 percent in the atorvastatin group and 23 percent in the placebo group; relative risk, 0.81; 95 percent confidence interval, 0.64 to 1.03; P=0.08). Eleven percent (70) of the patients in the atorvastatin group had a nonfatal myocardial infarction, as compared with 12 percent (79) of those in the placebo group (relative risk, 0.88; 95 percent confidence interval, 0.64 to 1.21; P=0.42). More patients (27) died of stroke in the

Table 1. (Continued.)

Characteristic	Placebo Group (N=636)	Atorvastatin Group (N=619)
Lipid values — mg/dl		
Total cholesterol	220±42	218±43
LDL cholesterol	127±30	125±29
HDL cholesterol	36±14	36±13
Triglycerides	267±168	261±165
LDL cholesterol levels — no. (%)		
<100	120 (18.9)	122 (19.7)
100–129	241 (37.9)	252 (40.7)
130–159	186 (29.2)	169 (27.3)
≥160	89 (14.0)	76 (12.3)
Antihypertensive medication — %		
ACE inhibitors	47	49
Angiotensin II–receptor antagonists	12	12
Beta-blockers	38	37
Calcium antagonists	40	41
Antiplatelet therapy	50	54
Use of erythropoietin — %	81	81
Dose per wk — IU	6.225	6.202

* Plus–minus values are means ±SD. To convert hemoglobin values to millimoles per liter, multiply by 0.6206. To convert values for calcium to millimoles per liter, multiply by 0.250. To convert values for phosphate to millimoles per liter, multiply by 0.3229. To convert values for total cholesterol, low-density lipoprotein (LDL) cholesterol, and high-density lipoprotein (HDL) cholesterol to millimoles per liter, multiply by 0.02586. To convert values for triglycerides to millimoles per liter, multiply by 0.01129. CABG denotes coronary-artery bypass grafting, PTCA percutaneous transluminal coronary angioplasty, TIA transient ischemic attack, and ACE angiotensin-converting enzyme.
† Types of disease and intervention are not mutually exclusive.
‡ Disease was documented by coronary angiography.
§ Most of the patients had New York Heart Association class II heart failure.
¶ The body-mass index is the weight in kilograms divided by the square of the height in meters.

atorvastatin group than in the placebo group (13; relative risk, 2.03; 95 percent confidence interval, 1.05 to 3.93; P=0.04). Nonfatal stroke was distributed equally in the two groups (33 patients in the atorvastatin group and 32 patients in the placebo group; relative risk, 1.04; 95 percent confidence interval, 0.64 to 1.69; P=0.89) (Table 2).

SECONDARY OUTCOMES

Death from all causes was similar in the two groups (48 percent in the atorvastatin group, as compared with 50 percent in the placebo group; relative risk, 0.93; 95 percent confidence interval, 0.79 to 1.08; P=0.33). Of nominal significance, the risk of all cardiac events combined was reduced by 18 percent in the atorvastatin group, with a total event rate of 33 percent, as compared with 39 percent in the placebo group (relative risk, 0.82; 95 percent confidence interval, 0.68 to 0.99; P=0.03) (Table 2). This result was driven mainly by differences in the rates of coronary-artery bypass grafting and percutaneous transluminal coronary angioplasty. The incidence of all cerebrovascular events combined in the atorvastatin group was not different from that in the placebo group (relative risk, 1.12; 95 percent confidence interval, 0.81 to 1.55; P=0.49) (Table 2).

ADHERENCE, TOLERABILITY, AND ADVERSE EVENTS

The mean (±SD) duration of exposure to placebo was 27.2±17.9 months (range, 0.03 to 70.2), and to atorvastatin, 28.5±18.6 months (range, 0.07 to

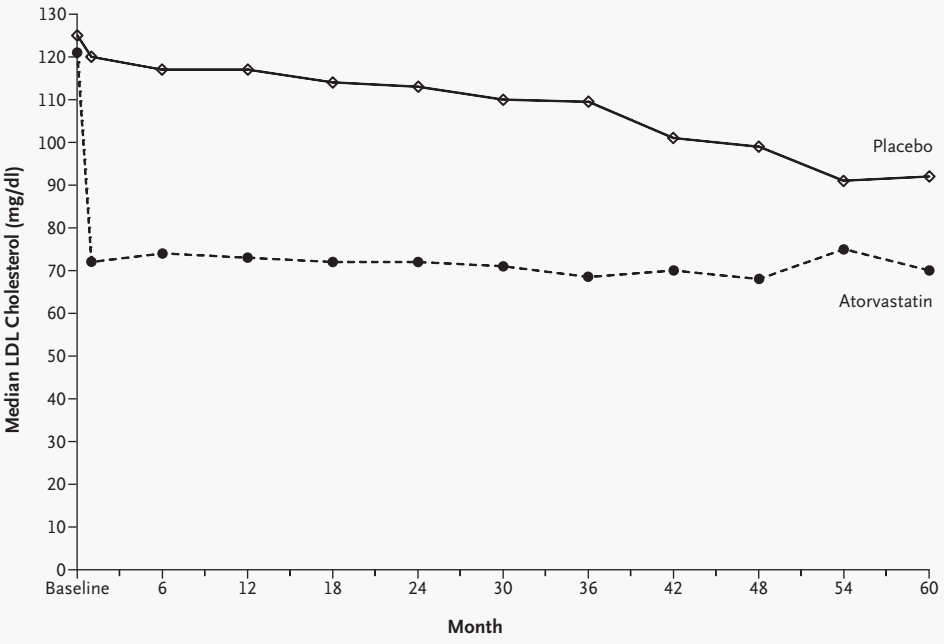


Figure 2. Median Level of Low-Density Lipoprotein (LDL) Cholesterol from Baseline to the End of the Study.
To convert values for LDL cholesterol to millimoles per liter, multiply by 0.02586.

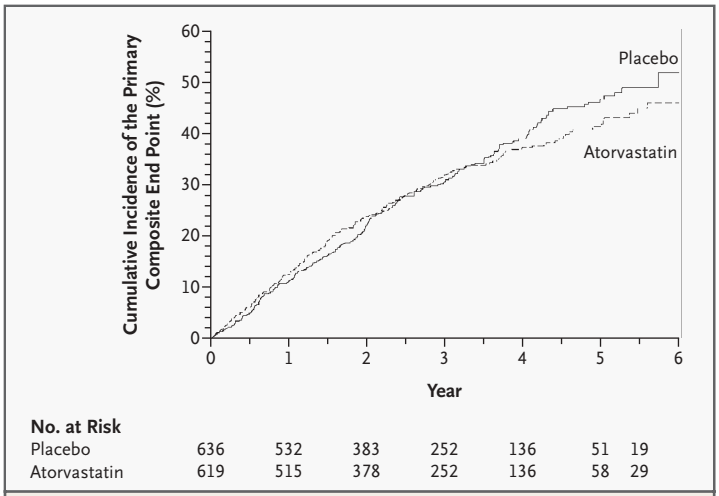


Figure 3. Estimated Cumulative Incidence of the Composite Primary End Point.

69.9). In the placebo group, 82 percent of patients took the study medication without interruption, and in the atorvastatin group, 80 percent of patients did so. The average number of days that treatment was interrupted was 12±36 in the placebo group and 13±40 in the atorvastatin group. During treatment, the dose of atorvastatin or matching placebo was halved when administered to 190 patients (15 percent). During the study, 98 patients in the placebo group (15 percent) began nonstudy statins, as compared with 10 percent of those in the atorvastatin group. The proportion of patients who continued to receive the study drug at one and two years, expressed as a percentage of those who remained alive and free of a primary event, was 74 percent (459 patients) and 51 percent (317 patients), respectively, in the atorvastatin group and 74 percent (469 patients) and 48 percent (303 patients), respectively, in the placebo group.

Table 2. Rates of Primary and Secondary End Points.*

End Point	Placebo Group (N=636)	Atorvastatin Group (N=619)	RR (95% CI)	P Value
	no. (%)			
Primary	243 (38)	226 (37)	0.92 (0.77–1.10)	0.37
Death from cardiac causes	149 (23)	121 (20)	0.81 (0.64–1.03)	0.08
Sudden death	83 (13)	77 (12)		
Fatal myocardial infarction	33 (5)	23 (4)		
Death due to congestive heart failure	24 (4)	17 (3)		
Death after interventions to treat coronary heart disease	4 (0.6)	3 (0.5)		
Other death due to coronary heart disease	5 (0.8)	1 (0.2)		
Nonfatal myocardial infarction	79 (12)	70 (11)	0.88 (0.64–1.21)	0.42
Silent	50 (8)	41 (7)		
Nonsilent	35 (6)	33 (5)		
Fatal stroke	13 (2)	27 (4)	2.03 (1.05–3.93)	0.04
Ischemic	7 (1)	18 (3)		
Hemorrhagic	5 (0.8)	3 (0.5)		
Other (not classified)	1 (0.2)	6 (1)		
Nonfatal stroke	32 (5)	33 (5)	1.04 (0.64–1.69)	0.89
Secondary				
All cardiac events combined	246 (39)	205 (33)	0.82 (0.68–0.99)	0.03
Death from cardiac causes	149 (23)	121 (20)		
Nonfatal myocardial infarction	79 (12)	70 (11)		
PTCA	45 (7)	34 (5)		
CABG	30 (5)	24 (4)		
Other interventions to treat coronary heart disease	0	1 (0.2)		
All cerebrovascular events combined	70 (11)	79 (13)	1.12 (0.81–1.55)	0.49
Stroke	44 (7)	59 (10)	1.33 (0.90–1.97)	0.15
Ischemic	33 (5)	47 (8)		
Hemorrhagic	8 (1)	5 (1)		
Other (not classified)	6 (1)	10 (2)		
TIA or PRIND	31 (5)	26 (4)		
Death from all causes	320 (50)	297 (48)	0.93 (0.79–1.08)	0.33
Death from causes other than cardiovascular or cerebrovascular disease	158 (25)	149 (24)	0.95 (0.76–1.18)	0.62
Fatal infection	68 (11)	60 (10)		
Fatal cancer	19 (3)	17 (3)		
Other	71 (11)	72 (12)		

* The total number of patients reaching the primary end point does not equal the sum of the numbers for each component of the primary end point, because only the first event per patient is included in the primary end point. Thus, a patient who had a stroke and a myocardial infarction was counted once in the primary end point, but appears in the separate totals for stroke and myocardial infarction. RR denotes relative risk, CI confidence interval, CABG coronary-artery bypass grafting, TIA transient ischemic attack, and PRIND prolonged reversible ischemic neurologic deficit.

Patients receiving hemodialysis generally have many adverse and serious adverse events (Table 3), but no cases of rhabdomyolysis or severe liver disease were detected in either group. The study medication was discontinued by the investigators in one patient receiving placebo because of a report of myalgia in combination with elevated creatine kinase levels.

DISCUSSION

We examined the value of lowering the level of LDL cholesterol in patients receiving hemodialysis who have type 2 diabetes mellitus, among whom the average annual incidence of myocardial infarction or death from coronary heart disease is 8.2 percent. This incidence rate exceeds the average annual rates of major coronary events that were reported in the placebo group of the Scandinavian Simvastatin Survival Study (6.6 percent) and is the highest rate of cardiovascular events in a long-term prospective trial of statin therapy.[19] Atorvastatin (20 mg daily) lowered LDL cholesterol levels by 42 percent, to 72 mg per deciliter, which is close to the target value of 70 mg per deciliter (1.81 mmol per liter) recommended by the Third Adult Treatment Panel of the National Cholesterol Education Program for persons at very high risk of cardiovascular disease. Despite the high rate of cardiovascular events and the pronounced LDL cholesterol–lowering activity of atorvastatin, a significant reduction in the incidence of the composite primary end point was not achieved.

Of nominal significance, more cases of fatal stroke occurred in the atorvastatin group (27) than in the placebo group (13). This finding is unexplained and could be a chance finding, particularly in view of the data from CARDS, which indicate that atorvastatin lowers the incidence of stroke.[3] That study reported a relative risk for stroke of 0.52 (95 percent confidence interval, 0.31 to 0.89) in persons with type 2 diabetes mellitus who were taking atorvastatin. The rate of fatal and nonfatal stroke decreased from 2.8 to 1.5 percent (39 vs. 21 patients), whereas in the present study, it increased from 7.0 to 9.7 percent (44 vs. 59 patients).

The complete absence of a stroke benefit and the increase in fatal strokes contribute considerably to the finding that the treatment effect on the primary end point was less than predicted. A possible reason for the unexpected results with regard to the primary end point might be related to the LDL cholesterol concentration at baseline. In general, the absolute risk reduction attained by lowering LDL cholesterol by a given percentage is less when pretreatment concentrations are low than when they are high.[20] The baseline levels of LDL cholesterol among patients in our study were, on average, above the target (126 mg per deciliter [3.25 mmol per liter]). Given the log-linear relation between LDL cholesterol and coronary heart disease, reducing levels of LDL cholesterol by 40 percent from a starting level of 125 mg per deciliter would result in an approximate relative risk reduction of 30 percent or more.[20] This estimate is empirically supported by the results of CARDS[3] and the British Heart Protection Study[21] and is very close to our initial assumption of a risk reduction of 27 percent.

Since we did not fully achieve this benefit, we speculate that the pathogenesis of vascular events in patients with diabetes mellitus who are receiving hemodialysis may, at least in part, be different from that in patients without end-stage renal disease. Subgroup analyses showed no difference in outcome for any LDL cholesterol level or patients with and patients without cardiovascular disease. Interestingly, there was a continuous decrease in LDL

Table 3. Adverse Events.*

Event	Placebo Group	Atorvastatin Group
	no. of events	
Total	2255	2276
Serious events	1060	1073
Events requiring hospitalization	942	949
Events requiring discontinuation of study drug	52	73
Drug-related serious events	1	1
Diagnosis of cancer	44	39
Severe hyperkalemia	9	3
Severe hypoglycemia	4	6
Ventricular fibrillation or tachycardia	13	7
Myalgia or myopathy	5	7
Creatine kinase level		
3 to 5 times the upper limit of normal	3	11
>5 to 10 times the upper limit of normal	1	1
Alanine aminotransferase level >4 times the upper limit of normal	1	5

* Some patients had more than one event.

cholesterol levels over time among patients in both groups. Some malnutrition cannot be ruled out during the course of the study, although there was no decrease in the body-mass index.

The extremely high rate of death from cardiovascular causes among patients receiving dialysis[22] is explained by more than the traditional coronary risk factors. Apart from the presence of many aggravating coexisting factors, such as inappropriate left ventricular hypertrophy, cardiac fibrosis, cardiac microvessel disease,[23] and sympathetic overactivity, among others, there are also indications that atherosclerosis itself is promoted by risk factors other than the traditional cardiovascular risk factors.[24,25] The most plausible explanation for the absence of a significant effect on mortality from cardiac causes and cardiac end points in this study is the presence of additional pathogenetic pathways in cardiovascular disease. The dose of atorvastatin in the present study was 20 mg, which is lower than the high dose used in a recent study by LaRosa et al.[26] in which intensive lipid-lowering therapy with atorvastatin at a dose of 80 mg per day was more effective than a dose of 10 mg per day in patients with stable coronary heart disease. However, whether such an advantage would accrue if patients with type 2 diabetes who were receiving dialysis were given a higher dose of atorvastatin is unknown.

Several important conclusions can be drawn from this study. First, we showed that it is difficult to rely on uncontrolled observational studies that show substantial advantages of statins in the treatment of patients receiving hemodialysis.[9,27] Second, and more important, is the conclusion that the benefit of atorvastatin is limited when intervention with statins is postponed until patients have reached end-stage renal disease. Subgroup analyses of major statin-intervention trials documented a cardiovascular benefit in patients with chronic kidney disease (stages 1, 2, and 3 according to the classification of the National Kidney Foundation).[28,29] According to CARDS, lowering LDL cholesterol levels early during the clinical course of type 2 diabetes mellitus is of benefit.[3] Third, there was no excess of serious adverse events; specifically, no cases of rhabdomyolysis occurred, but we found a nominally significant increase in fatal stroke.

We conclude that in persons with type 2 diabetes mellitus who are receiving maintenance hemodialysis and have LDL cholesterol values between 80 and 190 mg per deciliter, routine treatment with a statin to reduce the primary composite end point of death from cardiac causes, myocardial infarction, and stroke is not warranted. The initiation of lipid-lowering therapy in patients with type 2 diabetes mellitus who already have end-stage renal disease may come too late to translate into consistent improvement of the cardiovascular outcome.

Supported by Pfizer. The committee members and investigators did not receive remuneration for conducting the study, except for reimbursement of costs to participate in scientific meetings.

Dr. Wanner reports having received consulting fees and lecture fees from Genzyme; Dr. März, consulting fees, lecture fees, a research grant and stock options from Pfizer; and Dr. Mann, lecture fees from Aventis, Roche, and Janssen Cilag. Dr. Ritz is a member of the safety board of a trial sponsored by AstraZeneca and reports having received consulting fees from the company.

We are indebted to the German Association for Clinical Nephrology (K.-W. Kühn, chair) and the Association of German Nephrology Centers (H. Kütemeyer, chair).

APPENDIX

The following investigators and research coordinators participated in the study known as the 4D Study (a complete list is available at www.uni-wuerzburg.de/ nephrologie): *Steering committee:* C. Wanner, E. Ritz. *Clinical coordinator:* V. Krane. *Medical end-point monitors:* Z. Ülger, F. Swoboda. *Data and safety monitoring committee:* M. Wehling (chair), E. Keller (deceased), M. Schumacher, T. Eschenhagen. *Event committee:* J. Mann (chair), J. Bommer, P. Schanzenbächer, P. Schollmeyer, M. Schartl. *Electrocardiography monitoring board:* F. Heinrich, H. Mörl. *Biometric and statistical analysis:* University of Freiburg, M. Olschewski. *Central laboratory (lipid and safety core laboratory):* University of Freiburg, W. März. *Contract research organization:* Kendle, Munich, S. Reichmuth (*Project manager*); Datamap, Freiburg, J. Lilienthal. *Sponsor:* Pfizer, Karlsruhe, G. Ruf, B. Rauer (*Project manager*).

REFERENCES

1. Cheung BM, Lauder IJ, Lau CP, Kumana CR. Meta-analysis of large randomized controlled trials to evaluate the impact of statins on cardiovascular outcomes. Br J Clin Pharmacol 2004;57:640-51.
2. Armitage J, Bowman L. Cardiovascular outcomes among participants with diabetes in the recent large statin trials. Curr Opin Lipidol 2004;15:439-46.
3. Colhoun HM, Betteridge DJ, Durrington PN, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet 2004;364:685-96.
4. Schwenger V, Hofmann A, Kalifeh N, et al. Uremic patients — late referral, early death. Dtsch Med Wochenschr 2003;128:1216-20. (In German.)
5. Koch M, Thomas B, Tschöpe W, Ritz E. Survival and predictors of death in dialysed diabetic patients. Diabetologia 1993;36:1113-7.
6. Renal Data System. USRDS 2003 annual data report: atlas of end-stage renal disease in the United States. Bethesda, Md.: National Institute of Diabetes and Digestive and Kidney Disease, 2003. (Accessed June 27, 2005, at http://www.usrds.org.)

7. Prichard SS. Impact of dyslipidemia in end-stage renal disease. J Am Soc Nephrol 2003;14:Suppl 4:S315-S320.
8. Iseki K, Yamazato M, Tozawa M, Takishita S. Hypocholesterolemia is a significant predictor of death in a cohort of chronic hemodialysis patients. Kidney Int 2002;61:1887-93.
9. Seliger SL, Weiss NS, Gillen DL, et al. HMG-CoA reductase inhibitors are associated with reduced mortality in ESRD patients. Kidney Int 2002;61:297-304.
10. Baigent C, Burbury K, Wheeler D. Premature cardiovascular disease in chronic renal failure. Lancet 2000;356:147-52.
11. Lins RL, Matthys KE, Verpooten GA, et al. Pharmacokinetics of atorvastatin and its metabolites after single and multiple dosing in hypercholesterolaemic haemodialysis patients. Nephrol Dial Transplant 2003;18:967-76.
12. Wanner C, Krane V, Ruf G, März W, Ritz E. Rationale and design of a trial improving outcome of type 2 diabetics on hemodialysis. Kidney Int Suppl 1999;71:S222-S226.
13. Wanner C, Krane V, März W, et al. Randomized controlled trial on the efficacy and safety of atorvastatin in patients with type 2 diabetes on hemodialysis (4D Study): demographic and baseline characteristics. Kidney Blood Press Res 2004;27:259-66.
14. Nauck M, Winkler K, März W, Wieland H. Quantitative determination of high-, low-, and very-low-density lipoproteins and lipoprotein(a) by agarose gel electrophoresis and enzymatic cholesterol staining. Clin Chem 1995;41:1761-7.
15. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35:549-56.
16. Tschöpe W, Koch M, Thomas B, Ritz E. Serum lipids predict cardiac death in diabetic patients on maintenance hemodialysis: results of a prospective study. Nephron 1993;64:354-8.
17. Koch M, Kuthuhn B, Trenkwalder E, et al. Apolipoprotein B, fibrinogen, HDL cholesterol, and apolipoprotein(a) phenotypes predict coronary artery disease in hemodialysis patients. J Am Soc Nephrol 1997;8:1889-98.
18. Schulgen G, Olschewski M, Krane V, Wanner C, Ruf G, Schumacher M. Sample sizes for clinical trials with time-to-event endpoints and competing risks. Contemp Clin Trials 2005;26:386-96.
19. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet 1994;344:1383-9.
20. Grundy SM, Cleeman JI, Merz CNB, et al. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. Circulation 2004;110:227-39. [Erratum, Circulation 2004;110:763.]
21. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. Lancet 2002;360:7-22.
22. Foley RN, Parfrey PS, Sarnak MJ. Clinical epidemiology of cardiovascular disease in chronic renal disease. Am J Kidney Dis 1998;32:Suppl 3:S112-S119.
23. Amann K, Breitbach M, Ritz E, Mall G. Myocyte/capillary mismatch in the heart of uremic patients. J Am Soc Nephrol 1998;9:1018-22.
24. Takayama F, Aoyama I, Tsukushi S, et al. Immunohistochemical detection of imidazolone and N(epsilon)-(carboxymethyl) lysine in aortas of hemodialysis patients. Cell Mol Biol 1998;44:1101-9.
25. Himmelfarb J, Stenvinkel P, Ikizler TA, Hakim RM. The elephant in uremia: oxidant stress as a unifying concept of cardiovascular disease in uremia. Kidney Int 2002;62:1524-38.
26. LaRosa JC, Grundy SM, Waters DD, et al. Intensive lipid lowering with atorvastatin in patients with stable coronary disease. N Engl J Med 2005;352:1425-35.
27. Mason NA, Bailie GR, Satayathum S, et al. HMG-coenzyme A reductase inhibitor use is associated with mortality reduction in hemodialysis patients. Am J Kidney Dis 2005;45:119-26.
28. Tonelli M, Isles C, Curhan GC, et al. Effect of pravastatin on cardiovascular events in people with chronic kidney disease. Circulation 2004;110:1557-63.
29. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002;39:Suppl 1:S1-S266.

CORRECTION

Atorvastatin in Patients with Type 2 Diabetes Mellitus Undergoing Hemodialysis

Atorvastatin in Patients with Type 2 Diabetes Mellitus Undergoing Hemodialysis . On page 243, in the continuation of Table 1, the dose of erythropoietin per week (IU) should have read ``6225´´ for the placebo group and ``6202´´ for the atorvastatin group, rather than ``6.225´´ and ``6.202,´´ respectively, as printed.

N Engl J Med 2005;353:1640-a

BRIEF REPORT

Hyperinsulinemic Hypoglycemia with Nesidioblastosis after Gastric-Bypass Surgery

Geoffrey J. Service, M.D., Geoffrey B. Thompson, M.D., F. John Service, M.D., Ph.D., James C. Andrews, M.D., Maria L. Collazo-Clavell, M.D., and Ricardo V. Lloyd, M.D., Ph.D.

SUMMARY

We describe six patients (five women and one man; median age, 47 years; range, 39 to 54) with postprandial symptoms of neuroglycopenia owing to endogenous hyperinsulinemic hypoglycemia after Roux-en-Y gastric bypass surgery. Except for equivocal evidence in one patient, there was no radiologic evidence of insulinoma. Selective arterial calcium-stimulation tests, positive in each patient, were used to guide partial pancreatectomy. Nesidioblastosis was identified in resected specimens from each patient, and multiple insulinomas were identified in one. Hypoglycemic symptoms diminished postoperatively. We speculate that hyperfunction of pancreatic islets did not lead to obesity but that beta-cell trophic factors may have increased as a result of gastric bypass.

A CONSEQUENCE OF THE OBESITY EPIDEMIC IN THE UNITED STATES[1] IS the increasing use of gastric bypass surgery for patients with severe, medically complicated obesity.[2] Some patients who have undergone this procedure have postprandial symptoms that have been ascribed to rapid emptying of gastric contents.[3] This phenomenon, referred to as the dumping syndrome, is characterized by vasomotor symptoms of diaphoresis, weakness, dizziness, and flushing, but not neuroglycopenia.[4] In the past five years, we have treated six patients in whom postprandial symptoms of neuroglycopenia developed as a result of endogenous hyperinsulinemic hypoglycemia after gastric bypass. Their clinical presentation typified that of the noninsulinoma pancreatogenous hypoglycemia syndrome (postprandial neuroglycopenic hyperinsulinemic hypoglycemia and pancreatic nesidioblastosis)[5,6] and is also seen in some patients with insulinoma.[7] We attempted to determine whether hyperfunction of pancreatic islets as a result of nesidioblastosis, which is characteristic of the noninsulinoma pancreatogenous hypoglycemia syndrome, or insulinoma was the basis for the hypoglycemia and to determine the possible role of gastric bypass in the genesis of the abnormal islets.

From the Departments of Surgery (G.J.S., G.B.T.), Radiology (J.C.A.), and Pathology (R.V.L.) and the Division of Endocrinology, Metabolism, and Nutrition (F.J.S., M.L.C.-C.), Mayo Clinic and Foundation, Rochester, Minn. Address reprint requests to Dr. F. John Service at the Division of Endocrinology, Metabolism, and Nutrition, Mayo Clinic, 200 First St. SW, Rochester, MN 55905, or at service.john@mayo.edu.

N Engl J Med 2005;353:249-54.

METHODS

SUBJECTS

From 2000 to 2004, six patients (five women and one man; median age, 47 years; range, 39 to 54) who had undergone Roux-en-Y gastric bypass for extreme obesity were referred for evaluation of repeated episodes of postprandial hypoglycemia associat-

ed with symptoms of profound neuroglycopenia, which could not be controlled by lifestyle modification. Their median body-mass index (the weight in kilograms divided by the square of the height in meters) at the time of this evaluation was 28.2 (range, 23.1 to 39.5), in contrast to a median body-mass index of 50.1 (range, 44.4 to 62.5) before bypass, representing a median loss of 44 percent (range, 20 to 51 percent) of the preoperative weight. In the bypass circuit the alimentary limbs were a median of 100 cm in length (range, 80 to 250) from the anastomosis. Historical details regarding the development of symptoms in relation to gastric bypass and meals and the type of symptoms are provided in Table 1. The timing of symptoms in relation to gastric bypass was self-reported, since there was no documentation of hypoglycemia until the present evaluation. One patient recalled symptoms antedating gastric bypass that had worsened considerably after the procedure. Confirmation of the history of postprandial hypoglycemia was obtained by waiting for a spontaneous episode to occur, at which point venous blood samples were obtained either at our facility (four patients) or elsewhere (two patients). All laboratory analyses were performed at our facility.

Data relevant to the episodes of spontaneous postprandial hypoglycemia are shown in Table 1. Each patient was confirmed to have endogenous postprandial (one to four hours after eating) hyperinsulinemic hypoglycemia, defined as a serum insulin level of at least 3 μU per milliliter (18 pmol per liter) and a serum C-peptide level of at least 0.6 ng per milliliter (199 pmol per liter) with a concomitant serum glucose level of less than 55 mg per deciliter (3.1 mmol per liter) and the absence of sulfonylurea in the plasma (measured in five of the six patients).[7] Although postprandial hypoglycemia is an unusual occurrence in patients with insulinoma, the fact that it is a possibility warranted radiologic localization procedures, such as triple-phase spiral computed tomography and transabdominal ultrasonography of the pancreas. Endoscopic ultrasonography is often neither feasible nor useful in patients with a small, remnant gastric pouch. When the results of these conventional imaging procedures are negative or equivocal, the selective arterial calcium-stimulation test is recommended for the identification and regionalization of potentially hyperfunctioning beta cells.

LABORATORY ANALYSES

Serum levels of insulin[8] and C peptide[9] were measured with the use of immunochemiluminometric assays with a lower limit of detection of 0.1 μU per milliliter (0.6 pmol per liter) and 0.1 ng per milliliter (33 pmol per liter), respectively. Plasma sulfonylurea levels were measured by liquid chro-

Table 1. Historical Symptoms and Laboratory Values Obtained during Episodes of Spontaneous Postprandial Hypoglycemia.*

Patient No.	Historical Symptoms			Observed Episode of Spontaneous Postprandial Hypoglycemia			
	Timing after Gastric Bypass	Timing after Meals†	Type	Serum Glucose	Serum Insulin	Serum C Peptide	Plasma Sulfonylurea
	yr	*hr*		*mg/dl*	*μU/ml*	*ng/ml*	
1	2	1–3	Confusion	53	16.0	1.8	Not measured
2	1	2–3	Confusion	38	4.2	3.3	Undetectable
3	0.5	2	Loss of consciousness	31	28.0	1.4	Undetectable
4	1	1.5–2	Loss of consciousness	42	8.3	7.6	Undetectable
5	—‡	2	Confusion	39	3.1	3.6	Undetectable
6	8	1	Tunnel vision	44	3.4	3.7	Undetectable

* Criteria for endogenous hyperinsulinemic hypoglycemia are a serum insulin level of at least 3 μU per milliliter and a C peptide level of at least 0.6 ng per milliliter with a concomitant serum glucose level of less than 55 mg per deciliter and the absence of sulfonylurea in the plasma. 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 C peptide to picomoles per liter, multiply by 331.

† The timing is in the context of ad libitum feeding.

‡ This patient reported having symptoms three years before gastric bypass.

matographic tandem mass spectroscopy.[10,11] Serum glucose levels were measured according to a standard hexokinase method on a Hitachi Chemistry Analyzer (model 747-200, Roche Diagnostics).[12]

Selective Arterial Calcium-Stimulation Test

When selective arterial calcium-stimulation tests are performed as previously described,[5] a doubling of the basal insulin level in the right hepatic vein in response to the sequential injection of 0.025 mEq of calcium per kilogram of body weight into the splenic, superior mesenteric, and gastroduodenal arteries is considered to indicate hyperfunction of the beta cells in the vascular distribution of the artery studied. Although overlap can occur across vascular territories, such overlap can be identified from the angiographic findings. In general, the body and tail of the pancreas are within the splenic-artery distribution; the head and, secondarily, the uncinate process are within the gastroduodenal-artery distribution; and the uncinate process and, secondarily, the head are within the distribution of the superior mesenteric artery. The pattern of response to the intraarterial injection of calcium is expressed as a difference in response (gradient) from artery or arteries with a positive response to artery or arteries with a negative response. For instance, a positive response to the injection of calcium into the splenic artery, but not to injections into the superior mesenteric or gastroduodenal arteries, creates a gradient between the former and latter arterial distributions that can guide the surgeon in the extent of pancreatic resection.

Pathological Analysis

The resected pancreatic tissues were sectioned into 1-mm slices, fixed in buffered formalin, embedded in paraffin, and stained with hematoxylin and eosin. Histologic criteria for nesidioblastosis in adults (hypertrophic beta cells within enlarged or normal-appearing islets; small, scattered clusters of endocrine cells; and ductuloinsular complexes)[13] were used to evaluate the sections.

Immunohistochemical staining was performed on paraffin sections 5 μm thick with the use of the avidin–biotin–peroxidase complex system as previously described.[14] The antibodies used included chromogranin A (Boehringer Mannheim) at a dilution of 1:100, insulin (Dako) at a dilution of 1:750, glucagon (Dako) at a dilution of 1:3000, somatostatin (Dako) at a dilution of 1:1000, and gastrin (Dako) at a dilution of 1:1000. Positive controls for immunostaining consisted of normal pancreatic tissues. Negative controls consisted of sections in which the primary antibody had been omitted during the immunostaining procedure. The immunostained slides were analyzed for staining within the islets, ducts, and interacinar locations within the pancreas.

The size of the islets in each patient was estimated by measuring 50 islets per patient with a micrometer in the ocular of the microscope; the largest islet in each of the 50 randomly selected fields was selected. Then, the mean diameter of the largest islets was calculated. Pancreatic tissues with normal-sized islets from four obese patients (three women and one man; median body-mass index, 34.1; range, 33.2 to 36.3) without known endocrine disease were also analyzed as controls.

CONDUCT OF THE STUDY

Each author vouches for the data and analyses. Four provided direct care of each patient; one contributed to the bariatric aspects of the manuscript; and one, a surgical resident, assisted in the design, data acquisition, and writing of the manuscript. The institutional review board approved this minimal-risk study with waiver of informed consent in accordance with the Code of Federal Regulations (45 CFR 46.116), as noted in the Federal Policy for the Protection of Human Subjects from the Department of Health and Human Services.

RESULTS

CLINICAL INVESTIGATION

During spontaneous postprandial episodes, each patient had hypoglycemia related to endogenous hyperinsulinemia (Table 1). Because of the long half-life of C peptide (30 minutes), a finding of elevated postprandial levels is not proof of excessive secretion by beta cells; rather, it rules out the possibility of exogenous insulin administration, in which case the levels would be low or undetectable. Patient 1 had equivocal evidence of insulinoma on triple-phase spiral CT and transabdominal ultrasonography of the pancreas; the results of these procedures were entirely negative in the other five patients. Each patient therefore underwent selective arterial calcium-stimulation testing, with positive responses in one, two, or three arterial distributions (Table 2). Each patient underwent pancreatic exploration with complete mobilization and pal-

Table 2. Results of Selective Arterial Calcium Stimulation Tests.*

Artery Injected	Peak Hepatic-Vein Insulin					
	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
	multiple of the basal level					
Splenic	2.0	5.9	3.7	6.4	3.6	2.6
Superior mesenteric	1.0	2.5	2.8	1.1	2.9	1.1
Gastroduodenal	1.0	4.5	1.9	3.7	2.1	2.9

* Multiples of 2 or greater are considered to indicate a positive response.

pation of the pancreas and intraoperative ultrasonography. The latter showed an insulinoma in Patient 1 in the tail of the pancreas, as suggested by the preoperative radiologic assessment. Ultrasonography revealed no abnormalities in the other five patients.

SURGERY

Patient 1 underwent a spleen-preserving distal pancreatectomy. No insulinoma was detected by palpation and intraoperative ultrasonography in the other five patients; therefore, with guidance by the gradient between positive and negative responses in the selective arterial calcium-stimulation test, distal pancreatectomy in which the superior mesenteric vein was used as the distinguishing landmark was performed. When the positive response to the intraarterial injection of calcium was confined to the splenic-artery distribution, the distal resection was carried to the left of the superior mesenteric vein. When the distribution of the gastroduodenal artery, superior mesenteric vein, or both were also involved, a safe, subtotal distal resection was performed to the right of the superior mesenteric vein. With the use of this gradient guidance, Patients 2, 4, 5, and 6 underwent extended distal pancreatectomy. Although warranted in the case of Patient 3, dissection to the right of the superior mesenteric vein was thought to be unsafe because it would place the retrogastric Roux-en-Y limb at risk for devitalization; therefore, a distal pancreatectomy was performed.

PATHOLOGICAL FINDINGS

Patient 1 had multiple islet-cell tumors, some of which stained positive for insulin and were therefore considered to be functional insulinomas. Patient 5 had a 0.4-cm, nonfunctional islet-cell tumor with no staining for pancreatic islet-cell hormones but with staining for chromogranin A (a general neuroendocrine marker). The islets in most patients showed a variable pattern of islet-cell hypertrophy and hyperplasia. The mean (±SE) size of the islets was significantly larger in all the patients than in the obese controls (214±7.7 vs. 151±7.3 μm, P=0.001) (Fig. 1A and 1B).

Immunohistochemical staining of the islets for insulin, glucagon, somatostatin, and chromogranin A indicated that 60 to 80 percent of islets from both patients and controls were positive for insulin. Insulin-positive cells budding off the pancreatic ducts were noted in Patients 2, 3, 4, 5, and 6 and were more numerous (median, 7 per slide; range, 1 to 11) than in the controls (median, 3 per slide; range, 1 to 5), although the difference was not significant (Fig. 1C). Chromogranin A staining revealed more cells budding off the ducts than did insulin staining, suggesting that this was a subgroup of the chromogranin A–positive cells.

FOLLOW-UP

Over a median period of 20 months after partial pancreatectomy (range, 5 to 51), three patients were entirely free of any postprandial symptoms, and two patients had occasional mild, nondescript symptoms but no hypoglycemia. After almost a year of being symptom-free, Patient 3 had a recurrence of symptoms of postprandial hypoglycemia, which was confirmed by findings of low glucose levels on a reflectance meter, although the symptoms were less severe and frequent than they had been preoperatively. The recurrence of symptoms is probably due to the more conservative distal pancreatectomy performed in Patient 3, which probably did not remove all affected pancreatic tissue.

DISCUSSION

There is emerging recognition that postprandial hypoglycemia may be due to endogenous hyperinsulinemia from abnormal islets, as a result of either nesidioblastosis or insulinoma. Our finding extends that observation to patients with postprandial neuroglycopenia who have undergone gastric bypass as a treatment for severe obesity. The need for caution in ascribing postprandial symptoms to the dumping syndrome in patients who have undergone gastric bypass without considering the possibility of organic hyperinsulinism was emphasized by the authors of a case report of insulinoma after gastric bypass.[15]

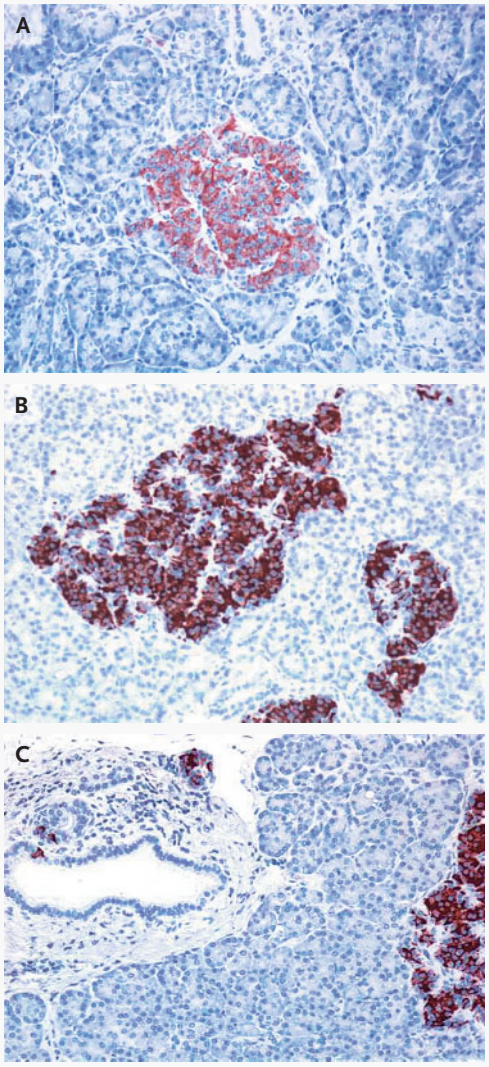


Figure 1. Islets from an Obese Control Subject (Panel A) and Patient 2 (Panel B) and Pancreatic Ducts from Patient 2 (Panel C).

Panel A shows a normal islet from an obese control subject with immunohistochemical staining for insulin. Panel B shows an enlarged islet from Patient 2 with immunohistochemical staining for insulin. The islet is about three times as large as the normal-sized islet in the lower right side of the figure. Panel C shows pancreatic ducts with insulin-positive cells (nesidioblastosis) from Patient 2. A portion of a hypertrophic islet with immunohistochemical staining for insulin is on the right. All three specimens are shown at the same magnification.

We initially considered our index case, in a patient who had functioning insulinomas and islet hypertrophy after gastric bypass, to be an unusual coincidence (postprandial symptoms and two types of pathological islet lesions). However, subsequent experience with patients who had postprandial hyperinsulinemic hypoglycemia as a result of nesidioblastosis after gastric bypass led us to raise the possibility of a link between the islet hyperfunction and the bypass surgery. The frequency of nesidioblastosis after gastric bypass exceeds that in the general population, since only nine adult patients without a history of gastric bypass had surgically confirmed nesidioblastosis at our institution during the same period in which the six patients in the present report were evaluated and treated. Thus, the latter group of patients accounted for 40 percent of our patients with confirmed cases of nesidioblastosis during that time, but only about 0.1 percent of the U.S. population has undergone gastric bypass procedures.[2]

It is possible that hyperinsulinemia from islet-cell hyperfunction led to the development of severe obesity or, alternatively, that the hyperinsulinemia was a consequence of the gastric bypass. Patients with insulinomas may gain weight, but rarely to a degree that would warrant gastric bypass. The median body-mass index of 58 patients with surgically confirmed insulinoma consecutively treated at our institution during the same five-year period in which the patients in this report were seen was 29.1 (range, 18.4 to 53.8). Six patients (10 percent) had a body-mass index of at least 40, and 12 patients (21 percent) had a body-mass index of at least 35. Among 17 patients at our institution who had the noninsulinoma pancreatogenous hypoglycemia syndrome and who did not undergo gastric bypass (some of whom have been described previously[5,6]), the median body-mass index was 27.2 (range, 19.5 to 35.6). These observations argue against islet hyperfunction, especially nesidioblastosis, as the trigger for severe obesity. Conversely, obesity does not appear to cause islet hypertrophy, as attested to by the normal size of islets in obese patients without hypoglycemia.

Persons with the dumping syndrome as a result of previous gastric surgery have been reported to have increased levels of glucagon-like peptide 1,[16-18] possibly owing to the rapid presentation of nutrients (a stimulus for the secretion of this peptide) to the distal ileum, the site of L cells, which are the source of glucagon-like peptide 1. Glucagon-like peptide 1 increases beta-cell mass in rodents through neogenesis and proliferation[19-22] and decreases apoptosis of islets in humans.[23] These findings provide support for the possibility

that beta-cell trophic factors may be brought into play after bypass surgery, leading to the growth of pancreatic beta cells and consequent hyperfunction of islets, ultimately culminating in postprandial hypoglycemia. Such a phenomenon, if ultimately confirmed, would affect not only the continued use of gastric bypass, but also the proposed use of glucagon-like peptide 1 or its analogues for the treatment of diabetes.[24] Furthermore, a potential disturbance in the negative association between insulin and grehlin[25] as a result of the expected reduction in grehlin levels after gastric bypass surgery may aggravate the propensity to hypoglycemia.

Postprandial hyperinsulinemic hypoglycemia and nesidioblastosis may occur in patients who have undergone Roux-en-Y gastric bypass for extreme obesity. Increased levels of a beta-cell–trophic polypeptide, such as glucagon-like peptide 1, may contribute to the hypertrophy of pancreatic beta cells in these patients.

Dr. F. John Service reports having received grant support from the National Institutes of Health; Dr. Lloyd, grant support from the National Institutes of Health and the Verto Institute; and Dr. Thompson, grant support from the National Institutes of Health, the National Cancer Institute (32385), the Food and Drug Administration, and the Mayo Clinic.

We are indebted to Mrs. Susan Hovey for assistance in the preparation of the manuscript.

REFERENCES

1. Manson JE, Skerret PJ, Greenland P, VanItallie TB. The escalating pandemics of obesity and sedentary lifestyle: a call to action for clinicians. Arch Intern Med 2004;164:249-58.
2. Steinbrook R. Surgery for severe obesity. N Engl J Med 2004;350:1075-9.
3. Monteforte MJ, Turkelson CM. Bariatric surgery for morbid obesity. Obes Surg 2000;10:391-401.
4. Sigstad H. A clinical diagnostic index in the diagnosis of the dumping syndrome: changes in plasma volume and blood sugar after a test meal. Acta Med Scand 1970;188:479-86.
5. Service FJ, Natt N, Thompson GB, et al. Noninsulinoma pancreatogenous hypoglycemia: a novel syndrome of hyperinsulinemic hypoglycemia in adults independent of mutations in Kir6.2 and SUR1 genes. J Clin Endocrinol Metab 1999;84:1582-9.
6. Thompson GB, Service FJ, Andrews JC, et al. Noninsulinoma pancreatogenous hypoglycemia syndrome (NIPHS): an update in 10 surgically treated patients. Surgery 2000;128:937-45.
7. Service FJ. Classification of hypoglycemic disorders. Endocrinol Metabol Clin North Am 1999;28:501-17.
8. Threatte GA, Henry JB. Carbohydrates. In: Henry JB, ed. Clinical diagnosis and management by laboratory methods. 19th ed. Philadelphia: W.B. Saunders, 1996:194-207.
9. Kao PC, Taylor RJ, Heser DW. C-peptide immunochemiluminetric assay developed from two seemingly identical polyclonal antisera. Ann Clin Lab Sci 1992;22:307-16.
10. Moura MRL, de Nucci G, Rath S, Reyes FGR. LC-APCI-MS-MS methodology for determination of glybenclamide in human plasma. Anal Bioanal Chem 2004;378:499-503.
11. Magni F, Marazzini L, Pereira S, Monti L, Galli KM. Identification of sulfonylureas in serum by electrospray mass spectrometry. Anal Biochem 2000;282:136-41.
12. Sacks DB. Carbohydrates. In: Burtis CA, Ashwood ER, eds. Tietz textbook of clinical chemistry. Philadelphia: W.B. Saunders, 1999:750-808.
13. Solcia E, Capella C, Kloppel G. Tumors of the pancreas. In: Atlas of tumor pathology. 3rd series, fascicle 20. Washington, D.C.: Armed Forces Institute of Pathology, 1997:120-44.
14. Lloyd RV, Scheithauer BW, Kovacs K, Roche PC. The immunophenotype of pituitary adenomas. Endocr Pathol 1996;7:145-50.
15. Zagury L, Moreira RO, Guedes EP, Coutinho WF, Appolonario JC. Insulinoma misdiagnosed as dumping syndrome after bariatric surgery. Obes Surg 2004;14:120-3.
16. Gebhard B, Holst JJ, Biegelmayer B, Miholic J. Postprandial GLP-1, norepinephrine, and reactive hypoglycemia in dumping syndrome. Dig Dis Sci 2001;46:1915-23.
17. Miholic JC, Orskov JJ, Holst JJ, Kotzerke J, Meyer HJ. Emptying of the gastric substitute, glucagon-like peptide-1 (GLP-1), and reactive hypoglycemia after total gastrectomy. Dig Dis Sci 1991;36:1361-70.
18. Andreasen JJ, Orskov C, Holst JJ. Secretion of glucagon-like peptide-1 and reactive hypoglycemia after partial gastrectomy. Digestion 1994;55:221-8.
19. Zhou J, Wang X, Pineyro MA, Egan JM. Glucagon-like peptide 1 and exendin-4 convert pancreatic AR42J cells into glucagon- and insulin-producing cells. Diabetes 1999;48:2358-66.
20. De Leon DD, Deng S, Madani R, Ahima RS, Drucker DJ, Stoffers DA. Role of endogenous glucagon-like peptide-1 in islet regeneration after partial pancreatectomy. Diabetes 2003;52:365-71.
21. List JF, Habener JF. Glucagon-like peptide 1 agonists and the development and growth of pancreatic β-cells. Am J Physiol Endocrinol Metab 2004;286:E875-E881.
22. Brubaker PL, Drucker DJ. Minireview: glucagon-like polypeptides regulate cell proliferation and apoptosis in the pancreas, gut and central nervous system. Endocrinology 2004;145:2653-9.
23. Farilla L, Bulotta A, Hirshberg B, et al. Glucagon-like peptide 1 inhibits cell apoptosis and improves glucose responsiveness of freshly isolated human islets. Endocrinology 2003;144:5149-58.
24. Zander M, Madsbad S, Madsen JL, Holst JJ. Effect of 6-week course of glucagon-like peptide 1 on glycaemic control, insulin sensitivity, and β-cell function in type 2 diabetes: a parallel-group study. Lancet 2002;359:824-30.
25. Fabio B, Gottero C, Benso A, et al. Ghrelin and the endocrine pancreas. Endocrine 2003;22:19-24.

SPECIAL ARTICLE

Quality of Care in U.S. Hospitals as Reflected by Standardized Measures, 2002–2004

Scott C. Williams, Psy.D., Stephen P. Schmaltz, Ph.D., David J. Morton, M.S., Richard G. Koss, M.A., and Jerod M. Loeb, Ph.D.

ABSTRACT

BACKGROUND

In July 2002, the Joint Commission on Accreditation of Healthcare Organizations implemented standardized performance measures that were designed to track the performance of accredited hospitals and encourage improvement in the quality of health care.

METHODS

We examined hospitals' performance on 18 standardized indicators of the quality of care for acute myocardial infarction, heart failure, and pneumonia. One measure assessed a clinical outcome (death in the hospital after acute myocardial infarction), and the other 17 measures assessed processes of care. Data were collected over a two-year period in more than 3000 accredited hospitals. All participating hospitals received quarterly feedback in the form of comparative reports throughout the study.

RESULTS

Descriptive analysis revealed a significant improvement ($P<0.01$) in the performance of U.S. hospitals on 15 of 18 measures, and no measure showed a significant deterioration. The magnitude of improvement ranged from 3 percent to 33 percent during the eight quarters studied. For 16 of the 17 process-of-care measures, hospitals with a low level of performance at baseline had greater improvements over the subsequent two years than hospitals with a high level of performance at baseline.

CONCLUSIONS

Over a two-year period, we observed consistent improvement in measures reflecting the process of care for acute myocardial infarction, heart failure, and pneumonia. Both quantitative and qualitative research are needed to explore the reasons for these improvements.

From the Joint Commission on Accreditation of Healthcare Organizations, Division of Research, Oakbrook Terrace, Ill. Address reprint requests to Dr. Williams at the Joint Commission on Accreditation of Healthcare Organizations, 1 Renaissance Blvd., Oakbrook Terrace, IL 60181, or at swilliams@jcaho.org.

N Engl J Med 2005;353:255-64.

THE 2001 REPORT OF THE INSTITUTE OF Medicine, *Crossing the Quality Chasm*, indicated that the American health care system "is currently functioning at far lower levels than it could and should."[1] To address shortcomings in the health care system, the report called for the "establishment of monitoring and tracking processes for use in evaluating the progress of the health system." The report emphasized the need to improve the effectiveness of health care through the consistent provision of services that are based on current scientific knowledge.

In 2003, the Agency for Healthcare Research and Quality (AHRQ) released the National Healthcare Quality Report (NHQR), which included results on a broad set of 57 performance measures that provided data on the trend in the quality of services for several clinical conditions.[2] Although improvement was reported in 20 of the 57 measures for which trend data were available, the use of disparate, preexisting data sources limited the analysis and skewed the representativeness of some samples. In 2003, Jencks et al. reported on the positive changes in care delivered to Medicare beneficiaries, on the basis of data collected during two periods.[3] These results, however, could not be confidently generalized beyond the Medicare population. In addition, neither the AHRQ nor the Centers for Medicare and Medicaid Services (CMS) provided feedback to contributing hospitals on their performance as a tool for continual quality improvement.

In 2002, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) implemented evidence-based standardized measures of performance in over 3000 accredited hospitals. The measures were designed to track hospitals' performance over time and encourage improvement through quarterly feedback in the form of comparative reports to all participating hospitals. Both qualitative and quantitative studies have demonstrated benefits associated with providing hospitals regular feedback on their performance on quality measures.[4-7] Comparative feedback has been particularly useful at an organizational level as a guide for improvement-oriented activities.[4,8,9]

While focusing on accredited hospitals, this report expands on the earlier work of the NHQR and CMS in three important ways. First, we did not limit patient populations to Medicare beneficiaries; rather, we included all patients. Second, data collected during the study were made available to hospitals through formal quarterly feedback reports, allowing hospitals to monitor their performance over time in comparison to national rates. Third, our use of hospital-level longitudinal analysis allowed us to compare the rates of change between hospitals that began the study with a low level of performance and those that began with a high level of performance.

METHODS

PARTICIPANTS

JCAHO accreditation accounts for more than 90 percent of the acute care medical–surgical hospital beds in the United States.[10] In July 2002, the JCAHO required 3377 of its 4644 accredited hospitals to submit data on standardized performance measures on their choice of at least two of the four initially available sets of measures: acute myocardial infarction, heart failure, pneumonia, and pregnancy and related conditions. (The pregnancy measures were not included in this analysis because two of the three measures address rare events and the third [vaginal birth after cesarean] is a subject of controversy.) The 27 percent of accredited hospitals exempted from this requirement either did not provide services addressed by any of the measure sets (e.g., psychiatric hospitals or specialty hospitals) or had an average daily census of fewer than 10 patients.

MEASURES AND DATA COLLECTION

The standardized performance measures for the quality of care for acute myocardial infarction, heart failure, and pneumonia (Table 1) were composed of precisely defined specifications and standardized data-collection protocols based on uniform medical language. The measures were designed to permit valid comparisons of health care organizations through the establishment of a national comparative database. Measures were identified with respect to published scientific evidence and consistency with established clinical-practice guidelines, since considerable gaps still exist between the practices described in clinical guidelines and the degree to which these practices are implemented during actual treatment.[11-14]

The JCAHO performed a pilot test of the measures from January through December 2001 through a collaborative effort among five state hospital associations and 83 hospitals in nine states.[15] After the national implementation of these measures, the reliability of the approach was reassessed

with the use of on-site reabstraction of medical records to evaluate the accuracy of the individual data elements collected in the third quarter of 2002. The average rates of agreement exceeded 90 percent.[16]

The JCAHO required accredited hospitals to collect data on performance measures for all eligible patients through the abstraction of medical records and, where applicable, the use of administrative or billing data. In a small number of hospitals (approximately 2 percent each of hospitals submitting data on acute myocardial infarction, heart failure, and pneumonia), patients' records were randomly sampled from all eligible patients. Only hospitals with at least 75 eligible patients per month were allowed to use sampling. Once these data were collected, hospitals submitted patient-level data to a third-party vendor, which compiled and transmitted hospital-level data to the JCAHO on a quarterly basis.[17,18] All participating hospitals received comparative feedback reports meeting standardized specifications on a quarterly basis. These reports were supplied by the third-party vendors and included, at a minimum, control charts to track variations in a hospital's performance over time and comparison charts to compare a hospital's rates for each measure against the national averages. Vendors often provided additional feedback for their hospital clients.

STATISTICAL ANALYSIS

For each measure, quarterly rates or means were calculated. Rates were based on a quarterly aggregation of data from all eligible patients. Rate-based measures are presented as a proportion in which the number of patients meeting the criteria for a specific measure is divided by the total number of patients. For continuous variables, national means were based on an aggregation of hospital means, weighted according to the number of patients, rather than a simple grand mean. Continuous variables are presented as a mean value (i.e., the number of minutes) for all patients who qualified for a given measure. For these measures, 2 percent of the aggregated data were identified as extreme outliers (i.e., monthly data points exceeding a threshold in which the mean time to thrombolysis was greater than 6 hours, the mean time to percutaneous coronary intervention was greater than 24 hours, or the mean time to the administration of antibiotics for pneumonia was greater than 36 hours) and were removed from the analysis. The national trends were analyzed with the use of ordinary least-squares regression analysis to quantify the linear change over time, expressed as change per quarter on the percent scale. For rate-based measures that have a high overall rate of performance, there was evidence that rates approached an upper asymptote over time. For these measures, a nonlinear three-parameter logistic curve was fitted to the data to quantify this upper asymptote. To compare the time trend on the measure of smoking-cessation counseling or advice, an analysis of covariance was used.[19]

For the analysis of hospital temporal trends, a mixed random-coefficients model analysis was used to assess the time trend, with the specific form of the analysis depending on the type of measure being analyzed: generalized linear mixed models were used to analyze binomial counts for the rate-based measures,[20] normal mixed random-coefficient models were weighted with the use of the variance of each data point for the continuous variables,[21] and a Poisson general linear mixed model was used for the risk-adjusted measures.[22] The influence of baseline values of a measure on the change in hospitals' performance over time was assessed by adding an independent baseline variable to the model, as well as an interaction between this baseline variable and the linear effect of time. The interpretation of this interaction is the change in the linear effect of time per unit change in the baseline value of the measure. All reported P values are two-sided and are not adjusted for multiple testing.

RESULTS

To ensure longitudinal comparability, we limited the analysis to hospitals that submitted data from the third quarter of 2002 (the first quarter of the study) through the second quarter of 2004 (the eighth and final quarter of the study). As a result, of the 3377 hospitals initially identified as participants, 3087 were included in the analysis; 1473 of the 3087 hospitals submitted data on acute myocardial infarction (only 1258 submitted data for the mean time to thrombolysis, and only 688 submitted data for the mean time to percutaneous coronary intervention, since not all hospitals submitting data on acute myocardial infarction provided these services), 1946 hospitals submitted data on heart failure, and 1797 submitted data on pneumonia. The decision to limit our analysis to hospitals with complete data sets led to the removal of 69 hospitals from the analysis of data on acute myocardial infarction, 82 hospitals from the analysis of data on heart failure, and 95 hospitals from the analysis of data on pneumonia.

Table 1. Core Measures of the JCAHO for the Quality of Care for Acute Myocardial Infarction, Heart Failure, and Pneumonia.*

Measure	Patients Included	Patients Excluded
Acute myocardial infarction	Patients 18 years of age or older with a principal ICD-9-CM discharge diagnosis of acute myocardial infarction	
Aspirin within 24 hours after admission†‡		Patients who were transferred to another acute care hospital on day of arrival or transferred from another hospital, including another emergency department; patients who were discharged, died, or left against medical advice on day of arrival; patients with one or more contraindications to aspirin
Aspirin prescribed at discharge†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice; patients with one or more contraindications to aspirin
ACE inhibitor prescribed at discharge for patients with left ventricular systolic dysfunction†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice; patients participating in a clinical trial of ACE inhibitors; patients with one or more contraindications to ACE inhibitors
Smoking-cessation counseling or advice†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice
Beta-blocker within 24 hours after admission†‡		Patients who were transferred to another acute care hospital on day of arrival or transferred from another hospital, including another emergency department; patients who were discharged, died, or left against medical advice on day of arrival; patients with one or more contraindications to beta-blockers
Beta-blocker prescribed at discharge†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice; patients with one or more contraindications to beta-blockers
Mean time from arrival to thrombolysis‡§	Patients with ST-segment elevation or left bundle-branch block on the electrocardiogram obtained closest to hospital arrival who received thrombolytic therapy within the first 6 hours after arrival	Patients who were transferred to another acute care hospital on day of arrival or transferred from another hospital, including another emergency department; patients who received thrombolytic therapy more than 6 hours after arrival
Mean time from arrival to PCI‡§	Patients with ST-segment elevation or left bundle-branch block on the electrocardiogram obtained closest to hospital arrival; patients with a valid ICD-9-CM procedure code for PCI; patients who underwent PCI within the first 24 hours after arrival	Patients who were transferred to another acute care hospital on day of arrival or transferred from another hospital, including another emergency department; patients who underwent PCI more than 24 hours after hospital arrival; patients given thrombolytic agents
Inpatient death†¶		Patients who were transferred to another acute care hospital on day of arrival or transferred from another hospital; patients who died in the emergency department
Heart failure	Patients 18 years of age or older with a principal ICD-9-CM discharge diagnosis of heart failure	
Discharge instructions regarding medications, diet, weight, worsening of symptoms, follow-up, and activity†‡		Patients discharged or transferred anywhere except home, home care, or home intravenous therapy
Assessment of left ventricular function†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice; patients with documented reasons for the absence of an assessment of left ventricular function

ACE inhibitor prescribed at discharge for patients with left ventricular systolic dysfunction†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice; patients participating in a clinical trial of ACE inhibitors; patients with one or more contraindications to ACE inhibitors
Smoking-cessation counseling or advice†‡		Patients who were transferred to another acute care hospital, died, left against medical advice, or were discharged to a hospice
Pneumonia	Patients with a principal ICD-9-CM discharge diagnosis of pneumonia (or a principal discharge diagnosis of septicemia or respiratory failure with a secondary diagnosis code of pneumonia)	
Oxygenation assessment within 24 hours after admission†‡		Patients less than 18 years of age; patients who were transferred from another acute care or critical care access hospital; patients who had no working diagnosis of pneumonia at admission; patients who were receiving comfort measures only
Pneumococcal screening, vaccination, or both by discharge†‡		Patients less than 65 years of age; patients who were transferred from another acute care or critical care access hospital; patients who were transferred to another acute care hospital or federal hospital, died, left against medical advice, or were discharged to a hospice; patients who had a documented allergy or sensitivity to pneumococcal vaccine; patients who had no working diagnosis of pneumonia at admission; patients who were receiving comfort measures only
Blood cultures collected before initiation of antibiotic therapy†‡		Patients less than 18 years of age; patients who were transferred from another acute care or critical care access hospital; patients who had no working diagnosis of pneumonia at admission; patients who were receiving comfort measures only; patients for whom no blood cultures were obtained
Smoking-cessation counseling or advice†‡		Patients less than 18 years of age; patients who were transferred to another acute care hospital or federal hospital, died, left against medical advice, or were discharged to a hospice; patients who had no working diagnosis of pneumonia at admission; patients who were receiving comfort measures only
Mean time from arrival to initial antibiotic administration‡§		Patients less than 29 days of age; patients who were transferred from another acute care or critical care access hospital; patients who had no working diagnosis of pneumonia at admission; patients who were receiving comfort measures only; patients who did not receive antibiotics during hospitalization or whose initial antibiotic was administered more than 36 hours after arrival

* Detailed specifications for the core measures, including descriptions of each measure, definitions of data elements, programming algorithms, code tables, and risk-adjustment methods, are available at www.jcaho.org/pms/core+measures/information+on+final+specifications.htm. Two of the original core measures were not included in this study. The pneumonia measure "pediatric smoking-cessation advice or counseling" was omitted owing to the small and unstable population size. The pregnancy measure "vaginal birth after cesarean section" was omitted owing to the changes in the evidence base for the measure and the resulting difficulties in interpretation. JCAHO denotes Joint Commission on Accreditation of Healthcare Organizations; ICD-9-CM *International Classification of Diseases, 9th Revision, Clinical Modification*; ACE angiotensin-converting enzyme; and PCI percutaneous coronary intervention.

† This rate-based measure is presented as a proportion, in which the number of patients who met the criterion is divided by the total number of patients.

‡ This measure reflects a clinical care process (as opposed to an outcome of care).

§ This continuous variable is presented as a mean value for all patients who met the criterion.

¶ This measure targets a specific outcome of clinical care.

These hospitals had very small samples, leading to particular quarters in which no cases of these conditions were reported.

NATIONAL ANALYSIS

National rates for each measure are shown according to quarter in Figure 1 and Table 2. Rate-based measures are shown for each condition. The three measures expressed as continuous variables are displayed together in Figure 1D. The national rates and means include data from all participating hospitals. On a national scale, performance for 15 of the 18 standardized measures demonstrated a significant trend of improvement (change per quarter) over

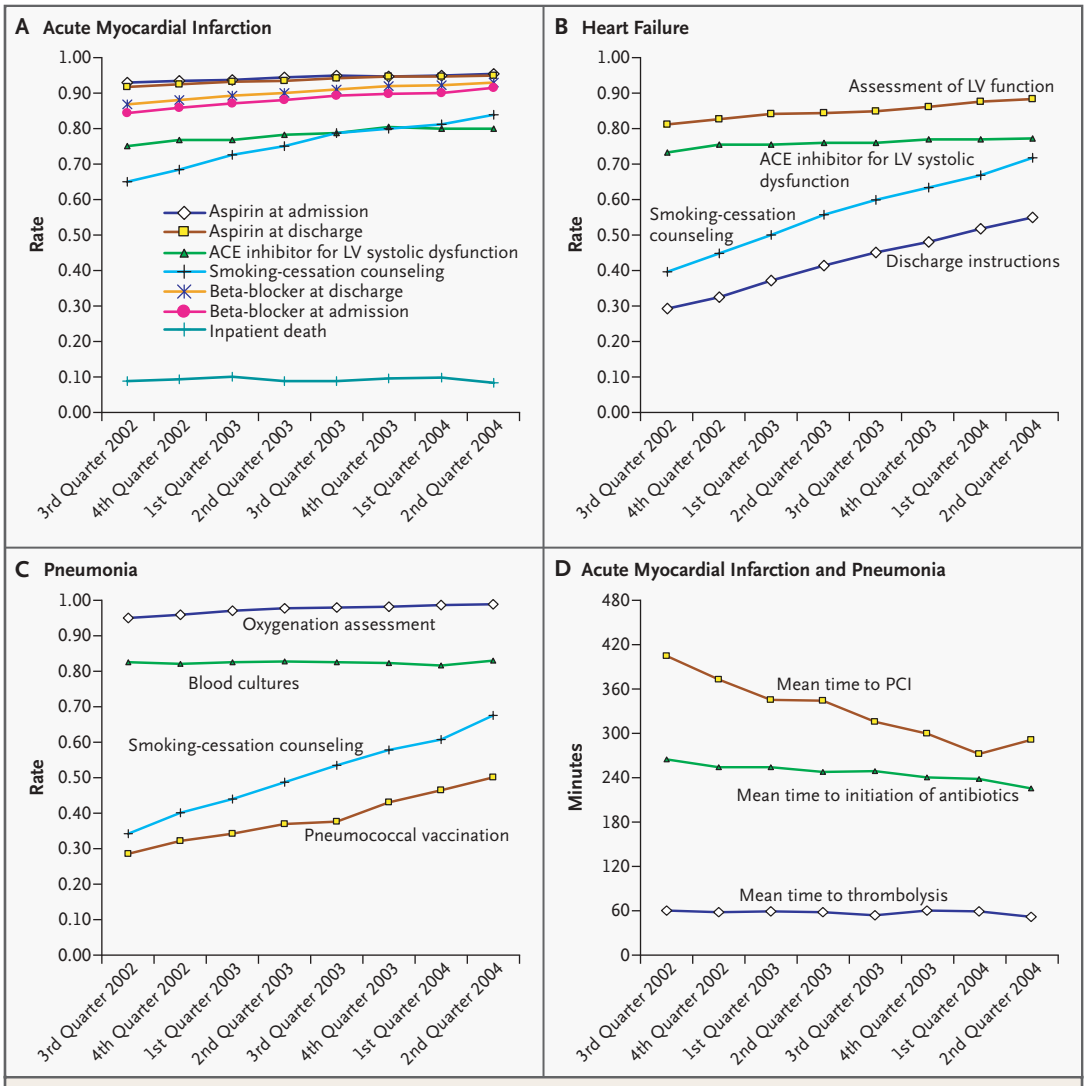


Figure 1. Trends in the Measures of the Quality of Care for Acute Myocardial Infarction (Panels A and D), Heart Failure (Panel B), and Pneumonia (Panels C and D) at U.S. Hospitals from July 2002 to June 2004.

Data in Panels A, B, and C are based on aggregate calculations for rate-based measures (dividing the number of patients who met the criterion by the total number of patients) for all participating hospitals. An improvement is reflected by a positive slope for all rate-based measures except inpatient death after acute myocardial infarction, for which an improvement is reflected by a negative slope. Panel D shows weighted mean values (in minutes) for the continuous variables. Mean values for each hospital were weighted by the total number of patients included by each hospital. For this panel, an improvement is reflected by a negative slope. ACE denotes angiotensin-converting enzyme, LV left ventricular, and PCI percutaneous coronary intervention.

the eight-quarter period (P<0.01), and no measure showed significant deterioration. The overall rates for four of the measures for acute myocardial infarction and one of the pneumonia measures approached an upper limit: aspirin at admission (96 percent), aspirin at discharge (96 percent), beta-blocker at admission (95 percent), beta-blocker at discharge (96 percent), and oxygenation assessment (99 percent).

Among the 18 measures studied, the most dramatic improvement occurred in the three measures of counseling for smoking cessation. A 19 percent, 32 percent, and 33 percent absolute difference from the first to the last quarter studied was observed for acute myocardial infarction, heart failure, and pneumonia, respectively. Interestingly, comparisons of the measures of counseling for smoking cessation, which were identically defined in the three categories, revealed significant differences in the national performance rates in each. Performance on the measure of smoking-cessation counseling for acute myocardial infarction was superior to that for heart failure, which was, in turn, better than that for pneumonia (P<0.001). The rates of change (slope) over the eight quarters were also significantly different (P<0.001) between the measures of counseling for smoking cessation for acute myocardial infarction (3 percent per quarter) and the measures of coun-

Table 2. Mean Values and Overall Changes in Measures of the Quality of Care during the Eight Quarters.*

Measure	Average No. of Patients per Hospital per Quarter	1st Quarter	8th Quarter	Absolute Difference	P Value
Acute myocardial infarction					
Aspirin at admission (%)	40	93	95	3	0.002
Aspirin at discharge (%)	48	92	95	3	<0.001
ACE inhibitor for LV systolic dysfunction (%)	13	75	80	5	<0.001
Smoking cessation counseling (%)	18	65	84	19	<0.001
Beta-blocker at admission (%)	35	84	91	7	<0.001
Beta-blocker at discharge (%)	47	87	93	6	<0.001
Mean time to thrombolysis (min)	4	62	54	−8	0.53
Mean time to PCI (min)	10	406	293	−113	<0.001
Inpatient death (%)	43	9	8	−1	0.58
Heart failure					
Discharge instructions (%)	57	29	55	26	<0.001
Assessment of LV function (%)	69	81	88	7	<0.001
ACE inhibitor for LV systolic dysfunction (%)	24	73	77	4	0.005
Smoking-cessation counseling (%)	11	39	72	32	<0.001
Pneumonia					
Oxygenation assessment (%)	68	95	99	4	<0.001
Pneumococcal vaccination (%)	37	28	50	22	<0.001
Blood cultures (%)	49	82	83	<1	0.31
Smoking-cessation counseling (%)	13	34	67	33	<0.001
Mean time to initiation of antibiotics (min)	65	266	227	−39	<0.001

* The first quarter of the study was the third quarter of 2002, and the eighth and final quarter of the study was the second quarter of 2004. Mean rates for rate-based measures are based on aggregate calculations (dividing all patients who met the criterion by the total number of patients) for all participating hospitals. An improvement is reflected by an increase in the rate for all rate-based measures except inpatient death after acute myocardial infarction, for which an improvement is reflected by a decrease in the rate. Mean values for continuous variables are based on aggregated mean rates for each participating hospital weighted by the total number of patients included by each hospital. An improvement in these measures is reflected by a decrease in the value. All numbers (except P values) were rounded to the nearest integer. ACE denotes angiotensin-converting enzyme, LV left ventricular, and PCI percutaneous coronary intervention.

seling for smoking cessation for heart failure and pneumonia (4 percent and 5 percent per quarter, respectively).

HOSPITAL-LEVEL ANALYSIS

The analyses involving a mixed random-coefficients model demonstrated that the degree of hospitals' improvement was significantly positively associated with baseline performance and linear time for all rate-based process measures and two of the three measures expressed as continuous variables ($P<0.05$). Only the mean time to thrombolysis did not reveal a significant relationship between baseline performance and linear time. More simply stated, the performance of hospitals generally tended to improve over time, and hospitals that began the study with lower baseline rates tended to improve at faster rates than hospitals with higher baseline rates. Table 3 illustrates this trend by stratifying hospitals into three groups according to their baseline percentile ranks for each measure.

DISCUSSION

Our data demonstrate a steady improvement in the performance of U.S. hospitals over a period of eight quarters in measures reflecting the quality of care for acute myocardial infarction, heart failure, and pneumonia. Improvement was observed in 15 of 18 measures, including 3 measures that already had mean performance rates of over 90 percent in the first quarter (e.g., the rate for aspirin at admission was 93 percent in the third quarter of 2002). Moreover, hospital-level analysis revealed that, for 16 of the 17 process measures, hospitals that began the study as low-level performers tended to improve at faster rates than those that started the study with higher levels of performance. With each passing quarter, low-level performers improved more quickly. In contrast, high-level performers generally maintained their high level of performance or improved at slower rates.

The faster rate of improvement among low-level performers represents an important finding. Whereas low-level performers have the most room for improvement, one might have expected different results, since such hospitals may be less likely to focus on quality or make an effort to improve performance than their counterparts with a higher level of performance. Our results support the results of previous work and lend support to the use of these measures as a means for encouraging improvement in hospitals and as tools for monitoring hospitals' performance, as called for by the Institute of Medicine.[1,3,7-9,23] Receiving quarterly national comparative data may have been an added stimulus for poor-performing hospitals to improve.

The improvement observed in some measures may have resulted in part from increased attention to documentation, rather than better patient care. Although this possibility cannot be definitively discounted, it could not explain the reductions noted in the time to thrombolysis and time to percutaneous coronary intervention.

A review of the measures of the quality of care for acute myocardial infarction reveals an apparent contradiction between improvement in the process measures and the lack of improvement in the inpatient-death measure. This discrepancy is misleading for two important reasons. First, it is highly unlikely that improvements in these process measures would have an important effect on inpatient death. In fact, four of the eight process measures for acute myocardial infarction address discharge activities, and the patient populations targeted by the measures for the timing of thrombolytic therapy and percutaneous coronary intervention, which one might reasonably expect to be correlated with inpatient death, represent only a small fraction of the patients included in the inpatient-death measure. Second, the process measures were selected because they had a scientific evidence base, established through randomized clinical trials, that demonstrated their relationship to multiple measures of outcome. In such instances, process measures can be more sensitive to differences in quality than comparisons of outcomes.[24] The inpatient-death measure would therefore be expected to provide important information about an individual hospital's performance but would not be expected to mirror trends observed for the process measures.

This study has several limitations. First, it is dependent on self-reported data from hospitals. Although the reliability of the data was evaluated twice, once in the pilot test of the measures and a second time after national implementation of the approach,[15,16] the nature of self-reported data provides an opportunity to introduce bias into the results. However, the results are consistent with findings reported in studies that used data collected by independent sources.[3,6,7,23] The Veterans Affairs health care system, for example, used similar measures, collected by independent abstractors, to track changes in performance after the implementation of a systemwide reengineering program. Over a four-year period, dramatic improvements

in the quality of care were observed.[25] Second, although improvements must have been due to hospital-based efforts to upgrade the quality of care, there is no way to determine the degree to which feedback on performance measures stimulated these improvement initiatives. Certainly, the national attention directed at these high-risk, problem-prone patient populations by the CMS, the National Quality Forum, the JCAHO, and others contributed to the observed improvement, independently of the availability of feedback on performance measures.

In this issue of the *Journal*, Jha et al., who analyzed 10 measures from the CMS Hospital Quality Alliance, also reported modest differences in the performance of hospitals on the basis of specific characteristics, such as geographic location, teaching status, and for-profit or not-for-profit status.[26] These easily observable demographic characteristics, however, did not account for the majority of the variation in quality observed on the measures. Both quantitative and qualitative research are needed to evaluate the reasons for differences in

Table 3. Average Change in Measures of the Quality of Care over Time, According to Baseline Performance.*

Measure	Low Level of Performance at Baseline (0–25th percentile)		Average Level of Performance at Baseline (26th–75th percentile)		High Level of Performance at Baseline (76th–100th percentile)	
	1st Quarter	8th Quarter	1st Quarter	8th Quarter	1st Quarter	8th Quarter
Rate-based (%)						
Acute myocardial infarction						
Aspirin at admission	78	93	95	96	100	96
Aspirin at discharge	66	86	93	94	100	92
ACE inhibitor for LV systolic dysfunction	36	74	79	81	100	83
Smoking-cessation counseling	7	68	61	80	98	85
Beta-blocker at admission	61	86	88	92	100	93
Beta-blocker at discharge	59	85	89	93	100	93
Heart failure						
Discharge instructions	1	42	24	53	73	73
Assessment of LV function	50	72	81	87	97	93
ACE inhibitor for LV systolic dysfunction	42	70	76	77	98	84
Smoking-cessation counseling	2	63	37	69	87	81
Pneumonia						
Oxygenation assessment	82	97	96	99	100	99
Pneumococcal vaccination	0	35	25	50	67	66
Blood cultures	63	77	85	84	98	84
Smoking-cessation counseling	1	57	31	65	80	74
Continuous (min)						
Time to thrombolysis	138	63	52	56	24	49
Time to PCI	881	340	302	262	92	196
Time to initiation of antibiotics	380	254	247	218	159	190

* Hospitals were grouped into three strata on the basis of their percentile rank for each measure at baseline. The first quarter of the study was the third quarter of 2002, and the eighth and final quarter was the second quarter of 2004. An improvement in rate-based measures is reflected by an increase in the rate. An improvement in a continuous variable and the outcome measure is reflected by a decrease in the value. For 16 of the 17 measures, a low level of performance at baseline was significantly associated with greater improvement in performance over the subsequent seven quarters ($P<0.001$ for 15 measures and $P=0.03$ for the time to percutaneous coronary intervention [PCI]). Baseline performance was not associated with an improvement in the time to thrombolysis ($P=0.33$). These P values were based on a random-effects model: $y_t=\alpha+B_1\text{baseline}+B_2\text{time}+B_3\text{baseline}\cdot\text{time}$, where y_t represents a hospital's predicted value at a given time (the log base e odds ratio for rate-based measures and the value for continuous variables). ACE denotes angiotensin-converting enzyme, and LV left ventricular.

hospital performance and improvement. Our study included only hospitals that were accredited by the JCAHO and excluded hospitals with an average daily census of fewer than 10 patients. As a result, our findings may underrepresent the performance of very small hospitals.

Although the impetus for the improvement cannot be pinpointed, the improvement in measures reflecting the quality of care for acute myocardial infarction, heart failure, and pneumonia remains very encouraging. If the current rates of change were to be maintained (which is by no means a certainty), the mean performance for hospitals that began the study in the lowest performance quartile would be expected to reach rates of 90 percent for 11 of the 14 rate-based process measures by the first quarter of 2006. Given the very low starting point or slow rate of improvement observed for the remaining three rate-based process measures (discharge instructions, pneumococcal screening or vaccination, and blood cultures), the 90 percent mark would probably not be reached until 2007 or 2008. It is also important to note that the data were largely collected before the widespread public reporting of hospital data or the implementation of pay-for-performance initiatives. The role and influence of public reporting is a widely debated subject that varies depending on the intended audience and the purpose of the reporting.[9,23,27,28] As the JCAHO's database of performance measures expands with each passing quarter, it will offer an opportunity to track the effect of national public reporting and pay-for-performance initiatives to a degree that once was not possible.

REFERENCES

1. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C.: National Academies Press, 2001.
2. Agency for Healthcare Research and Quality. 2004 National Healthcare Quality Report. (Accessed June 27, 2005, at http://www.qualitytools.ahrq.gov/qualityreport/download_report.aspx.)
3. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. JAMA 2003;289:305-12. [Erratum, JAMA 2002;289:2649.]
4. Bradley EH, Holmboe ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. Data feedback efforts in quality improvement: lessons learned from US hospitals. Qual Saf Health Care 2004;13:26-31.
5. *Idem.* A qualitative study of increasing beta-blocker use after myocardial infarction: why do some hospitals succeed? JAMA 2001;285:2604-11.
6. Schade CP, Cochran BF, Stephens MK. Using statewide audit and feedback to improve hospital care in West Virginia. Jt Comm J Qual Saf 2004;30:143-51.
7. Marciniak TA, Ellerbeck EF, Radford MJ, et al. Improving the quality of care for Medicare patients with acute myocardial infarction: results from the Cooperative Cardiovascular Project. JAMA 1998;279:1351-7.
8. Gibberd R, Hancock S, Howley P, Richards K. Using indicators to quantify the potential to improve the quality of health care. Int J Qual Health Care 2004;16:Suppl 1:i37-i43.
9. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA 2000;283:1866-74.
10. Healthcare Quick Disk 2005: AHA annual survey database for fiscal year 2003. Chicago: American Hospital Association, 2005.
11. Antman EM, Anbe DT, Armstrong PW, et al. ACC/AHA guidelines for the management of patients with ST elevation myocardial infarction — executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to revise the 1999 guidelines for the management of patients with acute myocardial infarction). J Am Coll Cardiol 2004;44:671-719. [Erratum, J Am Coll Cardiol 2005;45:1376.]
12. Hunt SA, Baker DW, Chin MH, et al. ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure). J Am Coll Cardiol 2001;38:2101-13.
13. Bartlett JG, Breiman RF, Mandell LA, File TM Jr. Community-acquired pneumonia in adults: guidelines for management. Clin Infect Dis 1998;26:811-38.
14. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.
15. Joint Commission on Accreditation of Healthcare Organizations. A comprehensive review of development and testing for national implementation of hospital core measures. (Accessed June 27, 2005, at http://www.jcaho.org/pms/core+measures/cr_hos_cm.htm.)
16. Watt A, Williams S, Lee K, Robertson J, Koss RG, Loeb JM. Keen eye on core measures: Joint Commission data quality study offers insights into data collection, abstracting processes. J AHIMA 2003;74:20-5.
17. Braun BI, Koss RG, Loeb JM. Integrating performance measure data into the Joint Commission accreditation process. Eval Health Prof 1999;22:283-97.
18. Joint Commission on Accreditation of Healthcare Organizations. Performance measurement system requirements for ORYX listing. (Accessed June 27, 2005, at http://www.jcaho.org/pms/reference+materials/pms+requirements.htm.)
19. Winer BJ. Statistical principles in experimental design. 2nd ed. New York: McGraw-Hill, 1971.
20. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. J Am Stat Assoc 1993;88:9-25.
21. SAS/STAT user's guide, version 8. Cary, N.C.: SAS Institute, 2002.
22. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 1987;43:671-81.
23. Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? Health Aff (Millwood) 2003;22(2):84-94.
24. Palmer RH. Using health outcomes data to compare plans, networks and providers. Int J Qual Health Care 1998;10:477-83.
25. Jha AK, Perlin JB, Kizer KW, Dudley RA. Effect of the transformation of the Veterans Affairs health care system on the quality of care. N Engl J Med 2003;348:2218-27.
26. Jha AK, Zhonghe L, Orav EJ, et al. Care in U.S. hospitals — the Hospital Quality Alliance program. N Engl J Med 2005;353:265-74.
27. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care: a survey of cardiovascular specialists. N Engl J Med 1996;335:251-6.
28. *Idem.* Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998;279:1638-42.

SPECIAL ARTICLE

Care in U.S. Hospitals — The Hospital Quality Alliance Program

Ashish K. Jha, M.D., M.P.H., Zhonghe Li, M.A., E. John Orav, Ph.D., and Arnold M. Epstein, M.D.

ABSTRACT

BACKGROUND

The Hospital Quality Alliance (HQA) is the first initiative that routinely reports data on hospitals' performance nationally. Heretofore, such data have been unavailable.

METHODS

We used data collected by the Centers for Medicare and Medicaid Services on 10 indicators of the quality of care for acute myocardial infarction, congestive heart failure, and pneumonia. The main outcome measures were hospitals' performance with respect to each indicator and summary scores for each clinical condition. Predictors of a high level of performance were determined with the use of multivariable linear regression.

RESULTS

A total of 3558 hospitals reported data on at least one stable measure (defined as information obtained from discharge data from at least 25 patients) during the first half of 2004. Median performance scores (expressed as the percentage of patients who satisfied the criterion) were at least 90 percent for 5 of the 10 measures but lower for the other 5. Performance varied moderately among large hospital-referral regions, with the top-ranked regions scoring 12 percentage points (for acute myocardial infarction) to 23 percentage points (for pneumonia) higher than the bottom-ranked regions. A high quality of care for acute myocardial infarction predicted a high quality of care for congestive heart failure but was only marginally better than chance at predicting a high quality of care for pneumonia. Characteristics associated with small but significant increases in performance included being an academic hospital, being in the Northeast or Midwest, and being a not-for-profit hospital.

CONCLUSIONS

Analysis of data from the new HQA national reporting system shows that performance varies among hospitals and across indicators. Given this variation and small differences based on hospitals' characteristics, performance reporting will probably need to include numerous clinical conditions from a broad range of hospitals.

From the Department of Health Policy and Management, Harvard School of Public Health (A.K.J., Z.L., A.M.E.); the Division of General Medicine, Brigham and Women's Hospital (A.K.J., E.J.O., A.M.E.); and the Boston Veterans Affairs Healthcare System (A.K.J.) — all in Boston. Address reprint requests to Dr. Jha at the Department of Health Policy and Management, Harvard School of Public Health, 677 Huntington Ave., Boston, MA 02115, or at ajha@hsph.harvard.edu.

N Engl J Med 2005;353:265-74.

Numerous studies have now shown that the quality of health care is variable and often inadequate.[1-3] Initiatives to measure quality are an important focus for policymakers who believe that measurement can drive quality-improvement programs and guide the choice of provider by consumers and payers.[4,5]

For more than a decade, the National Committee for Quality Assurance has published annual data on the quality of care provided by health plans as measured by quality indicators in the Health Plan Employer Data and Information Set.[6] Until recently, however, we have lacked any national database that could provide analogous data on the quality of care provided by hospitals. Recently, a consortium of organizations, including the Centers for Medicare and Medicaid Services (CMS), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the American Hospital Association, and consumer groups such as the American Association of Retired Persons, initiated an effort now called the Hospital Quality Alliance (HQA) to fill this gap. Under the HQA, hospitals nationwide report data to the CMS on indicators of the quality of care for acute myocardial infarction, congestive heart failure, and pneumonia. Both the CMS and the JCAHO have collected data based on these indicators, albeit in some instances with slightly disparate specifications. Different, limited versions of these data became available on the Internet in late 2003 in the case of the CMS (at www.cms.hhs.gov) and in July 2004 in the case of the JCAHO (at www.qualitycheck.org). The national data from the CMS first became publicly available for research in November 2004.

Despite the intense effort that has gone into defining and collecting these measures of quality, little is known about how hospitals measure up. We used the HQA data to answer four important questions: How well do hospitals perform on the basis of these quality measures? How variable is performance across regions, and more specifically, are there certain local regions in which the level of performance is consistently high or low? What is the likelihood that a high level of performance in one condition (e.g., acute myocardial infarction) predicts a high level of performance in other conditions (e.g., congestive heart failure)? Finally, do certain characteristics of hospitals, including profit status, number of beds, presence or absence of academic involvement, and geographic region, predict a high level of performance?

METHODS

CONDITIONS AND MEASURES OF QUALITY

To initiate the reporting effort, the CMS selected 10 measures of the quality of care that have been widely endorsed[7-10] and that are considered valid and feasible for immediate public reporting. These 10 measures reflect the quality of care for three major clinical conditions: acute myocardial infarction, congestive heart failure, and pneumonia. There were five measures of the quality of care for acute myocardial infarction: the use or nonuse of aspirin within 24 hours before or after arrival at the hospital and at discharge, the use or nonuse of a beta-blocker within 24 hours after arrival and at discharge, and the use or nonuse of an angiotensin-converting–enzyme (ACE) inhibitor for left ventricular systolic dysfunction. Two measures were used for congestive heart failure: assessment of left ventricular function and the use or nonuse of an ACE inhibitor for left ventricular dysfunction. Three measures were used for pneumonia: the timing of initial antibiotic therapy, the presence or absence of pneumococcal vaccination, and assessment of oxygenation.

The Medicare Modernization Act, passed in 2003, established financial incentives for hospitals to provide the CMS with data on these 10 indicators of quality. On April 1, 2005, data became available on the performance of hospitals with respect to seven additional measures — three for acute myocardial infarction, two for congestive heart failure, and two for pneumonia. Because these seven additional measures were based on admissions during only one quarter of 2004 and were available for relatively few hospitals (fewer than 10 percent of hospitals reported data on five of the seven measures that were based on an adequate sample size), we describe their performance in the Supplementary Appendix (available with the full text of this article at www.nejm.org) but chose not to include them in our primary analyses.

All data collected by the HQA from every hospital are audited quarterly by the CMS Clinical Data Abstraction Centers, which abstract and reanalyze data from five charts per hospital per quarter. Specifications for the measures are provided in the Supplementary Appendix.

PERFORMANCE DATA

HQA data on 10 quality indicators first became publicly available on November 30, 2004, and were up-

dated on April 1, 2005, to reflect hospital admissions during the first half of 2004. For each of the 10 measures, a hospital's score reflects the proportion of patients who satisfied the criterion. We defined any hospital performance measure that was based on discharge data from at least 25 patients as a stable measure, to be consistent with the convention of the CMS to refer to such a measure as reliable.

CHARACTERIZATION OF HOSPITALS AND PERFORMANCE ON INDIVIDUAL INDICATORS

We linked the HQA data set to the database of the American Hospital Association, which has information on hospitals' characteristics with respect to profit status, number of beds, region, type of setting (urban vs. rural), status of membership in the Council of Teaching Hospitals, percentage of patients covered by Medicare and Medicaid, ratio of nurses to patient-days (calculated by dividing the number of nurses on staff by 1000 patient-days), and presence or absence of an intensive care unit.

STATISTICAL ANALYSIS

We used chi-square tests and analysis of variance to compare the characteristics of hospitals that reported no data to the HQA with those that reported some data for every measure but that were based on discharge data for fewer than 25 patients and those that reported adequate data for at least one measure. In addition, t-tests with unequal variance were used to compare performance measures between hospitals with adequate sample sizes and hospitals with inadequate sample sizes. For each hospital, we used both performance scores that were weighted (according to the number of patients) and performance scores that were unweighted and found that the differences were similar in magnitude and direction. Therefore, we chose to report unweighted mean performance scores.

Factor Analysis and Creation of Summary Scores

To reduce the 10 performance measures to manageable summary scores, we performed factor analysis. The factor analysis combined the five measures of the quality of care for acute myocardial infarction into a weighted average with almost identical weightings. Therefore, for simplicity, we used an equally weighted average of the five items as our summary score for acute myocardial infarction. Similar results for the congestive heart failure and pneumonia measures led us to use equally weighted summary scores for these two conditions as well. Because Cronbach's alpha (the degree of association among the measures) showed a very strong correlation (0.82) among the five measures for acute myocardial infarction and because of the need to retain as representative a sample of hospitals as possible, we also calculated a summary score for acute myocardial infarction for hospitals that had stable measures for four of the five indicators related to acute myocardial infarction. This gave us a total of 1537 hospitals with summary scores for acute myocardial infarction. Given lower Cronbach's alpha values for the two measures related to congestive heart failure (0.60) and for the three pneumonia measures (0.43), we included only the 1915 hospitals that had stable measures for both indicators related to congestive heart failure and the 3076 hospitals that had stable measures for all three pneumonia items in our calculation of summary scores for these two conditions.

Performance of Hospital-Referral Regions

We examined summary scores according to hospital-referral regions, which are based on regional markets for tertiary care and were previously described in the *Dartmouth Atlas of Health Care*.[11] In this calculation, we combined all patients with any of the three conditions who were treated in hospitals for which we had hospital summary scores and chose the 40 hospital-referral regions with the largest total numbers of patients. We then calculated an average summary score for each of the three clinical conditions in each hospital-referral region by averaging the summary scores of individual hospitals within each region. We subsequently ranked all regions according to their performance on quality measures for each condition and calculated the difference (with 95 percent confidence intervals) in performance between the top-ranked and bottom-ranked regions. Finally, we calculated Spearman correlation coefficients to determine how performance in one condition was correlated with performance in another condition across referral regions.

Predicting Quality across Conditions

We determined, on a hospital-by-hospital basis, how performance in one condition related to performance in other conditions. We used information on the hospitals that had summary scores for both acute myocardial infarction and congestive heart failure and the hospitals that had summary scores for both acute myocardial infarction and pneumonia. For each comparison, we first catego-

rized each hospital's performance according to the summary score for acute myocardial infarction. We then calculated the proportion of hospitals in the top decile, top quartile, bottom quartile, and bottom decile of performance measures for acute myocardial infarction that scored in the top quartile, top half, or bottom quartile of performance measures for each of the other two conditions (congestive heart failure and pneumonia).

Hospital Characteristics and Performance

We examined whether four characteristics of hospitals — profit status (for profit vs. not for profit), academic status (member of the Council of Teaching Hospitals vs. nonmember), number of beds, and region of the country — that have previously been found to be associated with the quality of hospital care[12-14] were associated with performance with respect to each of the three conditions. We built separate multivariable linear regression models with the summary scores for acute myocardial infarction, congestive heart failure, and pneumonia as outcomes. The models were simultaneously adjusted for each of the four primary predictors as well as other available characteristics that might be associated with performance: the proportion of patients with Medicare insurance, the proportion of patients with Medicaid insurance, the ratio of nurses to patient-days, the presence or absence of an intensive care unit, and setting (urban vs. rural).

RESULTS

Of the 4203 hospitals in the HQA database, 4002 hospitals reported on at least one measure to the CMS. The 201 hospitals that reported no data were mostly specialty surgical centers and orthopedic hospitals. A total of 444 hospitals reported only information that was based on discharge data from fewer than 25 patients, and 3558 hospitals reported information on one or more measures that was based on discharge data from at least 25 patients and thus considered stable. The three categories of hospitals differed significantly in terms of size, geo-

Table 1. Characteristics of the Hospitals.*

Characteristic	Hospitals That Reported ≥1 Stable Measure (N=3558)	Hospitals That Reported Only Nonstable Measures (N=444)	Hospitals That Reported No Data (N=201)	P Value†
No. of beds — no. (% of hospitals)				<0.001
<100	1093 (31)	339 (76)	100 (50)	
100–400	1908 (54)	51 (11)	15 (7)	
>400	423 (12)	4 (1)	0	
Region — no. (% of hospitals)				<0.001
West	1038 (29)	186 (42)	80 (40)	
Midwest	828 (23)	142 (32)	59 (29)	
South	1027 (29)	81 (18)	44 (22)	
Northeast	627 (18)	21 (5)	16 (8)	
For profit — no. (% of hospitals)	567 (16)	34 (8)	38 (19)	<0.001
Member of COTH — no. (% of hospitals)	277 (8)	7 (2)	0	<0.001
Urban setting — no. (% of hospitals)	2209 (62)	92 (21)	63 (31)	<0.001
Cardiac ICU — no. (% of hospitals)	1400 (39)	39 (9)	3 (1)	<0.001
Medical ICU — no. (% of hospitals)	2667 (75)	107 (24)	20 (10)	<0.001
Medicare insurance (% of patients)	0.45±0.13	0.50±0.22	0.39±0.23	<0.001
Medicaid insurance (% of patients)	0.17±0.10	0.15±0.13	0.13±0.14	<0.001
No. of nurses/1000 patient-days	5.80±2.81	7.75±18.29	31.44±209.44	<0.001

* Plus–minus values are means ±SD. Measures based on discharge data from at least 25 patients were defined as stable. Matching AHA data were not available for all hospitals; therefore, percentages do not necessarily sum to 100. COTH denotes the Council of Teaching Hospitals, and ICU intensive care unit.

† P values are for the difference among the three categories of hospital.

graphic region, status of membership in the Council of Teaching Hospitals, presence of an intensive care unit, and setting (Table 1).

The quality of care in the hospitals that reported at least 1 stable measure was higher on 9 of the 10 measures than in hospitals that reported no stable measures, although some of the differences were small and not significant (Table 2). The quality of care in the hospitals that reported at least 1 stable measure varied widely across the 10 measures, from a mean (±SD) of 98±5 percent for oxygenation assessment to 43±27 percent for pneumococcal vaccination (Table 2 and Fig. 1). The median score was at least 90 percent on four of the five performance measures for acute myocardial infarction (all except ACE-inhibitor therapy), neither of the two performance measures for congestive heart failure, and one of the three pneumonia measures (oxygenation assessment).

SUMMARY SCORES AND PERFORMANCE ACROSS HOSPITAL-REFERRAL REGIONS

Among the 1537 hospitals for which we could calculate a summary score for acute myocardial infarction, the mean score was 89±6 percent. Similarly, the mean summary score for congestive heart failure among the 1915 hospitals that were included in the analysis was 81±10 percent, and the mean pneumonia score among the 3076 hospitals included in the analysis was 71±11 percent.

Among the 40 largest hospital-referral regions in the database, we found substantial gaps in mean performance among the three conditions. The difference in the pneumonia composite score between the top-ranked region in this respect (Oklahoma City) and the bottom-ranked region (San Bernardino, Calif.) was 23±4 percentage points, the gaps between the top- and bottom-ranked performers for acute myocardial infarction (12±4 percentage points) and the top- and bottom-ranked performers for congestive heart failure (21±5 percentage points) were smaller (Table 3). There was a moderate correlation between the performance of a hospital-referral region with respect to acute myocardial infarction and its performance with respect to congestive heart failure (Spearman correlation coefficient, 0.72; P<0.001) but a lower correlation between the performance of a region with respect to acute myocardial infarction and its performance with respect to pneumonia (Spearman correlation coefficient, 0.45;

Table 2. Mean Performance Scores.*

Measure	Hospitals That Reported ≥1 Stable Measure (N=3558)	Hospitals That Reported Only Nonstable Measures (N=444)	P Value
	percent		
Acute myocardial infarction			
Aspirin at admission	92±12	81±33	<0.001
Aspirin at discharge	87±19	76±36	<0.001
ACE inhibitor for left ventricular dysfunction	75±26	64±45	0.06
Beta-blocker at admission	85±20	75±35	<0.001
Beta-blocker at discharge	84±19	72±36	<0.001
Congestive heart failure			
Assessment of left ventricular function	80±18	58±33	<0.001
ACE inhibitor for left ventricular dysfunction	74±20	71±36	0.18
Pneumonia			
Oxygenation assessment	98±5	95±15	<0.001
Pneumococcal vaccination	43±27	38±35	0.001
Timing of initial antibiotic therapy	72±13	77±23	<0.001

* Plus–minus values are means ±SD. Scores reflect the percentage of patients who met the criterion. ACE denotes angiotensin-converting enzyme. Stable is defined as any performance measure that is based on discharge data from at least 25 patients.

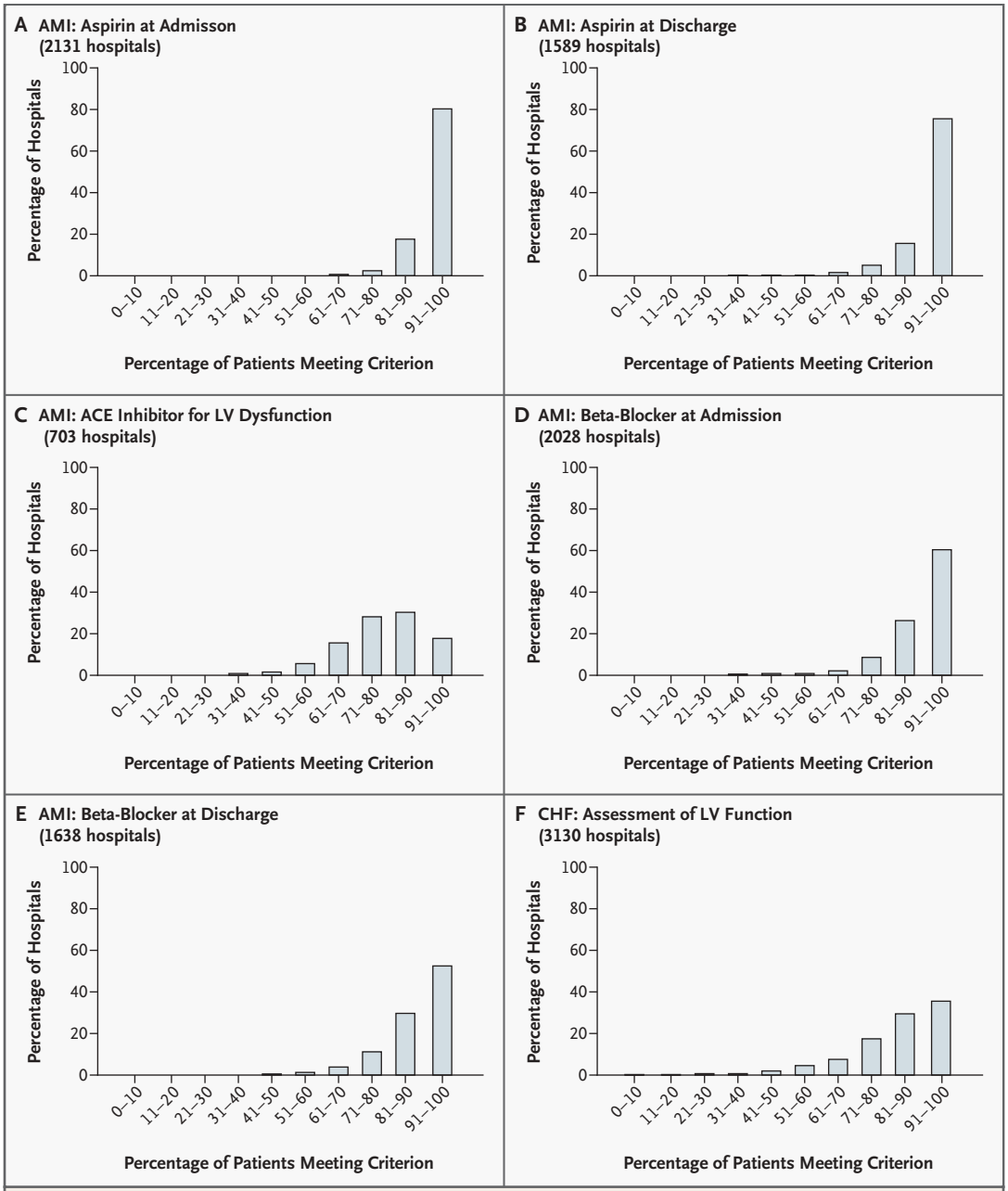


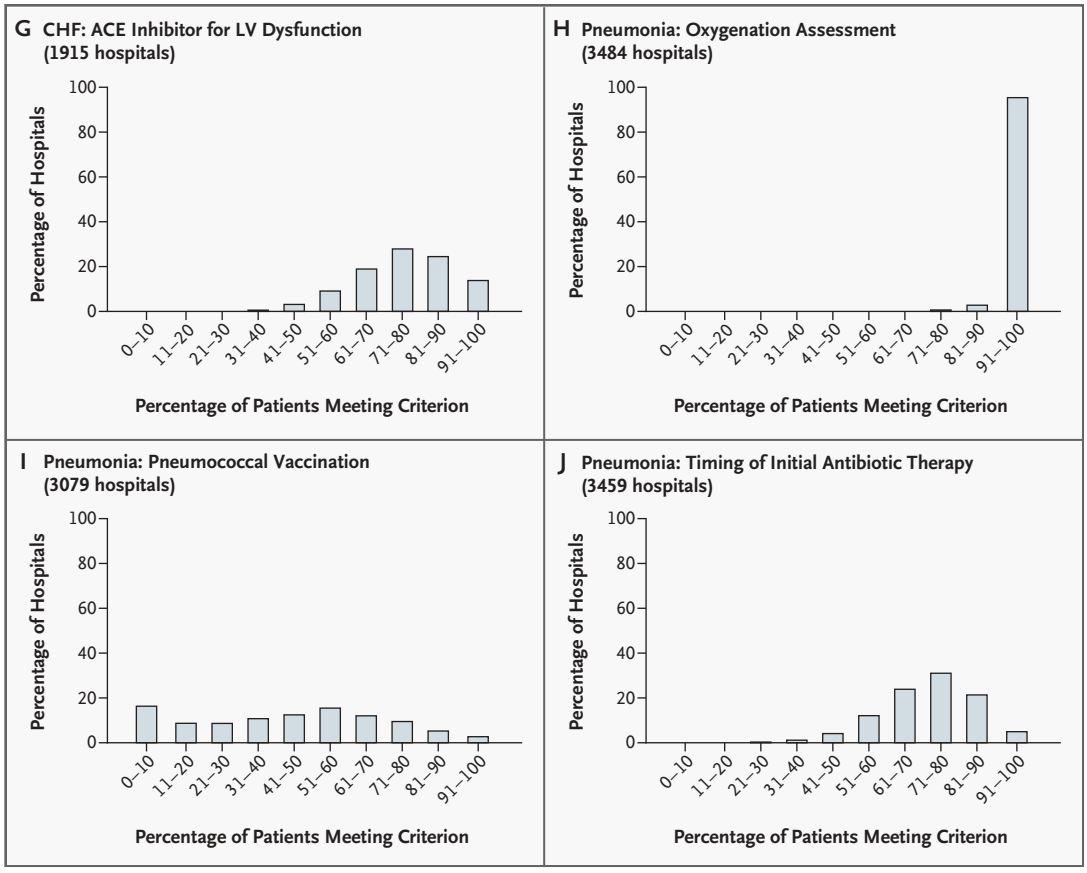
Figure 1. Distribution of Performance Scores for the 10 Core Measures of the Quality of Care for Acute Myocardial Infarction (Panels A, B, C, D, and E), Congestive Heart Failure (Panels F and G), and Pneumonia (Panels H, I, and J).

AMI denotes acute myocardial infarction, ACE angiotensin-converting enzyme, LV left ventricular, and CHF congestive heart failure.

P=0.004) and a lower correlation still between the performance of a region with respect to congestive heart failure and its performance with respect to pneumonia (Spearman correlation coefficient, 0.15; P=0.35).

PREDICTING QUALITY ACROSS CONDITIONS WITHIN HOSPITALS

Performance scores for acute myocardial infarction closely predicted performance scores for congestive heart failure but not for pneumonia. Seventy-three



percent of hospitals that were in the top decile of performance scores for acute myocardial infarction were in the top quartile of performance scores for congestive heart failure, and 91 percent of such hospitals were in the top half of performance scores for congestive heart failure, whereas less than 1 percent were in the bottom quartile (Table 4). However, only 33 percent of hospitals in the top decile of performance scores for acute myocardial infarction were in the top quartile of performance scores for pneumonia, and 41 percent were in the bottom half.

CHARACTERISTICS AND PERFORMANCE OF HOSPITALS

We subsequently examined the relationship between performance scores and hospitals' academic status (as reflected by membership or nonmembership in the Council of Teaching Hospitals), profit status, geographic region, and number of beds (Table 5). We found that after adjustment for potential confounders (as well as the other variables of interest), academic hospitals had higher performance scores for acute myocardial infarction than nonacademic hospitals (91 percent vs. 89 percent, $P<0.001$) and congestive heart failure (85 percent vs. 81 percent, $P<0.001$), but lower scores for pneumonia (69 percent vs. 71 percent, $P=0.02$). Not-for-profit hospitals had significantly higher scores for all three conditions than did for-profit hospitals, and there were significant regional differences in scores for each of the three conditions, with the Midwest and Northeast outperforming the West and South. The number of beds was significantly associated only with the pneumonia scores ($P=0.001$), with the smallest hospitals having the highest scores.

DISCUSSION

We evaluated the national HQA data set launched by the CMS and found that the quality of care in American hospitals varied greatly according to the indicator of quality and the condition. For five of the quality indicators — especially those for acute myocardial infarction — half the hospitals scored above 90 percent. However, the level of performance with respect to other measures of quality was much

Table 3. The Top-Ranked and Bottom-Ranked Performances in Measures of the Quality of Care for AMI, CHF, and Pneumonia among the 40 Largest Hospital-Referral Regions.*

Hospital-Referral Region	AMI Score	No. of Hospitals	Hospital-Referral Region	CHF Score	No. of Hospitals	Hospital-Referral Region	Pneumonia Score	No. of Hospitals
	%			%			%	
Top-ranked			**Top-ranked**			**Top-ranked**		
Boston	95	39	Boston	89	39	Oklahoma City	82	32
Minneapolis	94	37	Detroit	88	14	Indianapolis	79	31
Kansas City, Mo.	94	39	Baltimore	87	21	Kansas City, Mo.	78	35
Albany, N.Y.	93	21	Camden, N.J.	87	21	Camden, N.J.	78	20
Indianapolis	92	33	Cleveland	86	24	Knoxville, Tenn.	77	26
Bottom-ranked			**Bottom-ranked**			**Bottom-ranked**		
Little Rock, Ark.	86	31	San Diego, Calif.	77	22	Miami	63	26
Orlando, Fla.	86	22	Nashville	76	37	Chicago	61	29
Miami	85	26	Orlando, Fla.	74	22	San Diego, Calif.	60	21
Memphis, Tenn.	84	26	Little Rock, Ark.	69	28	Los Angeles	60	66
San Bernardino, Calif.	83	21	Lexington, Ky.	68	30	San Bernardino, Calif.	59	19

* AMI denotes acute myocardial infarction, and CHF congestive heart failure.

lower. There was substantial variability in the quality of care provided by hospitals in different metropolitan areas. A high quality of care for acute myocardial infarction closely predicted a high quality of care for congestive heart failure but not for pneumonia. There were significant but small differences in performance between academic and nonacademic hospitals and for-profit and not-for-profit hospitals, as well as among hospitals in various geographic regions, but there was no consistent association between performance and the size of the hospital.

The HQA is the first national public reporting system that provides detailed performance data for each hospital. All but 1 of the 10 quality indicators we evaluated (oxygenation assessment) were examined previously in a state-level analysis by Jencks and colleagues using data on Medicare beneficiaries from 2000 through 2001.[1] We found that the level of performance was higher than that described by Jencks et al. for all but one measure (the timing of initial antibiotic therapy in patients with pneumonia), which is consistent with the results of Williams et al.,[15] whose findings in this issue of the *Journal* demonstrate temporal improvements in performance using comparable data reported to JCAHO. The variability in performance for different quality indicators may be due to several important factors, including the length of time the process measure has been considered high quality, the importance that clinicians place on the measure, and the difficulty in providing the specific aspect of appropriate care. Further studies of these issues will be critical to future efforts to improve the quality of health care.

Our findings indicate that quality measures had only moderate predictive ability across the three conditions. Although a high quality of care for acute myocardial infarction predicted a high quality of care for congestive heart failure, the former was only marginally better than chance at identifying a high quality of care for pneumonia. These data do not provide support for the notion that "good" hospitals are easy to identify or consistent in their performance across conditions. Our data suggest that evaluations of hospitals' performance will most likely need to be based on a large number of conditions.

On the basis of the literature, one might predict that the quality of care would be higher in large, academically oriented,[16,17] not-for-profit[18,19] hospitals. However, we found moderate associations between these characteristics and hospitals' performance. The quality of care in teaching hospitals has been an especially controversial topic, since care in such hospitals costs much more than care in nonteaching institutions. Our data, based on a limited number of measures but on a much larger sample of hospitals than in most previous studies, suggest

Table 4. Ability of Performance Scores for AMI to Predict High Performance Scores for CHF and Pneumonia.*

Performance Score for AMI	Performance Score for CHF			Performance Score for Pneumonia		
	Top Quartile	Top Half	Bottom Quartile	Top Quartile	Top Half	Bottom Quartile
		percent			*percent*	
Top decile	73	91	<1	33	59	25
Top quartile	60	85	4	26	54	27
Bottom quartile	5	19	51	8	28	39
Bottom decile	2	10	64	4	18	50

* Deciles and quartiles refer to hospital ranking. AMI denotes acute myocardial infarction, and CHF congestive heart failure.

Table 5. Adjusted Performance Scores for AMI, CHF, and Pneumonia, According to Select Characteristics of the Hospitals.*

Characteristic of Hospital	AMI Score	P Value	CHF Score	P Value	Pneumonia Score	P Value
	%		%		%	
Academic status		<0.001		<0.001		0.02
Member of COTH	91		85		69	
Not a member of COTH	89		81		71	
Profit status		0.03		<0.001		0.02
For profit	88		80		70	
Not for profit	90		82		71	
Region		<0.001		<0.001		<0.001
West	89		80		70	
Midwest	91		83		74	
South	88		81		71	
Northeast	90		83		71	
No. of beds		0.74		0.23		0.001
<100 beds	89		81		72	
100–400 beds	89		82		71	
>400 beds	90		81		70	

* Values were adjusted by means of linear regression for each of the other three primary predictor variables as well as for the proportion of patients with Medicare insurance, the proportion of patients with Medicaid insurance, the ratio of nurses to patient-days, the presence or absence of an intensive care unit, and location (urban vs. rural). AMI denotes acute myocardial infarction, CHF congestive heart failure, and COTH the Council of Teaching Hospitals.

that the extra money spent on teaching institutions does not necessarily buy a higher quality of some important components of care. Of course, the training function of teaching hospitals is important in itself. The moderate differences in performance associated with hospitals' characteristics suggest the need to target a large breadth of hospitals for improvements in the quality of care.

Our study has important limitations. First, we could evaluate only 10 measures of the quality of care for three clinical conditions, although these conditions account for 15 percent of Medicare admissions. The CMS plans to expand the HQA database to include additional conditions. Second, our data on hospitals' characteristics do not provide potentially important details, such as a hospital's management structure or quality-management programs that might be associated with a high level of

performance. Finally, our analyses provide results on process measures and not on patient outcomes.

In summary, we found that the quality of hospital care in the United States varies widely across different indicators of quality and that individual hospitals vary in their performance according to indicators and conditions. Although the public reporting of quality measures in the HQA database represents an important start, our results provide a hint of the hard work that lies ahead. The variability of hospitals' performance across conditions and hospitals indicates that we will need to expand our data-collection efforts to include many more conditions and that we will most likely need to focus quality-improvement efforts on a large set of hospitals.

Supported by the Commonwealth Fund, New York.

REFERENCES

1. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. JAMA 2003;289:305-12. [Erratum, JAMA 2002; 289:2649.]
2. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348: 2635-45.
3. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C.: National Academies Press, 2001.
4. Galvin R, Milstein A. Large employers' new strategies in health care. N Engl J Med 2002;347:939-42.
5. Milstein A, Galvin RS, Delbanco SF, Salber P, Buck CR Jr. Improving the safety of health care: the Leapfrog initiative. Eff Clin Pract 2000;3:313-6. [Erratum, Eff Clin Pract 2001;4:94.]
6. HEDIS 2005. Vol. 2. HEDIS technical specifications. No. 10284-100-05. Washington, D.C.: National Committee for Quality Assurance, 2005.
7. British Thoracic Society Standards of Care Committee. BTS guidelines for the management of community acquired pneumonia in adults. Thorax 2001;56:Suppl 4:IV-1–IV-64.
8. Niederman MS, Mandell LA, Anzueto A, et al. Guidelines for the management of adults with community-acquired pneumonia: diagnosis, assessment of severity, antimicrobial therapy, and prevention. Am J Respir Crit Care Med 2001;163:1730-54.
9. Mandell LA, Bartlett JG, Dowell SF, File TM Jr, Musher DM, Whitney C. Update of practice guidelines for the management of community-acquired pneumonia in immunocompetent adults. Clin Infect Dis 2003; 37:1405-33.
10. Antman EM, Anbe DT, Armstrong PW, et al. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of Patients with Acute Myocardial Infarction). Circulation 2004; 110:e82-e292.
11. Wennberg J, Cooper M, eds. The Dartmouth atlas of health care. Chicago: American Hospital Association Press, 1999.
12. Meehan TP, Fine MJ, Krumholz HM, et al. Quality of care, process, and outcomes in elderly patients with pneumonia. JAMA 1997; 278:2080-4.
13. Marciniak TA, Ellerbeck EF, Radford MJ, et al. Improving the quality of care for Medicare patients with acute myocardial infarction: results from the Cooperative Cardiovascular Project. JAMA 1998;279:1351-7.
14. Bradley EH, Holmboe ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. A qualitative study of increasing beta-blocker use after myocardial infarction: why do some hospitals succeed? JAMA 2001;285: 2604-11.
15. Williams SC, Schmaltz SP, Morton DJ, Koss RG, Loeb JM. Quality of care in U.S. hospitals as reflected by standardized measures, 2002–2004. N Engl J Med 2005;353: 255-64.
16. Allison JJ, Kiefe CI, Weissman NW, et al. Relationship of hospital teaching status with quality of care and mortality for Medicare patients with acute MI. JAMA 2000;284: 1256-62.
17. Ayanian JZ, Weissman JS. Teaching hospitals and quality of care: a review of the literature. Milbank Q 2002;80:569-93.
18. Sloan FA, Trogdon JG, Curtis LH, Schulman KA. Does the ownership of the admitting hospital make a difference? Outcomes and process of care of Medicare beneficiaries admitted with acute myocardial infarction. Med Care 2003;41:1193-205.
19. Devereaux PJ, Schunemann HJ, Ravindran N, et al. Comparison of mortality between private for-profit and private not-for-profit hemodialysis centers: a systematic review and meta-analysis. JAMA 2002;288: 2449-57. [Erratum, JAMA 2004;291:186.]

CLINICAL TRIAL REGISTRATION

The *Journal* encourages investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors plan to consider clinical trials for publication only if they have been registered (see N Engl J Med 2004;351:1250-1). The National Library of Medicine's www.clinicaltrials.gov is a free registry, open to all investigators, that meets the committee's requirements.

REVIEW ARTICLE

CURRENT CONCEPTS

Benign Breast Disorders

Richard J. Santen, M.D., and Robert Mansel, M.D., Ph.D.

From the Department of Medicine, University of Virginia Health Sciences Center, Charlottesville (R.J.S.); and the Department of Surgery, University of Wales College of Medicine, Cardiff, Wales, United Kingdom (R.M.). Address reprint requests to Dr. Santen at the University of Virginia, P.O. Box 801416, Charlottesville, VA 22908, or at rjs5y@virginia.edu.

N Engl J Med 2005;353:275-85.

Hormones and growth factors act on stromal and epithelial cells to regulate the development, maturation, and differentiation of mammary-gland cells.[1,2] Broadly summarized, estrogen mediates the development and elongation of ductal tissue, progesterone facilitates ductal branching and lobulo–alveolar development, and prolactin regulates the production of milk protein. At puberty, estradiol and progesterone levels increase to initiate breast development. A complex tree-like structure results, comprising 5 to 10 primary milk ducts that originate at the nipple, 20 to 40 segmental ducts, and 10 to 100 subsegmental ducts that end in glandular units called terminal-duct lobular units.[3] In the adult breast, cyclic changes occur during the menstrual cycle that result in an increased rate of cell proliferation during the luteal phase.[4] An increase in breast size by up to 15 percent may occur during this phase.[4] At menopause, the total number of lobules diminishes.

Changes in the Normal Breast

The morphologic features of the breast undergo substantial change between early adolescence and menopause.[5] The spectrum of normal histologic features ranges from a predominance of ducts, lobules, and intralobular and interlobular stroma to features that exhibit mainly patterns of fibrous change and cyst formation, formerly called fibrocystic disease of the breast. The term "fibrocystic changes" is now preferred, since in up to 50 to 60 percent of women without breast disease, this histologic pattern may be evident.[6] The term "fibrocystic changes" implies that women with lumpy breasts or breasts with nondiscrete nodules do not have breast disease. The fibrocystic changes detected clinically entail no increased risk of breast cancer.

In women between adolescence and the mid-20s, the lobules and stroma in the breast may respond to hormonal stimuli in an exaggerated fashion with the development of single and multiple palpable fibroadenomas.[7] In autopsy series, 15 to 23 percent of women in this age group were found to have fibroadenomas, whereas specialized clinics have found 7 to 13 percent among female patients in this age group and epidemiologic studies have found 2.2 percent.[5,8,9] In the third and fourth decades of life, the degree of diffuse palpable nodularity may increase. In histologic terms, this increase represents adenosis — that is, enhancement of the amount of normal lobular tissue. The stroma may also undergo hypertrophy, resulting in palpable areas of ill-defined fullness, frequently in the axillary tail. In women between the middle of the fourth decade of life and menopause, glandular tissue may undergo further hypertrophy in association with an increase in stromal tissue. A higher prevalence of cyst formation is associated with late menopause, the use of hormone-replacement therapy, and a thin body composition.[8,10]

Classification of Benign Breast Lesions

In a practical classification, breast lesions that do not increase the risk of breast cancer are distinguished from those that confer a small increase in risk (relative risk, 1.5 to 2.0)

or a moderate increase (relative risk, >2.0) (Table 1).[12] The levels of risk have been established by means of long-term follow-up, and the specific lesions were defined by a consensus conference.[12,13] An important basis of the classification is the degree of cellular proliferation.[12-14]

CAUSAL FACTORS IN BENIGN BREAST DISORDERS

Clinical observations of women receiving estrogens and antiestrogen drugs suggest that hormonal events play a causative role in the development of benign lesions. Among postmenopausal women receiving estrogens with or without progestins for more than eight years, the prevalence of benign breast lesions increased by a factor of 1.70 (95 percent confidence interval, 1.06 to 2.72).[15] The antiestrogen drug tamoxifen, when used in the prevention of breast cancer, is associated with a reduction of 28 percent in the prevalence of benign breast lesions (relative risk, 0.72; 95 percent confidence interval, 0.65 to 0.79), including those classified as adenosis, cysts, mammary-duct ectasia, and hyperplasia.[16]

Underlying and acquired genetic changes are also associated with benign breast lesions (Fig. 1). Loss of heterozygosity, as a result of deletions of small segments of DNA,[18-22] is commonly found in benign breast lesions. Lesions are frequently multifocal, with each lesion exhibiting loss of heterozygosity of different regions of DNA. Women with mutations in the *BRCA1* or *BRCA2* genes have a high frequency of multiple benign or malignant breast lesions on meticulous examination of specimens obtained from bilateral mastectomy.[23] These findings support the current theory of an underlying predisposition to mutations in some patients as the cause of multiple breast lesions. This phenomenon (i.e., a predisposition to mutations) has been termed a field effect and, more recently, a mutator phenotype.[24]

PROGRESSION TO MALIGNANT DISEASE

By analogy to the linear model of the development of colon cancer,[25] breast lesions are believed by many investigators to progress in a linear fashion from usual ductal hyperplasia (ductal hyperplasia with-

Table 1. Classification of Benign Breast Lesions on Histologic Examination, According to the Relative Risk of Breast Cancer.

Risk	Proliferation	Histologic Findings
No increase	Minimal	Fibrocystic changes (within the normal range): cysts and ductal ectasia (72%), mild hyperplasia (40%), nonsclerosing adenosis (22%), and periductal fibrosis (16%)*; simple fibroadenoma (15–23%)†; and miscellaneous (lobular hyperplasia, juvenile hypertrophy, and stromal hyperplasia)
		Benign tumors: hamartoma, lipoma, phyllodes tumor,‡ solitary papilloma, neurofibroma, giant adenoma, and adenomyoepithelioma
		Traumatic lesions: hematoma, fat necrosis, and lesions caused by penetration by a foreign body
		Infections: granuloma and mastitis
		Sarcoidosis
		Metaplasia: squamous and apocrine
		Diabetic mastopathy
Small increase (relative risk, 1.5–2.0)	Proliferative without atypia	Usual ductal hyperplasia, complex fibroadenoma (containing cysts >3 mm in diameter, sclerosing adenosis, epithelial calcifications, or papillary apocrine changes), papilloma or papillomatosis, radial scar, and blunt duct adenosis
Moderate increase (relative risk, >2.0)	Proliferative with atypia	Atypical ductal hyperplasia and atypical lobular hyperplasia

* Percentages indicate the percentage of breasts examined at autopsy in which the lesion was found. Data are from Sandison.[11]
† Data are from Goehring and Morabia.[8]
‡ Most phyllodes tumors are considered to be benign fibroepithelial tumors, but some have malignant clinical and histologic features.

out atypia) or from unfolded lobules[20] to atypical ductal hyperplasia and then to ductal carcinoma in situ and invasive cancer. Several biologic and molecular changes (Fig. 1) have been observed in association with this progression, but causal relationships have not been established.

CLINICAL FEATURES OF BENIGN BREAST DISEASE

BREAST PAIN

Cyclic breast pain usually occurs during the late luteal phase of the menstrual cycle, in association with the premenstrual syndrome or independently,[26-31] and resolves at the onset of menses (Table 2).[26,27] In a study of 1171 healthy premenopausal American women, 11 percent had moderate-to-severe cyclic breast pain and 58 percent had mild discomfort.[28,31] Breast pain interfered with usual sexual activity among 48 percent of the patients, and among others it interfered with physical activities (37 percent), social activities (12 percent), and school activities (8 percent).[28,31] Whether caffeine, iodine deficiency, alterations in levels of fatty acid in the breast, fat intake in the diet, or psychological factors have a causative role in cyclic breast pain has not been established.[32,33]

Noncyclic breast pain is unrelated to the menstrual cycle. Detection of focal tenderness is helpful diagnostically and suggests a tender cyst, rupture through the wall of an ectatic duct, or a particularly tender area of breast nodularity. Acute enlargement of cysts and periductal mastitis may cause severe, localized pain with a sudden onset.

NONBREAST PAIN

Pain arising from the chest wall may be mistakenly attributed to the breast. Pain that is limited to a particular area and characterized as burning or knife-like in nature may arise from the chest wall. Several distinct types of pain can be distinguished, including localized or diffuse lateral chest-wall pain, radicular

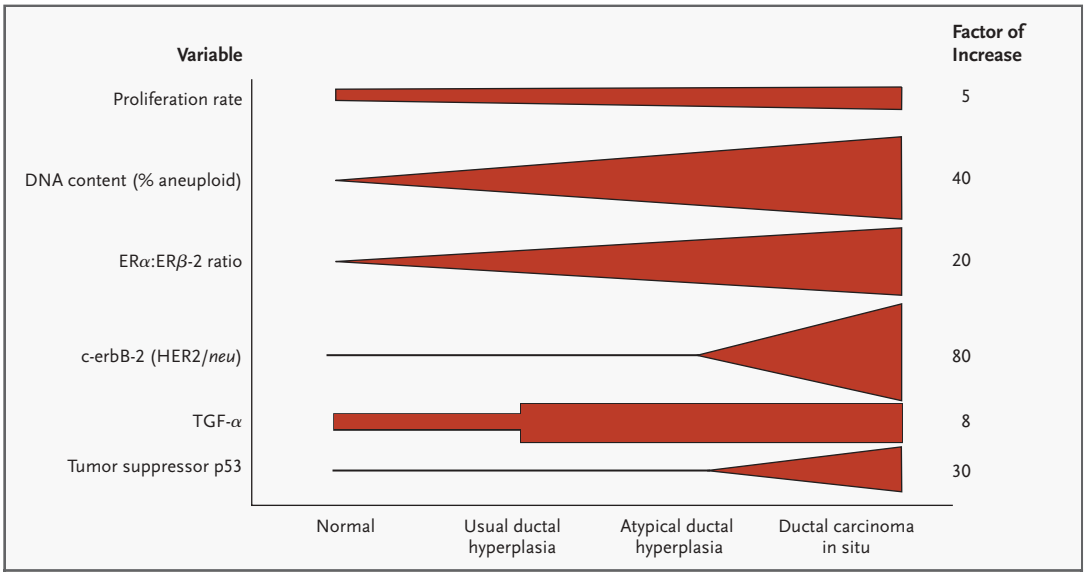


Figure 1. Progression from a Benign to a Malignant Lesion.

Progression to malignant breast disease is associated with the accumulation of an increasing number of genetic mutations. Numbers at the right are factors by which the items listed on the left are increased in women with ductal carcinoma in situ. The rate of proliferation increases by a factor of five when benign breast tissue is compared with ductal carcinoma in situ. The DNA content, as represented by aneuploidy (more or less DNA than would be expected for the presence of 46 chromosomes) gradually increases. Aneuploidy is commonly present in ductal carcinoma in situ. Several studies have shown that levels of estrogen receptor α (ERα) increase, and one large recent study showed a reduction in levels of estrogen receptor β (ERβ).[17] The oncogene for transforming growth factor α (TGF-α) increases concomitantly with the appearance of hyperplasia. There is an increase in c-erbB-2 (HER2/*neu*) and in mutations of the tumor-suppressor gene *p53* only in ductal carcinoma in situ. For references for data in this figure, see the Supplementary Appendix, available with the full text of this article at www.nejm.org.

Table 2. Common Benign Breast Disorders in Women.

Symptom or Finding	Possible Causes or Disorders
Breast pain	
Cyclic pain	Hormonal stimulation of normal breast lobules before menses
Noncyclic pain	Stretching of Cooper's ligaments Pressure from brassiere Fat necrosis from trauma Hidradenitis suppurativa Focal mastitis Periductal mastitis Cyst Mondor's disease (sclerosing periphlebitis of breast veins)
Nonbreast pain	
Chest-wall pain	Tietze's syndrome (costochondritis) Localized lateral chest-wall pain Diffuse lateral chest-wall pain Radicular pain from cervical arthritis
Non–chest-wall pain	Gallbladder disease Ischemic heart disease
Nipple discharge	
Presence of galactorrhea	
From multiple ducts bilaterally	Hyperprolactinemia from pituitary tumor, hypothyroidism, drugs*
Absence of galactorrhea	
From one duct — elicited or spontaneous and bloody, with occult blood, or serosanguineous	Intraductal papilloma Ductal carcinoma in situ Paget's disease of breast
From multiple ducts — elicited and bloody or nonbloody, bilateral, black or clear	Fibrocystic changes Ductal ectasia
Discrete solitary lump	
Age <30 yr	
Firm, rubbery lump	Most common lesion: fibroadenoma
Age 30–50 yr	
Firm, discrete lump	Most common lesions: fibroadenoma, cyst, fibrocystic changes, usual ductal hyperplasia, atypical ductal hyperplasia, atypical lobular hyperplasia†
Age >50 yr	
Firm, discrete lump	Most common lesions: cyst, ductal carcinoma in situ, invasive cancer
Diffuse lumpiness ("lumpy-bumpy")	
Absence of discrete lump	Fibrocystic changes

* Data on drugs with galactorrhea as an adverse effect are listed in standard textbooks.

† Usual ductal hyperplasia, atypical ductal hyperplasia, and atypical lobular hyperplasia may be detected incidentally in female patients undergoing biopsy of masses with other causes, such as fibroadenomas and fibrocystic changes.

pain from cervical arthritis, and pain from Tietze's syndrome (costochondritis).

NIPPLE DISCHARGE

Among female patients referred to physicians because of symptoms of breast disorder, 6.8 percent have nipple discharge. Although this symptom is particularly distressing to the patient, only 5 percent of these patients are found to have serious underlying disease.[34] Nipple discharge is considered to be pathologic if it is spontaneous, arises from a single duct, is persistent, and contains gross or occult blood. Age is an important factor with respect to the risk of malignant disease. In one series, among women with nipple discharge as their only symptom who were found to have cancer, 3 percent were younger than 40 years of age, 10 percent were between 40 and 60 years, and 32 percent were older than 60 years.[35]

FOCAL AND DIFFUSE BREAST LUMPS

Discrete lesions detected by palpation or on routine mammography are different entities in women who are less than 30 years of age, 31 to 50 years, or older than 50 years. On a statistical basis, 9 of 10 new nodules in premenopausal women are benign (Table 1). Diffuse symmetrical lumpiness is commonly found on physical examination and is associated with fibrocystic changes on histologic examination.

ABNORMALITIES ASSOCIATED WITH AN INCREASED RISK OF BREAST CANCER

Several pathologic entities are associated with an enhanced risk of breast cancer[36-38] (Table 1). A recent report suggests that there is a slight increase in the risk of breast cancer among women more than 50 years of age with benign lesions that are in the lower category of risk: cyst, adenosis, mammary-duct ectasia, fibrosis, metaplasia, fibroadenoma, mild-to-moderate or florid hyperplasia without atypia, and papilloma.[38] This study combined two groups of women in the lower category of risk: those with proliferative disease without atypia and those with nonproliferative disease (Table 1). A large study reported elsewhere in this issue of the *Journal*[13] found an increased risk of breast cancer only among women with proliferative disease. Among women with nonproliferative lesions, only those with a strong family history of breast cancer were at increased risk.

Breast density on mammographic screening is also a risk factor, with an increase in relative risk by a factor of five for the highest density[39-42] (Fig. 2). Dense breasts contain a higher proportion of stromal and glandular tissue as well as an increased number of lesions classified as usual ductal hyperplasia and atypical ductal hyperplasia.[40] According to classic studies in twins, heritability accounts for approximately 60 percent of the variation in breast density.[42]

The risk of breast cancer is also increased in association with high plasma levels of free estradiol and testosterone in postmenopausal women,[43,44] weight gain of 20 kg (44 lb) or more after menopause, early menarche, late menopause, late childbearing, and a family history of breast cancer.[45] However, it is not known whether the increased risk associated with breast density or specific histologic lesions adds to these other known risks or whether these factors are merely dependent variables.[44,45] *BRCA1* carriers have a 65 percent probability (95 percent confidence interval, 44 to 78 percent) of developing breast cancer by 70 years of age, and *BRCA2* carriers have a 45 percent probability (95 percent confidence interval, 31 to 56 percent).[46] Studies are under way to evaluate the use of nipple aspiration and ductal lavage as additional means of risk stratification.[47] Currently, the Gail and Claus models are commonly used as practical means of estimating the risk of breast cancer, but these models use only a limited number of variables and are not powerful in predicting disease.[48,49]

PRACTICAL MANAGEMENT

A detailed history and physical examination are used to evaluate systematically the entire breast and the chest wall and should focus on areas related to the patient's symptoms (Table 3). The sensitivity and specificity of the clinical breast examination are estimated to be 54 percent and 94 percent, respectively, and depend on the thoroughness and technique of the examiner.[50] Diagnostic studies may then be ordered. The "triple test" of lumps includes palpation, imaging, and percutaneous biopsy (i.e., core needle biopsy or fine-needle aspiration). Mammography, often in conjunction with ultrasonographic examination,[51-54] is required for evaluation of dis-

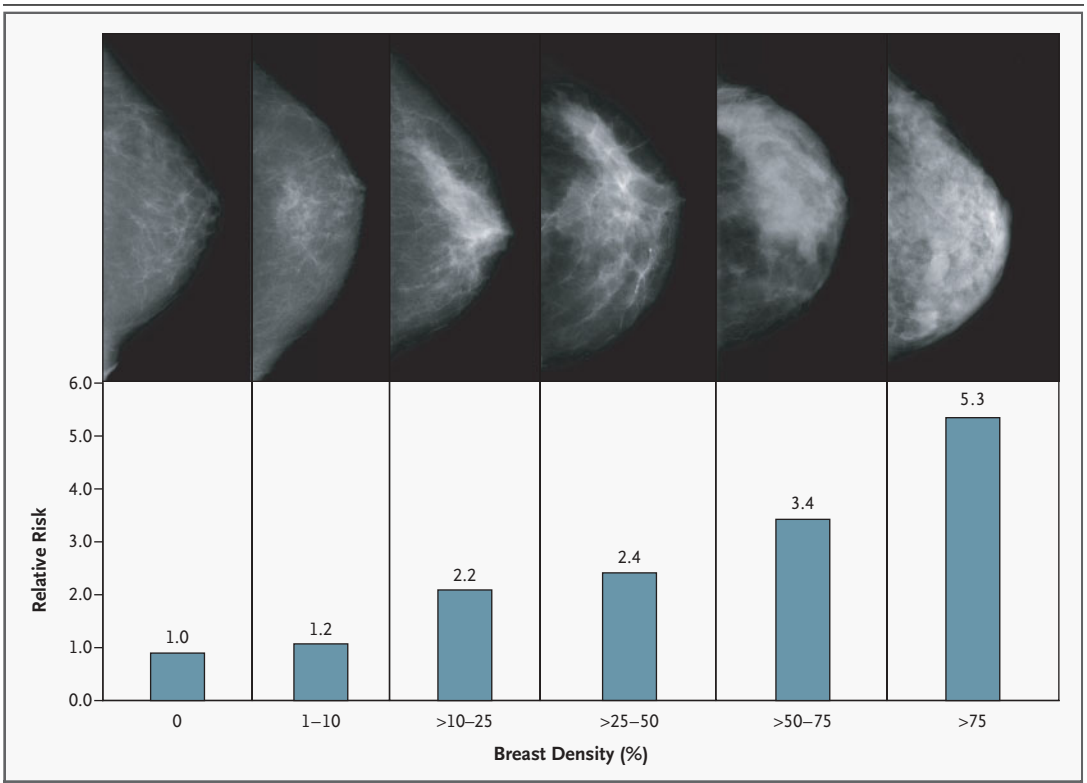


Figure 2. Risk of Breast Cancer According to Breast Density in Premenopausal and Postmenopausal Women.

crete palpable lesions in women more than 35 years of age; ultrasonography provides an optional substitute among younger women. Round dense lesions detected on mammography often are cysts that require ultrasonographic examination to distinguish them from solid lesions.

For solid lesions, core needle biopsy directed with the use of radiographic or ultrasonographic techniques provides highly discriminative information with regard to the presence or absence of malignant disease. In core needle biopsy, a large cutting needle is used with a spring-loaded, automated biopsy instrument to obtain tissue specimens suitable for histologic analysis. Fine-needle aspiration yields cellular material suitable for cytologic evaluation, but the technique must be used by an experienced cytopathologist and the specimen obtained is insufficient for diagnosis in up to 36 percent of cases in which there are nonpalpable lesions.[55]

The roles of magnetic resonance imaging (MRI) and digital mammography in the evaluation of breast lesions are currently being investigated. Galactography (also called ductography) is useful in the detection of focal lesions within a single duct. Cytologic examination of nipple discharge is of limited value, with a sensitivity for detecting malignant disease of only 35 to 47 percent.[34,56] In the treatment of all patients with benign breast disease, clinical judgment is required to provide the proper balance between the intense and frequent surveillance needed for some patients and the risk of overdiagnosis and treatment for others.

Table 3. Clinical Examination of a Patient with Benign Breast Disease.

History

Characterize symptoms

Identify risk factors for breast cancer
- Age
  - At menarche
  - At first live birth
- Number of relatives with breast cancer or ovarian cancer
  - Age at diagnosis
- Number of previous breast biopsies
- Presence of atypical hyperplasia or lobular carcinoma in situ on previous breast biopsy
- Weight gain after menopause
- Waist-to-hip ratio
- Results of bone-density testing

If patient is postmenopausal
- Age at menopause
- Duration of use of estrogen or progestin therapy

Physical examination

Palpate the four breast quadrants while patient is sitting and lying down
- Identify discrete lumps and examine for regional nodes
- Determine whether consistency is doughy with vague nodularity — findings consistent with fibrocystic changes
- Determine whether a discrete lesion has distinctly marginated borders — a finding consistent with fibroadenoma

Examine overlying skin, areola, and axilla

Determine degree of symmetry (asymmetry suggests underlying disease)

Examine nipple and seek to elicit discharge
- Determine whether galactorrhea is present
- Determine whether discharge is from one duct or from multiple ducts
- Determine whether discharge is viscous, watery, serosanguineous, grossly bloody, clear, blue–black, or green
- Determine whether occult blood is present

Seek to elicit chest-wall pain
- Examine costochondral junctions (Tietze's syndrome)
- Examine lateral chest wall while patient is lying on her side (at 90 degrees), to move breast away from chest wall
- Compare pain elicited by squeezing breast tissue with pain elicited by palpation of chest wall

TREATMENT

CYCLIC BREAST PAIN

The most important issue in the management of cyclic breast pain is to decide whether to treat. In the absence of a mass or discharge, mild symptoms warrant reassuring the patient regarding the absence of serious disease.[57] Among 85 percent of women evaluated in large referral clinics,[58] watchful waiting without treatment was considered acceptable after their anxiety about malignant disease was alleviated, whereas the remaining 15 percent requested treatment. Several clinics specializing in breast disorders administer tamoxifen and danazol for breast pain and have conducted randomized, placebo-controlled, clinical trials to demonstrate the efficacy of these strategies. Data on secondary end points from the International Breast Cancer Intervention Study, involving 7152 women who received tamoxifen for adjuvant therapy for breast cancer, provide additional evidence of the efficacy of this agent in relieving mastalgia.[59] Several other therapies are probably beneficial, on the basis of physiologic principles. Precise fitting of a brassiere to provide support for pendulous breasts may provide pain relief.[60] Lowering the dose of estrogens in the treatment of postmenopausal women or the addition of an androgen to estrogen-replacement therapy appears to be beneficial in reducing breast pain. The use of oral contraceptives has not been systematically studied, but preparations that contain low-dose estrogen (20 μg of ethinyl estradiol) and 19-nor progestins may produce relief.[61,62]

No standard regimen for moderate-to-severe breast pain has been widely accepted. Initial recom-

mendations may include the use of mild analgesic agents such as acetaminophen, nonsteroidal antiinflammatory drugs (NSAIDs), or aspirin.[32,33,56,63-67] Other approaches include tamoxifen, at a dose of 10 mg daily for three to six months,[68,69] and among patients in whom there is no response to treatment, a change to danazol, at a dose of 200 mg daily (or only during the luteal phase of the menstrual cycle).[70] Evening primrose oil has been used, at oral doses of 1 to 3 g daily, on the basis of two randomized studies,[71,72] but recent trials question its efficacy.[73,74] Gonadotropin-releasing hormone agonists have been used successfully for severe pain.[75,76]

NONCYCLIC PAIN

When pain is truly arising from the breast, the approach outlined for cyclic pain is used. However, a musculoskeletal cause is present in 40 percent of women referred to specialized mastalgia clinics for pain thought to arise from the breast. In two thirds of women with diffuse chest-wall pain, the condition responds to oral or topical NSAIDs.[77] Among the remaining patients, 85 percent gain temporary or permanent relief from the use of a combination of anesthetic and steroidal drugs injected into the tender site.[77]

FOCAL LESIONS

Careful examination distinguishes between solitary, discrete, dominant, persistent masses and vague nodularity and thickening. The Society of Surgical Oncology practice guidelines[78] recommend the following evaluation. In women 35 years of age or younger, all dominant, discrete, palpable lesions require referral to a surgeon. If vague nodularity, thickening, or asymmetric nodularity is present, the examination should be repeated at midcycle after one or two menstrual cycles. If the abnormality resolves, the patient should be reassured, and if it does not, the patient should be referred to a surgeon. Breast imaging may be appropriate. Women older than 35 years of age with a dominant mass should undergo diagnostic mammography (and frequently, ultrasonography)[51,54,55] and should then be referred to a surgeon. When vague nodularity or thickening is present, mammographic screening is required, with physical examination repeated at midcycle one to two months afterward and referral to a surgeon if the abnormality persists.

Postmenopausal women are referred for surgical consultation after undergoing mammography. For gross cysts, the guidelines suggest fine-needle aspiration with imaging studies repeated within six months. Nonbloody fluid is discarded, but if the same cyst refills, surgical consultation is warranted. If bloody, the fluid should be sent for cytologic analysis and consultation with a surgeon should be requested.

Usual practice requires the triple test (palpation; mammography, often in conjunction with ultrasonography; and biopsy) for women more than 35 years of age with dominant masses. When the results of mammographic screening are negative but a dominant mass is present, biopsy is required to rule out breast cancer, since lobular carcinoma may not be visible on mammography. Among younger women, mammography may be omitted if the results of ultrasonographic testing and biopsy yield definitive information. Many experts omit biopsy in the evaluation of younger women with lesions characteristic of fibroadenoma on ultrasonography and elect to follow these patients carefully with serial ultrasonography at six-month intervals for a period of two years and once yearly thereafter. Since careful studies have shown that a lesion that appears to be benign on mammography and ultrasonography is benign more than 99 percent of the time, some experienced clinicians opt for follow-up without biopsy.[51-54] However, other experienced surgeons disagree and believe that all fibroadenomas require diagnostic core needle biopsy or fine-needle aspiration, especially among carriers of a *BRCA* mutation, in whom medullary cancer may be found. Confirmation on biopsy of fibroadenoma eliminates the need for serial ultrasonography. For patients with a diagnosis of atypical ductal hyperplasia on fine-needle aspiration or core needle biopsy, excisional biopsy is required, because more complete resection often changes the diagnosis to ductal carcinoma in situ.

NIPPLE DISCHARGE

A practical algorithm (Fig. 3) divides discharge into two categories according to the presence or absence of galactorrhea (defined as milk production more than one year after weaning or in nulligravid or menopausal women). The presence of a discharge in association with a palpable mass and positive results on mammography or ultrasonography warrants evaluation of the mass. Galactorrhea is considered pathologic if spontaneous. A workup for galactorrhea includes measurement of prolactin and thyrotropin levels and appropriate endocrinologic evaluation and treatment if the levels are ele-

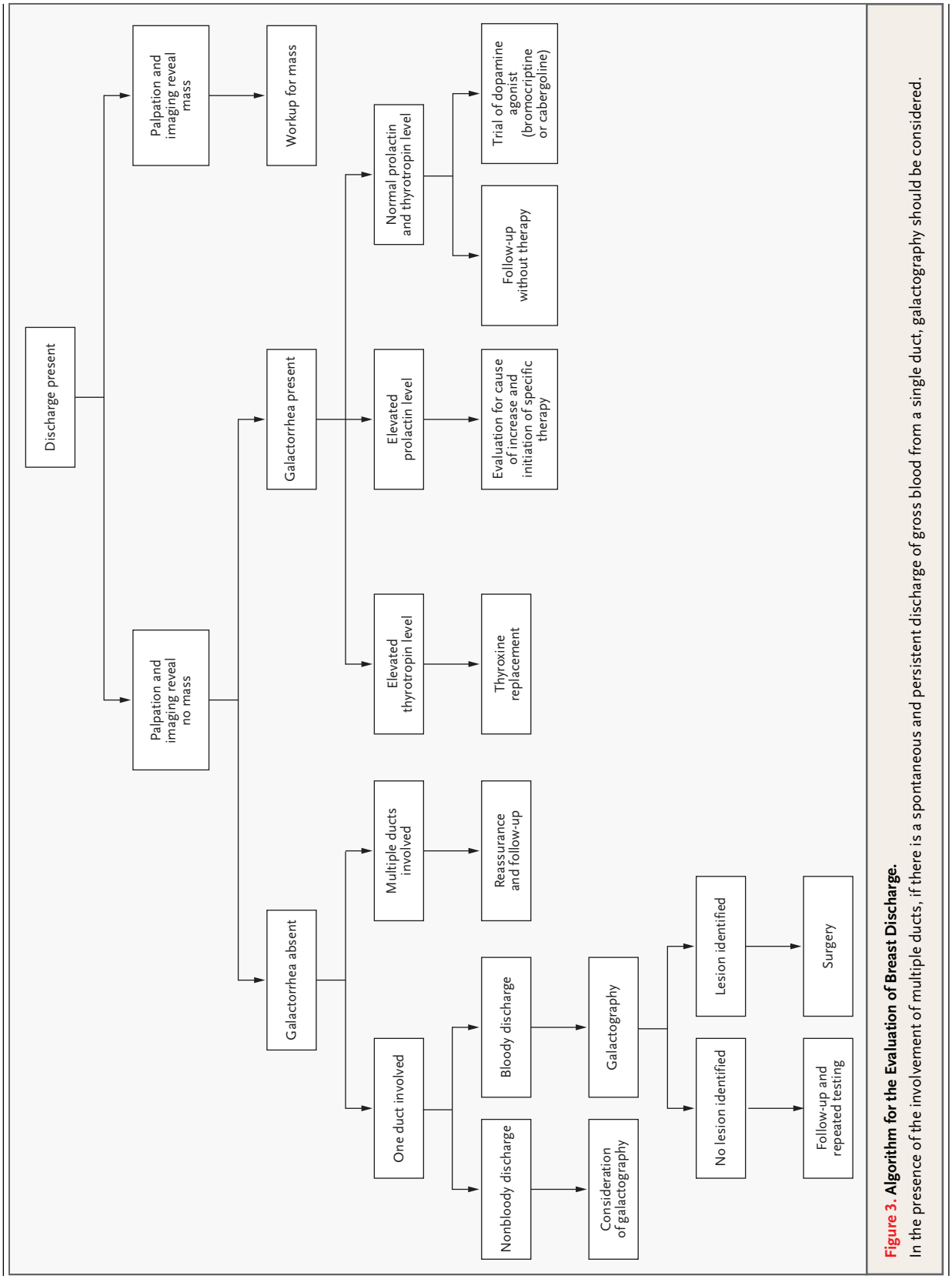


Figure 3. Algorithm for the Evaluation of Breast Discharge.
In the presence of the involvement of multiple ducts, if there is a spontaneous and persistent discharge of gross blood from a single duct, galactography should be considered.

vated.[79,80] If the levels of both are normal, treatment with dopaminergic agonists may be initiated if the patient desires to reduce the fluid leak.

A discharge in the absence of galactorrhea is considered to be ductal in origin and is classified as either uniductal or multiductal.[80] When the discharge is from one duct, and particularly if it is grossly bloody or the results of testing for occult blood are positive, a further workup is needed. Galactography with the use of cannulation and insertion of dye into the single duct emitting blood at the nipple allows visualization of a space-occupying lesion. Not all imaging centers have expertise with this technique, and alternatively, surgical biopsy can be used to define the lesion. Direct examination of the ducts by means of mini-fiberoptic endoscopy with a 0.65-mm (outer diameter) cannula (ductoscopy) is currently being evaluated in research centers.[81] Ductal exploration allows the removal of pathologic lesions and cessation of the discharge. Multiductal discharge that is clear, serous, green–black, or nonbloody requires only reassurance of the patient. Blood arising predominantly from one or two ducts should be evaluated further. Lesions that are commonly found are listed in Table 2.

PREVENTION OF BREAST CANCER

Patients with benign breast lesions that impart an increased risk of breast cancer can be offered tamoxifen as a preventive strategy. The risk of breast cancer is determined with the use of the Gail model (or among women with a family history that includes second-degree maternal or paternal relatives with breast cancer, the Claus model). Decisions are based on an evaluation of the benefits as compared with the risks of tamoxifen (http://smi-web.stanford.edu/people/pcheng/breastca).[48,49] Risk factors not included in either the Gail or the Claus model include the degree of breast density, plasma levels of free estradiol, bone density, weight gain after menopause, and waist-to-hip ratio.[44,45,82,83] Current recommendations suggest that women with a five-year risk of breast cancer of more than 1.67 percent and no contraindications to tamoxifen therapy should be informed about the option of taking tamoxifen for five years.[84-88] A recent overview of breast-cancer–prevention trials showed a reduction of 50 percent in the relative risk of breast cancer with tamoxifen, but the benefits may be offset by the increased risk of thromboembolic phenomena, endometrial cancer, and the maturation of cataracts.[88] The ongoing Study of Tamoxifen and Raloxifene (STAR) is addressing whether raloxifene might be preferable to tamoxifen.[89] More intensive and frequent screening with the use of multimodality imaging (e.g., digital or standard mammography plus ultrasonography or MRI) may be required for high-risk patients.

REFERENCES

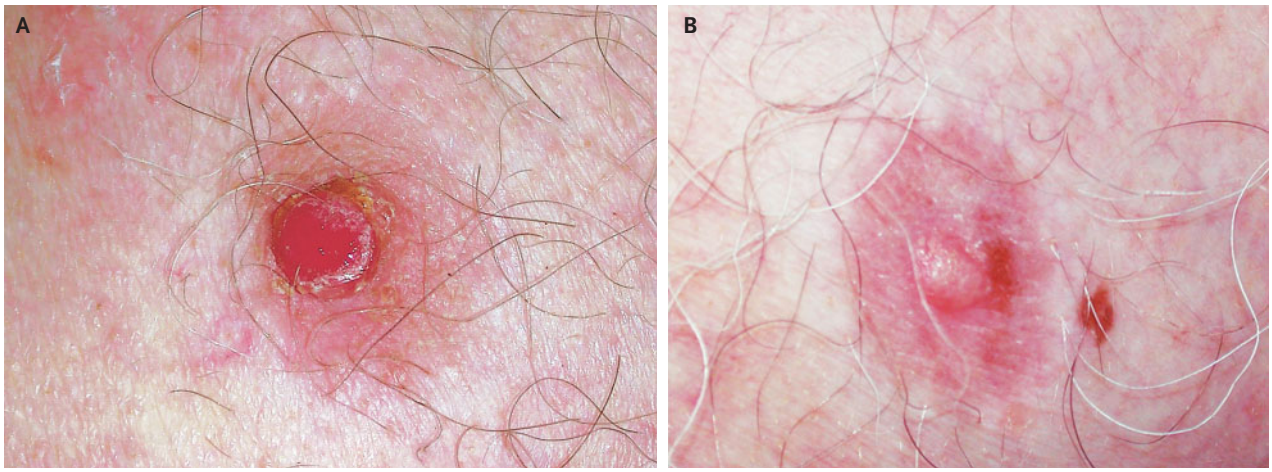
1. Russo J. Hormonal control of breast development. In: DeGroot LJ, Jameson JL, Burger H, et al., eds. Endocrinology. Philadelphia: W.B. Saunders, 2001:2181-8.
2. Ginger MR, Gonzalez-Rimbau MF, Gay JP, Rosen JM. Persistent changes in gene expression induced by estrogen and progesterone in the rat mammary gland. Mol Endocrinol 2001;15:1993-2009.
3. Osborne MP. Breast anatomy and development. In: Harris JR, Osborne CK, Morrow M, Lippman ME, eds. Diseases of the breast. Philadelphia: Lippincott Williams & Wilkins, 2000:1-13.
4. Potten CS, Watson RJ, Williams GT, et al. The effect of age and menstrual cycle upon proliferative activity of the normal human breast. Br J Cancer 1988;58:163-70.
5. Hughes LE, Mansel RE, Webster DJ. Aberrations of normal development and involution (ANDI): a new perspective on pathogenesis and nomenclature of benign breast disorders. Lancet 1987;2:1316-9.
6. Love SM, Gelman RS, Silen W. Fibrocystic "disease" of the breast — a nondisease? N Engl J Med 1982;307:1010-4.
7. Houssami N, Cheung MN, Dixon JM. Fibroadenoma of the breast. Med J Aust 2001;174:185-8.
8. Goehring C, Morabia A. Epidemiology of benign breast disease, with special attention to histologic types. Epidemiol Rev 1997;19:310-27.
9. Dent DM, Hacking EA, Wilkie W. Benign breast disease: clinical classification and disease distribution. Br J Clin Pract 1988;42:Suppl:69-71.
10. Costantini L, Bucchi L, Dogliotti L, et al. Cohort study of women with aspirated gross cysts of the breast — an update. In: Mansel RE, ed. Recent developments in the study of benign breast disease. London: Parthenon, 1993:227-39.
11. Sandison AT. An autopsy study of the adult human breast; with special reference to proliferative epithelial changes of importance in the pathology of the breast. In: National Cancer institute monograph 8. Washington, D.C.: Government Printing Office, 1962:1-90.
12. Is 'fibrocystic disease' of the breast precancerous? Arch Pathol Lab Med 1986;110:171-3.
13. Hartmann LC, Sellers TA, Frost MH, et al. Benign breast disease and the risk of breast cancer. N Engl J Med 2005;353:229-37.
14. Shaaban AM, Sloane JP, West CR, et al. Histopathologic types of benign breast lesions and the risk of breast cancer: case-control study. Am J Surg Pathol 2002;26:421-30.
15. Rohan TE, Miller AB. Hormone replacement therapy and risk of benign proliferative epithelial disorders of the breast. Eur J Cancer Prev 1999;8:123-30.
16. Tan-Chiu E, Wang J, Costantino JP, et al. Effects of tamoxifen on benign breast disease in women at high risk for breast cancer. J Natl Cancer Inst 2003;95:302-7.
17. Roger P, Sahla ME, Makela S, Gustafsson JA, Baldet P, Rochefort H. Decreased ex-

pression of estrogen receptor beta protein in proliferative preinvasive mammary tumors. Cancer Res 2001;61:2537-41.
18. Allred DC, Mohsin SK. Biological features of human premalignant breast disease. In: Harris JR, Osborne CK, Morrow M, Lippman ME, eds. Diseases of the breast. Philadelphia: Lippincott Williams & Wilkins, 2000:355-66.
19. O'Connell P, Pekkel V, Fuqua SA, Osborne CK, Clark GM, Allred DC. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. J Natl Cancer Inst 1998;90:697-703.
20. Allred DC, Mohsin SK, Fuqua SA. Histological and biological evolution of human premalignant breast disease. Endocr Relat Cancer 2001;8:47-61.
21. Lundin C, Mertens F. Cytogenetics of benign breast lesions. Breast Cancer Res Treat 1998;51:1-15.
22. Micale MA, Visscher DW, Gulino SE, Wolman SR. Chromosomal aneuploidy in proliferative breast disease. Hum Pathol 1994;25:29-35.
23. Hoogerbrugge N, Bult P, de Widt-Levert LM, et al. High prevalence of premalignant lesions in prophylactically removed breasts from women at hereditary risk for breast cancer. J Clin Oncol 2003;21:41-5.
24. Boone CW, Kelloff GJ, Freedman LS. Intraepithelial and postinvasive neoplasia as a stochastic continuum of clonal evolution, and its relationship to mechanisms of chemopreventive drug action. J Cell Biochem Suppl 1993;17G:14-25.
25. Bardelli A, Parsons DW, Silliman N, et al. Mutational analysis of the tyrosine kinome in colorectal cancers. Science 2003; 300:949.
26. Preece PE, Mansel RE, Bolton PM, Hughes LM, Baum M, Gravelle IH. Clinical syndromes of mastalgia. Lancet 1976;2: 670-3.
27. Khan SA, Apkarian AV. The characteristics of cyclical and non-cyclical mastalgia: a prospective study using a modified McGill Pain Questionnaire. Breast Cancer Res Treat 2002;75:147-57.
28. Ader DN, South-Paul J, Adera T, Deuster PA. Cyclical mastalgia: prevalence and associated health and behavioral factors. J Psychosom Obstet Gynecol 2001;22:71-6.
29. Kessel B. Premenstrual syndrome: advances in diagnosis and treatment. Obstet Gynecol Clin North Am 2000;27:625-39.
30. Goodwin PJ, Miller A, Del Giudice ME, Ritchie K. Breast health and associated premenstrual symptoms in women with severe cyclic mastopathy. Am J Obstet Gynecol 1997;176:998-1005.
31. Ader DN, Browne MW. Prevalence and impact of cyclic mastalgia in a United States clinic-based sample. Am J Obstet Gynecol 1997;177:126-32.
32. Goodwin PJ, Neelam M, Boyd NF. Cyclical mastopathy: a critical review of therapy. Br J Surg 1988;75:837-44.
33. Santen R, Pinkerton J. Benign breast disorders. In: DeGroot LJ, Jameson JL, Burger H, et al., eds. Endocrinology. Philadelphia: W.B. Saunders, 2002:2189-98.
34. Ambrogetti D, Berni D, Catarzi S, Ciatto S. The role of ductal galactography in the differential diagnosis of breast carcinoma. Radiol Med (Torino) 1996;91:198-201. (In Italian.)
35. Seltzer MH, Perloff LJ, Kelley RI, Fitts WT Jr. The significance of age in patients with nipple discharge. Surg Gynecol Obstet 1970;131:519-22.
36. Dupont WD, Parl FF, Hartmann WH, et al. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 1993;71:1258-65.
37. London SJ, Connolly JL, Schnitt SJ, Colditz GA. A prospective study of benign breast disease and the risk of breast cancer. JAMA 1992;267:941-4. [Erratum, JAMA 1992;267:1780.]
38. Wang J, Costantino JP, Tan-Chiu E, Wickerham DL, Paik S, Wolmark N. Lower-category benign breast disease and the risk of invasive breast cancer. J Natl Cancer Inst 2004;96:616-20.
39. Wolfe JN, Saftlas AF, Salane M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. AJR Am J Roentgenol 1987;148:1087-92.
40. Boyd NF, Jensen HM, Cooke G, Han HL, Lockwood GA, Miller AB. Mammographic densities and the prevalence and incidence of histological types of benign breast disease. Eur J Cancer Prev 2000;9:15-24.
41. Byrne C, Schairer C, Brinton LA, et al. Effects of mammographic density and benign breast disease on breast cancer risk (United States). Cancer Causes Control 2001;12:103-10.
42. Boyd NF, Dite GS, Stone J, et al. Heritability of mammographic density, a risk factor for breast cancer. N Engl J Med 2002; 347:886-94.
43. Hankinson SE, Willett WC, Manson JE, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 1998;90:1292-9.
44. Key T, Appleby P, Barnes I, Reeves G, Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002;94:606-16.
45. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. Maturitas 2001;38:103-13.
46. Antoniou A, Pharoah PDP, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet 2003;72:1117-30. [Erratum, Am J Hum Genet 2003;73:709.]
47. Dooley WC, Ljung BM, Veronesi U, et al. Ductal lavage for detection of cellular atypia in women at high risk for breast cancer. J Natl Cancer Inst 2001;93:1624-32.
48. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879-86.
49. McTiernan A, Kuniyuki A, Yasui Y, et al. Comparisons of two breast cancer risk estimates in women with a family history of breast cancer. Cancer Epidemiol Biomarkers Prev 2001;10:333-8.
50. Barton MB, Harris R, Fletcher SW. Does this patient have breast cancer? JAMA 1999; 282:1270-80.
51. Shetty MK, Shah YP, Sharman RS. Prospective evaluation of the value of combined mammographic and sonographic assessment in patients with palpable abnormalities of the breast. J Ultrasound Med 2003;22: 263-8.
52. Moy L, Slanetz PJ, Moore R, et al. Specificity of mammography and US in the evaluation of a palpable abnormality: retrospective review. Radiology 2002;225:176-81.
53. Soo MS, Rosen EL, Baker JA, Vo TT, Boyd BA. Negative predictive value of sonography with mammography in patients with palpable breast lesions. AJR Am J Roentgenol 2001;177:1167-70.
54. Flobbe K, Bosch AM, Kessels AG, et al. The additional diagnostic value of ultrasonography in the diagnosis of breast cancer. Arch Intern Med 2003;163:1194-9.
55. Venta LA. Image-guided biopsy of nonpalpable breast lesions. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the breast. 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2000:149-64.
56. Groves AM, Carr M, Wadhera V, Lennard TWJ. An audit of cytology in the evaluation of nipple discharge: a retrospective study of 10 years' experience. Breast 1996;5: 96-9.
57. Gateley CA, Miers M, Mansel RE, Hughes LE. Drug treatments for mastalgia: 17 years experience in the Cardiff Mastalgia Clinic. J R Soc Med 1992;85:12-5.
58. Hughes LE, Mansel RE, Webster DJT. Benign breast disorders and diseases of the breast. 2nd ed. London: W.B. Saunders, 1999.
59. Cuzick J, Forbes J, Edwards R, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): a randomised prevention trial. Lancet 2002;360:817-24.
60. Wilson MC, Sellwood RA. Therapeutic value of a supporting brassiere in mastodynia. Br Med J 1976;2:90.
61. Charreau I, Plu-Bureau, Bachelot A, Contesso G, Guinebretiere JM, Le MG. Oral contraceptive use and risk of benign breast disease in a French case-control study of young women. Eur J Cancer Prev 1993;2: 147-54.
62. Mauvais-Jarvis P. Mastodynia and fibrocystic disease. Curr Ther Endocrinol Metab 1988;3:280-4.

63. Millet AV, Dirbas FM. Clinical management of breast pain: a review. Obstet Gynecol Surv 2002;57:451-61.
64. BeLieu RM. Mastodynia. Obstet Gynecol Clin North Am 1994;21:461-77.
65. Zylstra S. Office management of benign breast disease. Clin Obstet Gynecol 1999; 42:234-48.
66. Steinbrunn BS, Zera RT, Rodriguez JL. Mastalgia: tailoring treatment to type of breast pain. Postgrad Med 1997;102:183-4.
67. Wetzig NR. Mastalgia: a 3 year Australian study. Aust N Z J Surg 1994;64:329-31.
68. Fentiman IS, Caleffi M, Brame K, Chaudary MA, Hayward JL. Double-blind controlled trial of tamoxifen therapy for mastalgia. Lancet 1986;1:287-8.
69. Messinis IE, Lolis D. Treatment of premenstrual mastalgia with tamoxifen. Acta Obstet Gynecol Scand 1988;67:307-9.
70. O'Brien PM, Abukhalil IE. Randomized controlled trial of the management of premenstrual syndrome and premenstrual mastalgia using luteal phase-only danazol. Am J Obstet Gynecol 1999;180:18-23.
71. Pashby NL, Mansel RE, Hughes LE, Hanslip JI, Preece P. A clinical trial of evening primrose oil in mastalgia. Br J Surg 1981;68:801-24.
72. Preece PE, Hanslip JI, Gilbert L. Evening primrose oil (Efamol) for mastalgia. In: Horrobin DF, ed. Clinical uses of essential fatty acids. Montreal: Eden Press, 1982:147-54.
73. Goyal A, Mansel RE, Efamast Study Group. A randomized multicenter study of gamolenic acid (Efamast) with or without antioxidants, vitamins and minerals in the management of mastalgia. Breast J 2005;11: 41-7.
74. Blommers J, de Lange-De Klerk ES, Kuik DJ, Bezemer PD, Meijer S. Evening primrose oil and fish oil for severe chronic mastalgia: a randomized, double-blind, controlled trial. Am J Obstet Gynecol 2002;187:1389-94.
75. Hamed H, Caleffi M, Chaudary MA, Fentiman IS. LHRH analogue for treatment of recurrent and refractory mastalgia. Ann R Coll Surg Engl 1990;72:221-4.
76. Mansel RE, Goyal A, Preece P, et al. European randomized, multicenter study of goserelin (Zoladex) in the management of mastalgia. Am J Obstet Gynecol 2004;191: 1942-9.
77. Kollias J, Sibbering DM, Blamey RW. Topical non-steroidal anti-inflammatory gel for diffuse chest wall pain in mastalgia patients. In: Mansel RE, ed. Recent developments in the study of benign breast disease. London: Parthenon Publishing, 1997:119-24.
78. Morrow M, Bland KI, Foster R. Breast cancer surgical practice guidelines: Society of Surgical Oncology practice guidelines. Oncology (Huntingt) 1997;11:877-81.
79. Mah PM, Webster J. Hyperprolactinemia: etiology, diagnosis, and management. Semin Reprod Med 2002;20:365-74.
80. Falkenberry SS. Nipple discharge. Obstet Gynecol Clin North Am 2002;29:21-9.
81. Mokbel K, Elkak AE. The evolving role of mammary ductoscopy. Curr Med Res Opin 2002;18:30-2.
82. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. Cancer Epidemiol Biomarkers Prev 1998;7:1133-44.
83. Kuller LH, Cauley JA, Lucas L, Cummings S, Browner WS. Sex steroid hormones, bone mineral density, and risk of breast cancer. Environ Health Perspect 1997; 105:Suppl 3:593-9.
84. Chlebowski RT, Collyar DE, Somerfield MR, Pfister DG. American Society of Clinical Oncology technology assessment on breast cancer risk reduction strategies: tamoxifen and raloxifene. J Clin Oncol 1999; 17:1939-55.
85. Chlebowski RT, Col N, Winer EP, et al. American Society of Clinical Oncology technology assessment of pharmacologic interventions for breast cancer risk reduction including tamoxifen, raloxifene, and aromatase inhibition. J Clin Oncol 2002;20: 3328-43.
86. Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. J Natl Cancer Inst 1999;91:1829-46. [Erratum, J Natl Cancer Inst 2000;92:275.]
87. Levine M, Moutquin JM, Walton R, Feightner J. Chemoprevention of breast cancer: a joint guideline from the Canadian Task Force on Preventive Health Care and the Canadian Breast Cancer Initiative's Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. CMAJ 2001;164:1681-90.
88. Cuzick J, Powles T, Veronesi U, et al. Overview of the main outcomes in breast-cancer prevention trials. Lancet 2003;361: 296-300.
89. Pappas SG, Jordan VC. Chemoprevention of breast cancer: current and future prospects. Cancer Metastasis Rev 2002;21: 311-21.

IMAGES IN CLINICAL MEDICINE

Adenocarcinoma of the Breast in a Man



Andrew Birnie, M.R.C.P.
Sandeep Varma, M.R.C.P.

Queen's Medical Centre
Nottingham NG7 2UH, United Kingdom

An 83-year-old man presented with a one-month history of an enlarging lump on his right nipple. It was asymptomatic. Physical examination revealed an 11-mm erythematous, well-circumscribed nodule attached to the right nipple, which had begun to ulcerate in the center (Panel A), and a normal left nipple (Panel B). The nodule was excised with a 5-mm margin. Histologic examination showed an intraductal and invasive ductal adenocarcinoma of the breast, with the cancer invading the nipple from below. The patient underwent a complete mastectomy and axillary-node dissection. No residual carcinoma was found, and all lymph nodes were free of cancer, indicating a good prognosis. A delay in diagnosis is common in men with breast cancer because patients and physicians do not recognize the clinical features. Breast cancer affects more than 1600 men per year in the United States, and treatment is currently based on the same guidelines used for the treatment of breast cancer in women.

Web-only Images in Clinical Medicine are published every week in the *Journal*. They are listed (with e page numbers) in the table of contents on the cover of the printed *Journal* and can be seen at **www.nejm.org**.

CASE RECORDS of the MASSACHUSETTS GENERAL HOSPITAL

Founded by Richard C. Cabot
Nancy Lee Harris, M.D., *Editor*
Jo-Anne O. Shepard, M.D., *Associate Editor*
Sally H. Ebeling, *Assistant Editor*
Stacey M. Ellender, *Assistant Editor*
Christine C. Peters, *Assistant Editor*

Case 22-2005: An 81-Year-Old Man with Cough, Fever, and Altered Mental Status

Harry Hollander, M.D., Pamela W. Schaefer, M.D., and E. Tessa Hedley-Whyte, M.D.

PRESENTATION OF CASE

Dr. Sherry Chou (Neurology): An 81-year-old man was admitted to the Massachusetts General Hospital in September because of fever, chills, productive cough, and diffuse weakness.

Two weeks before admission, a cough developed, with small amounts of yellowish sputum. The patient was treated with azithromycin for five days, without improvement. Five days before admission, amoxicillin was begun. The cough continued to worsen, with increased production of sputum that was tinged with blood. High fevers and chills developed, with progressive anorexia, diffuse weakness, and mild confusion. He was brought to the emergency department and admitted.

A diagnosis of chronic lymphocytic leukemia (CLL) had been made six years earlier. The patient had been treated for two years with fludarabine, which was stopped because anemia and thrombocytopenia had developed. For three years before admission, he received no treatment for his leukemia and had been clinically stable, with slowly rising peripheral lymphocyte counts and falling serum immunoglobulin levels (Table 1). He had had several episodes of viral and pneumococcal pneumonia in the two years before admission. Other medical problems included atrial fibrillation, diabetes mellitus, hypertension, prostatic adenocarcinoma (which had been treated with radiation therapy seven years earlier), multiple cutaneous squamous-cell carcinomas (treated with surgical excision), and three melanomas on the face and neck. The most recent melanoma had been treated within the six months before admission with excision and skin grafting.

The patient lived at home with his wife in the Boston area. He had traveled to California six months before and to New York one month before becoming ill. He had no known contacts with sick people or recent exposures to animals or insects. He did not have nausea, vomiting, or a change in bowel habits. His medications on admission included potassium chloride, digoxin, hydrochlorothiazide, atenolol, warfarin, glyburide, and amoxicillin.

On examination, the patient appeared tired but was able to converse appropriately. The axillary temperature was 39.6° C. The pulse was irregular at 100 beats per minute, the blood pressure 170/90 mm Hg, and the respiratory rate 24 breaths per minute. Oxygen saturation was 98 percent with the patient breathing 2 liters of supplemental oxygen delivered by nasal cannula. The pupils were equal and reactive to light. The neck was

From the Division of Infectious Diseases, Department of Medicine, University of California at San Francisco, San Francisco (H.H.); the Departments of Radiology (P.W.S.) and Pathology (E.T.H.-W.), Massachusetts General Hospital, Boston; and the Departments of Radiology (P.W.S.) and Pathology (E.T.H.-W.), Harvard Medical School, Boston.

N Engl J Med 2005;353:287-95.

Table 1. Hematologic and Immunologic Laboratory Data.

Variable	Normal Range	5 Yr before Admission	3 Yr before Admission	On Admission	Hospital Day 5	Hospital Day 12
White cells (per mm³)	4500–11,000	23,700	36,100	42,000	95,900	54,900
Hematocrit (%)	41.0–53.0	39.6	37.7	37.1	31.3	33.6
Hemoglobin (g/dl)	13.5–17.5	13.8	12.6	12.5	10.3	11.8
Red cells (per mm³)	4,500,000–5,900,000	4,570,000	4,490,000	4,820,000	4,050,000	3,970,000
Platelets (per mm³)	150,000–350,000	361,000	96,000	97,000	189,000	99,000
Mean corpuscular volume (μm³)	80–100	87	84	77	77	85
Differential count (%)						
Neutrophils		36	16	13	19	11
Band forms			4	4	2	2
Lymphocytes		57	70	51	54	78
Atypical lymphocytes			7	30	11	8
Monocytes		5		2	2	1
Eosinophils		1	3	0	3	0
Basophils		1	0	0	9	0
Prothrombin time (sec)				16.6	15.5	12.8
Activated prothrombin time (sec)				49.4	52.0	55.9
Prothrombin time (international normalized ratio)				1.8	1.6	1.1
IgG (mg/dl)	614–1295	475	310		130	
IgA (mg/dl)	60–309	110	86		50	
IgM (mg/dl)	53–334	28	19		6	

supple. There was no palpable cervical lymphadenopathy, but slightly enlarged lymph nodes were present in the axillae. The chest was clear to auscultation. There was a 2/6 systolic ejection murmur at the right upper sternal border. The abdomen was nontender, nondistended, and soft. The tip of the spleen was palpable. The remainder of the examination revealed no abnormalities. On neurologic examination, the patient was mildly confused and was not oriented to place or time. He did not follow commands. The strength in the arms and legs was 3/5 bilaterally. There was marked truncal weakness, and he was unable to sit up in bed.

A chest radiograph showed prominence of the right paratracheal stripe and right hilum and bilateral small pleural effusions. Blood levels of electrolytes and serum levels of urea nitrogen, creatinine, glucose, albumin, bilirubin, alkaline phosphatase, and aminotransferase were normal. The amylase level was 106 units per liter and the lipase level 7.1 units per liter; the globulin level was 2.3 g per deciliter. Hematologic laboratory values are shown in Table 1. Flow cytometry of peripheral-blood lymphocytes revealed that they were CD19+ and CD20+ B cells coexpressing CD5 and CD23 and monotypic lambda immunoglobulin light chains. Specimens of urine, blood, nasal secretions, and sputum were sent for culture and viral testing. Ampicillin, gentamycin, and metronidazole were administered.

On the second hospital day, the patient remained febrile, with a temperature that peaked at 39.1°C. He continued to be confused, and by the third day he was somnolent. His degree of arousal waxed and waned. Right-sided weakness developed. Computed tomographic (CT) scanning of the head after the administration of intravenous contrast material revealed no abnormalities except for mild mucosal thickening of the ethmoid air cells. CT scanning of the chest showed mediastinal and axillary lymphadenopathy that had enlarged since a study obtained five months earlier. There was no hilar lymphadenopathy or pulmonary abnormalities. A lumbar puncture was performed (Table 2). Flow cytometry showed that the lymphocytes were T cells (90 percent) with a normal immunophenotype and a normal CD4:CD8 ratio; there were a small number of B

Table 2. Results of Cerebrospinal Fluid Analysis.*

Variable	Hospital Day 2	Hospital Day 5
Opening pressure (cm H_2O)	24	
Glucose (mg/dl)	61	89
Protein (mg/dl)	72	65
White cells (per high-power field)	253	161
Neutrophils (%)	14	1
Lymphocytes (%)	85	90
Monocytes (%)	1	9
Red cells (per high-power field)	9	161
Gram stain	No organisms	
Fluid culture	Negative	
Fungi		
Wet preparation	No fungi seen	
Culture	Negative	
Acid-fast bacilli (smear)	Negative	Negative
Mycobacteria (culture)	Negative	
Cryptococcal antigen	Negative	
Histoplasma antigen		Negative
General viral culture	Negative	
Encephalitis (antibody panel)	Negative	
Syphilis (VDRL test)	Nonreactive	
Bartonella antibody		
B. henselae		Negative
B. quintana		Negative
HSV (PCR)	Negative	
Enterovirus (PCR)	Negative	
West Nile virus (PCR)	Negative	

* VDRL denotes Venereal Disease Research Laboratory, HSV herpes simplex virus, and PCR polymerase chain reaction.

cells with an immunophenotype identical to that of cells associated with CLL.

Acyclovir and ceftriaxone were begun. On the fourth hospital day, the patient became progressively more somnolent, and hypoxemic respiratory failure developed (the partial pressure of oxygen was 63 mm Hg with the fraction of inspired oxygen at 40 percent). The trachea was intubated, and the patient was transferred to the medical intensive care unit. He had only minimal motor response to pain. He had a weak grimace in response to nasal tickle and a cough reflex. Nuchal rigidity was present. Repeated CT scanning of the head showed no changes. An abdominal ultrasonographic study showed no gallstones; the spleen was enlarged. A culture of sputum that had been obtained on the first hospital day grew *Pseudomonas aeruginosa*. The results of other laboratory tests are shown in Tables 1, 2, and 3. Ceftriaxone was discontinued and ceftazidime was begun.

By the fifth hospital day, the patient was comatose. Magnetic resonance imaging (MRI) of the brain showed a region of restricted diffusion in the left thalamus with surrounding T_2 hyperintensity that involved most of the left thalamus. There was a 5-mm-long, enhancing focus, T_1 isointense and T_2 hyperintense, that was located just inferior to the left hypoglossal canal and thought to be a schwannoma. Magnetic resonance angiography showed mild irregularity, including foci of slight ectasia of the major cerebral arteries, that was consistent with atherosclerosis. Contrast-enhanced CT scanning of the abdomen and pelvis disclosed a small calcification in the spleen, multiple cysts in both kidneys, and an enlarged periaortic lymph node that measured 20 mm by 27 mm.

Over the course of the next seven days, the patient remained comatose and unresponsive. A dose of intravenous immunoglobulin was given on the seventh hospital day. After discussion with his family, ventilator support was withdrawn on the 12th hospital day, and the patient died. An autopsy was performed.

DIFFERENTIAL DIAGNOSIS

Dr. Harry Hollander: May we review the imaging studies?

Dr. Pamela W. Schaefer: Fluid-attenuated inversion-recovery images and T_2-weighted images from the MRI examination performed on the fifth hospital day (Fig. 1) show subtle hyperintensity throughout most of the left thalamus, with mild expansion. On diffusion-weighted images, there is a punctate focus of decreased diffusion within the left thalamus, but most of the left thalamus has relatively normal diffusion. No abnormal enhancement is seen.

The differential diagnosis for the left thalamic lesion includes viral encephalitis and bartonella infection, among other infectious and inflammatory processes, low-grade glioma, subacute arterial or venous infarction, paraneoplastic syndromes, and acute disseminated encephalomyelitis.

Dr. Hollander: A central issue in framing the differential diagnosis in this case is the accurate characterization of the central nervous system syndrome.

Table 3. Microbiologic Test Results.*

Test	Source	Result (Hospital Day)
Adenovirus antigen	Nasal swab	Negative (day 2)
Influenza A and B antigen	Nasal swab	Negative (day 2)
Parainfluenza virus 1, 2, and 3 antigens	Nasal swab	Negative (day 2)
RSV antigen	Nasal swab	Negative (day 2)
CMV antigen	Blood	Negative (day 2)
Lyme antibody	Blood	Negative (day 2)
Heterophile antibody	Blood	Negative (day 2)
Encephalitis antibody panel (West Nile virus IgG and IgM, eastern equine encephalitis IgG and IgM, LAC, SLE, POW/TBE IgM)	Blood	Negative (day 2)
Cryptococcal antigen	Blood	Negative (day 3)
EBV antibody panel (IgG, IgM, EBV antigen, EBV nuclear antigen)	Blood	Negative (day 3)
Fungal antibodies	Blood	
Histoplasmosis		Negative
Blastomycosis		Negative
Coccidioidomycosis		Negative
Varicella antibody (IgG)	Blood	Negative (day 6)
Legionella culture	Sputum	Negative
Acid-fast smear	Sputum	Negative
Mycobacterial culture	Sputum	Negative
Fungal wet preparation	Sputum	Negative
Fungal culture	Sputum	Negative
Adenovirus antigen	Sputum	Negative
Influenza A and B antigen	Sputum	Negative
Parainfluenza virus 1, 2, or 3 antigen	Sputum	Negative
RSV antigen	Sputum	Negative
Mycoplasma pneumoniae PCR	Sputum	Negative

* RSV denotes respiratory syncytial virus, CMV cytomegalovirus, LAC La Crosse encephalitis virus, SLE St. Louis encephalitis virus, POW Powassan encephalitis virus, TBE tick-borne encephalitis, EBV Epstein–Barr virus, and PCR polymerase chain reaction.

DEFINING THE NEUROLOGIC SYNDROME

The presence of fever and the physical findings of meningeal irritation and cerebrospinal fluid pleocytosis all support the conclusion that a component of meningitis could be part of the syndrome under discussion. Perturbations of cognition and level of consciousness are often seen in patients with meningitis. However, altered mental status that precedes the onset of meningeal findings and worsens in the setting of relatively mild, stable abnormalities of the cerebrospinal fluid suggests that encephalitis is actually the primary process.

The profound peripheral weakness that was observed in this patient cannot be explained on the basis of meningoencephalitis alone and suggests involvement of additional levels of the neuraxis. I assume that the weakness was hypotonic and that the respiratory arrest later in the course probably represented ventilatory failure caused by involvement of the respiratory musculature. A lesion in the cervical spinal cord would probably be accompanied by sensory changes and sphincter dysfunction. Polyradiculopathy tends to be asymmetric, often preferentially affects the lumbosacral plexus, and is characterized by mixed motor and sensory findings rather than isolated motor involvement. Patients with a demyelinating neuropathy often have pure motor findings, but typically, the cerebrospinal fluid does not show the degree of inflammation observed here. Thus, the most likely site of neuropathology that explains this degree of weakness is diffuse involvement of anterior horn cells, the ventral roots of the spinal cord, or both. In summary, this patient's clinical findings indicate a syndrome of meningomyeloencephalitis or encephalomyelitis, and this must be the focus of the differential diagnosis.

COMPLICATIONS OF CLL

CLL usually follows an indolent, minimally symptomatic course for many years, as it did in this man. When a patient with this underlying diagnosis presents with new clinical problems, there are several types of complications that must be considered: autoimmune manifestations, direct effects of tumor infiltration or biologic transformation (Richter's syndrome), complications of treatment, and intercurrent infections.

Most autoimmune complications of CLL are hematologic, with worsening thrombocytopenia or anemia as a result of autoantibody formation. Once hypogammaglobulinemia develops, autoimmune disease is a less likely explanation for new clinical complications. This patient's syndrome and the absence of radiographic findings indicating demyelination argue against an autoimmune cause for his problems. Finally, CLL is not a tumor associated with paraneoplastic autoimmune manifestations such as the anti-Hu syndrome.

A more difficult issue is determining whether this patient's decline is a result of increasingly aggressive behavior or biologic transformation of

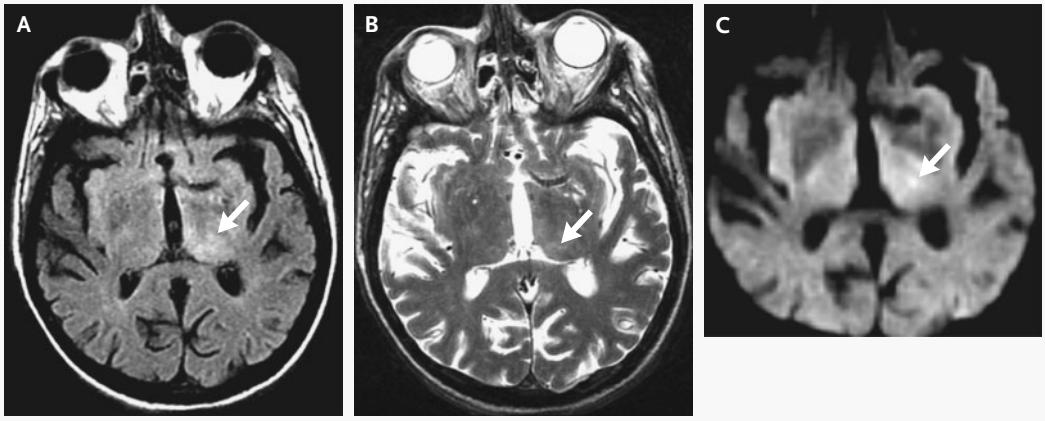


Figure 1. Magnetic Resonance Images on the Fifth Hospital Day.

A fluid-attenuated inversion-recovery image (Panel A) and a T_2-weighted image (Panel B) show subtle hyperintensity in the left thalamus (arrow, Panels A, B, and C). On the diffusion-weighted image (Panel C), there is a punctate hyperintense focus that suggests ischemia. Most of the left thalamus is relatively isointense to, and has diffusion similar to, the remainder of the brain parenchyma.

the CLL. Progressive leukocytosis and hypogammaglobulinemia, increasing mediastinal and retroperitoneal lymphadenopathy, and, perhaps, the development of pleural effusions all raise the possibility of transformation to high-grade lymphoma (Richter's syndrome). The stable immunophenotype of the circulating leukemic cells argues against this, and tissue specimens must be obtained to establish the diagnosis of high-grade lymphoma. Unlike myeloid leukemias, CLL rarely produces intravascular leukostasis leading to organ-system dysfunction, even with extreme elevation of the peripheral white-cell count. Treatment complications are not an issue in this case, because the patient had received no recent therapy for CLL.

Infections cause high morbidity and mortality in CLL. Early in the course of the illness, encapsulated bacteria, especially pneumococcus, may cause severe invasive disease because of hypogammaglobulinemia. With therapy, the incidence of viral infections, particularly herpesviruses, rises. Pulmonary infections are common, cause substantial mortality, and constitute a common indication for hospitalization. In the central nervous system, infection with the JC papovavirus may cause progressive multifocal leukoencephalopathy, but I will not consider this diagnosis further because of the rapid course, absence of typical imaging findings, and evidence of inflammation in the cerebrospinal fluid.

In the patient under consideration, the absence of a typical mass lesion and the predominance of T cells in the central nervous system both strongly suggest that the neurologic syndrome is not a direct complication of CLL.

INFECTIOUS CAUSES OF ENCEPHALOMYELITIS

The causes of meningoencephalitis are legion and may have tremendous overlap in clinical presentation, laboratory results, and imaging findings, which can make diagnosis extremely difficult.[1] In this patient, the presence of profound humoral immunodeficiency presents even greater diagnostic challenges, owing to the substantial likelihood of false negative serologic studies. Thus, the accurate etiologic diagnosis of encephalitis leans heavily on epidemiologic clues and unique clinical or radiologic findings.

The presence of a poliomyelitis-like syndrome with flaccid paralysis should focus attention on viruses, which are the only pathogens likely to cause this syndrome, with or without concomitant encephalitis. Herpesvirus infections must be briefly considered because of the frequency with which these viruses cause encephalitis and because of the common association with late-stage CLL. Herpes simplex virus 1 (HSV-1) is probably not the cause of this man's encephalitis because there were no temporal-lobe abnormalities on imaging studies, and the negative results of the polymerase-chain-reaction (PCR) assay for HSV-1 in the cerebrospinal flu-

id. Both of these indicators have excellent negative predictive value.[2,3] Varicella zoster may cause acute strokes or a chronic leukoencephalitis in immunocompromised patients,[4] but neither pattern fits this clinical picture. Cytomegalovirus disease of the central nervous system is largely confined to patients with human immunodeficiency virus infection. Human herpesvirus 6 and human herpesvirus 7 most commonly cause encephalitis in recipients of transplanted stem cells. Moreover, none of the above viruses is commonly associated with a poliomyelitis-like illness.

In the late summer and early fall, enteroviruses are a common cause of aseptic meningitis, and they can occasionally cause more severe neurologic disease, including encephalitis and flaccid paralysis. Enterovirus 71 has been associated with outbreaks of brain-stem encephalitis and flaccid paralysis in Asian children.[5] Hypogammaglobulinemia may predispose patients to severe meningoencephalitis associated with enteroviruses. This predisposition is most commonly seen in children with X-linked hypogammaglobulinemia but has also been documented in adults with common variable immunodeficiency.[6,7] PCR analysis of cerebrospinal fluid for enteroviruses has a sensitivity of more than 98 percent for detecting infection of the central nervous system.[8] Thus, the negative cerebrospinal fluid study in this patient provided important evidence against this diagnosis. However, since enteroviral infection was still a possibility on clinical grounds, administration of intravenous immune globulin was justified, since this has been successfully used as adjunctive therapy in cases in which there is underlying immunodeficiency.

Finally, arthropod-borne infection must be considered, given the season in which the patient presented, the greater likelihood of severe disease in elderly and immunocompromised patients, and the presence of several pathogens causing viral encephalitis in Massachusetts, including St. Louis virus, eastern equine encephalitis virus, Powassan virus, and West Nile virus. Of these, only West Nile virus has been reported to cause flaccid paralysis, either as an isolated neurologic manifestation or in conjunction with encephalitis. Lower motor-neuron weakness is reported in 20 to 60 percent of patients infected with West Nile virus that affects the central nervous system.[9,10] The thalamic abnormality revealed on MRI of this patient adds weight to this diagnosis, because an increasing number of flaviviruses have been associated with basal-ganglion involvement.[11] A parkinsonian syndrome developed in a minority of patients during or after recovery from West Nile encephalitis,[10,11] attesting to the clinical relevance of the imaging findings. A respiratory prodrome is atypical of West Nile virus infection, however, and its presence makes me wonder whether it was due to an unrelated infection in this case.

The best diagnostic test for West Nile encephalitis is generally a measurement of IgM antibody in the cerebrospinal fluid, which has a sensitivity of approximately 80 percent on the first sample.[12] PCR-based assays of cerebrospinal fluid are less sensitive.[13] The IgM measurement in the cerebrospinal fluid was negative in this patient, but I believe it was a false negative result caused by his hypogammaglobulinemia. There are no established therapies for severe West Nile infection, and supportive care is the cornerstone of management. Intravenous immune globulin has been given in some cases,[14] but its efficacy is not known.

Dr. Nancy Lee Harris: Dr. Singhal and Dr. Daskalakis, could you give us your impressions at the time that you cared for this patient?

Dr. Aneesh Singhal (Neurology): When neurology consultants first saw this patient, he was intubated and comatose, and it was difficult to determine whether his weakness resulted from a central or a peripheral nervous system process. The history of fever, cough, and a rapidly progressive encephalopathy, along with the results of the cerebrospinal fluid studies, raised the possibility of a viral meningoencephalitis. Cerebral complications of chronic leukemia were considered unlikely for the reasons outlined by Dr. Hollander. The MRI finding of diffuse hyperintensity in the left thalamus was consistent with a viral encephalitis and has been reported in patients infected with West Nile virus, eastern equine encephalitis virus, and Japanese B encephalitis virus, as well as in bartonella, and in several other types of encephalitis. Given this patient's geographic location, we considered West Nile virus and eastern equine encephalitis virus infections to be the leading differential diagnoses. We thought that the small hyperintense lesion seen on diffusion-weighted imaging was a region of intense inflammation within the surrounding T_2-hyperintense region of encephalitis, rather than an area of ischemia.

Dr. Demetre C. Daskalakis (Infectious Disease): This patient's initial presentation was similar to that of

many elderly patients with community-acquired infections, such as pneumonia or urinary tract infection. Though weakness and confusion may complicate such infections, the progressive confusion and weakness prompted a lumbar puncture that revealed a lymphocytic pleocytosis, suggesting meningoencephalitis or lymphomatous meningitis. Although the most common cause of viral meningoencephalitis is infection with herpes simplex virus, we considered this an unlikely cause for the reasons outlined by Dr. Hollander. We also considered partially treated bacterial meningitis, since both pneumococcus and meningococcus could have responded to azithromycin. Listeria infection of the central nervous system may cause a meningoencephalitis very similar to that from viral causes and may occur more frequently in patients with underlying hematologic cancer than in normal hosts. This patient's presentation in early autumn led us to consider arboviruses. The viral meningoencephalitis most associated with flaccid paralysis is West Nile virus. Central nervous system involvement by CLL or high-grade lymphoma was still being seriously considered by the clinical team caring for this patient at the time of his death.

CLINICAL DIAGNOSES

Viral meningoencephalitis due to enterovirus or West Nile virus infection.
Alternate diagnosis: Central nervous system involvement by CLL or high-grade lymphoma (Richter's syndrome).

DR. HARRY HOLLANDER'S DIAGNOSIS

Severe viral encephalomyelitis with flaccid paralysis, probably due to West Nile virus, with underlying hypogammaglobulinemia due to CLL.

PATHOLOGICAL DISCUSSION

Dr. E. Tessa Hedley-Whyte: The postmortem examination revealed extensive replacement of lymph nodes, bone marrow, and spleen by CLL with focal infiltrates in the lung, kidney, liver, heart, and prostate gland. The mass in the hypoglossal canal was an accumulation of lymphoid cells, consistent with an infiltrate of CLL, with no sign of a schwannoma.

The brain was externally and internally unremarkable on gross examination. Gross inspection of the spinal cord revealed congestion and retraction of the anterior horns, more marked in the cervical and lumbar areas and less so in the thoracic area. Microscopical examination of the brain and spinal cord revealed an extensive encephalomyelitis, with the most substantial damage in the anterior horns of the spinal cord and motor nuclei of the brain stem. Multiple areas of necrosis and inflammation were present, characterized by loss of neurons, accumulation of CD3+ lymphocytes and macrophages, and microglial nodules (Fig. 2A through 2D). The thalamus, cerebellum, and cerebral cortex contained scattered microglial nodules (Fig. 2E). Staining for the presence of infectious agents that might have caused encephalitis in an immunosuppressed patient, including toxoplasma, fungi, acid-fast bacilli, and Epstein–Barr virus, was negative.

Because the findings on admission included severe motor weakness, especially of the trunk, we thought that this condition might well be West Nile virus encephalomyelitis — despite the negative serologic assays. We sent unstained sections of the spinal cord, thalamus, and hippocampus to Dr. Juan Bilbao at the University of Toronto, who performed immunohistochemical staining for West Nile virus antigens.[15] Clusters of neurons in the hippocampus and the thalamus were positive for West Nile virus antigens (Fig. 2F). The spinal cord was negative except for an occasional neurite. These results established the diagnosis of West Nile virus encephalomyelitis. Cerebrospinal fluid obtained at postmortem examination was sent to the Massachusetts State Laboratory for testing for West Nile virus. The cerebrospinal fluid was positive for West Nile virus DNA by PCR, confirming the diagnosis of West Nile virus infection. The results of tests for antibodies to West Nile virus remained negative.

The first report of West Nile virus encephalomyelitis in 1999[16] noted that the infected patients were weak, but since examination of the spinal cord is often omitted as a part of the routine autopsy, initial pathology reports did not recognize involvement of the spinal cord as the explanation for the weakness[17]; this phenomenon was later recognized and reported.[18]

Twenty-three of 44 consecutive autopsy samples of the central nervous system submitted to the Centers for Disease Control and Prevention with a diagnosis of encephalitis between August and December 2002 tested positive for West Nile virus encephali-

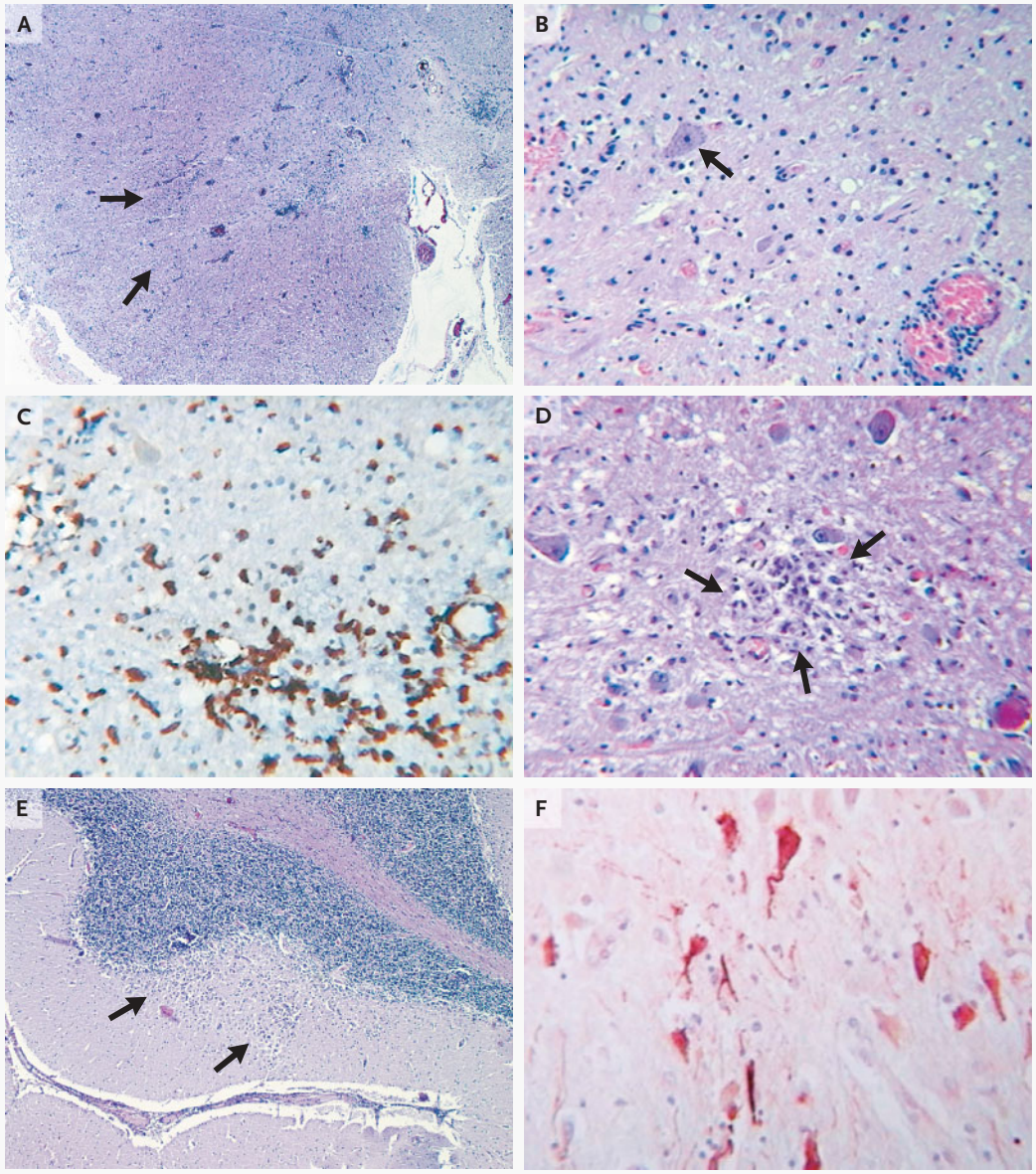


Figure 2. Microscopical Examination of the Brain and Spinal Cord at Autopsy.
A cross-section of the cervical spinal cord (Panel A) shows a marked inflammatory infiltrate that is most severe in the anterior horn (delineated by arrows). At higher magnification (Panel B), there is marked destruction of anterior-horn neurons with perivascular lymphocytic cuffing seen at lower right of the panel. (The arrow indicates a remaining neuron.) The infiltrating cells in Panel C are CD3+ T cells (immunoperoxidase stain for CD3). A section of 12th nerve nucleus (Panel D) shows a microglial nodule (arrows) composed of macrophages and degenerating neurons being phagocytized (neuronophagia). A section of cerebellum (Panel E) shows focal infiltration by inflammatory cells (arrows) (Panels A, B, D, and E, hematoxylin and eosin stain). With the presence of antibody to the West Nile virus antigen, there is staining of neurons in the hippocampus (Panel F). The brown-colored reaction product fills the entire cytoplasm and dendrites of the positive neurons. (Immunoperoxidase stain for West Nile virus, Panel F, performed by Juan Bilbao, M.D., University of Toronto, Toronto.)

tis.[19] As in this case, underlying medical conditions were present in more than 80 percent of the patients, and 70 percent of the patients were men more than 70 years of age. The histologic abnormalities, as in this case, were primarily in the brain stem and the anterior horns of the spinal cord. Serologic assays for West Nile virus were positive in 18 of 20 cases tested. West Nile virus antigens were detected by immunohistochemical analysis in only about half of the cases — more often in cases tested within the first week of illness than in those tested later.

ANATOMICAL DIAGNOSES

Encephalomyelitis due to West Nile virus.
CLL with involvement of the lymph nodes, bone marrow, spleen, lung, kidney, liver, heart, prostate, and hypoglossal canal.

REFERENCES

1. Glaser CA, Gilliam S, Schnurr D, et al. In search of encephalitis etiologies: diagnostic challenges in the California Encephalitis Project, 1998-2000. Clin Infect Dis 2003;36:731-42.
2. Cinque P, Bossolasco S, Lundkvist A. Molecular analysis of cerebrospinal fluid in viral diseases of the central nervous system. J Clin Virol 2003;26:1-28.
3. Domingues RB, Fink MC, Tsanaclis AM, et al. Diagnosis of herpes simplex encephalitis by magnetic resonance imaging and polymerase chain reaction assay of cerebrospinal fluid. J Neurol Sci 1998;157:148-53.
4. Kleinschmidt-DeMasters BK, Gilden DH. The expanding spectrum of herpesvirus infections of the nervous system. Brain Pathol 2001;11:440-51.
5. Huang C-C, Liu C-C, Chang Y-C, Chen C-Y, Wang S-T, Yeh T-F. Neurologic complications in children with enterovirus 71 infection. N Engl J Med 1999;341:936-42.
6. Misbah SA, Spickett GP, Ryba PC, et al. Chronic enteroviral meningoencephalitis in agammaglobulinemia: case report and literature review. J Clin Immunol 1992;12:266-70.
7. Rudge P, Webster AD, Revesz T, et al. Encephalomyelitis in primary hypogammaglobulinaemia. Brain 1996;119:1-15.
8. Carroll KC, Taggart B, Robinson J, Byington C, Hillyard D. Evaluation of the Roche AMPLICOR enterovirus PCR assay in the diagnosis of enteroviral central nervous system infections. J Clin Virol 2000;19:149-56.
9. Nash D, Mostashari F, Fine A, et al. The outbreak of West Nile virus infection in the New York City area in 1999. N Engl J Med 2001;344:1807-14.
10. Pepperell C, Rau N, Krajden S, et al. West Nile virus infection in 2002: morbidity and mortality among patients admitted to hospital in southcentral Ontario. CMAJ 2003;168:1399-405.
11. Solomon T, Fisher AF, Beasley DWC, et al. Natural and nosocomial infection in a patient with West Nile encephalitis and extrapyramidal movement disorders. Clin Infect Dis 2003;36:E140-E145.
12. Petersen LR, Roehrig JT, Hughes JM. West Nile virus encephalitis. N Engl J Med 2002;347:1225-6.
13. Lanciotti RS, Kerst AJ, Nasci RS, et al. Rapid detection of West Nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. J Clin Microbiol 2000;38:4066-71.
14. Haley M, Retter AS, Fowler D, Gea-Banacloche J, O'Grady NP. The role for intravenous immunoglobulin in the treatment of West Nile virus encephalitis. Clin Infect Dis 2003;37:e88-e90.
15. Bilbao JM, Chiasson D, Young B. West Nile virus encephalitis: pathology of seven cases. J Neuropathol Exp Neurol 2003;62:538.
16. Outbreak of West Nile-like viral encephalitis — New York, 1999. MMWR Morb Mortal Wkly Rep 1999;48:845-9.
17. Sampson BA, Ambrosi C, Charlot A, Reiber K, Veress JF, Armbrustmacher V. The pathology of human West Nile virus infection. Hum Pathol 2000;31:527-31.
18. Sampson BA, Nields H, Armbrustmacher V, Asnis DS. Muscle weakness in West Nile encephalitis is due to destruction of motor neurons. Hum Pathol 2003;34:628-9.
19. Guarner J, Shieh W-J, Hunter S, et al. Clinicopathologic study and laboratory diagnosis of 23 cases with West Nile virus encephalomyelitis. Hum Pathol 2004;35:983-90.

EDITORIALS

Benign Breast Disease — The Risks of Communicating Risk

Joann G. Elmore, M.D., M.P.H., and Gerd Gigerenzer, Ph.D.

The term "risk" appears in the title of more than 10,000 medical articles published in 2004 (2 percent of the total) — nine times as many as appeared in 1975. In this issue of the *Journal*, the article by Hartmann et al. on benign breast disease and the risk of breast cancer[1] continues the trend, as does this editorial. Hartmann et al. studied a cohort of women who had a benign breast lesion and found that the histologic appearance of the initial biopsy specimen was associated with the risk of breast cancer. As compared with women in the general population, women with nonproliferative findings on breast biopsy had a relative risk of breast cancer of 1.27, those with proliferative changes but no atypia had a relative risk of 1.88, and those with atypical hyperplasia had a relative risk of 4.24. The effect of atypia on the risk of cancer seemed to be independent of a family history of breast cancer.

These data solidify what has long been known about the risk of breast cancer among women with benign breast disease[2,3] and help stratify women with a benign lesion into high-risk and low-risk groups. The information will be useful for a surprising number of women: within a decade of starting annual screening, approximately 20 percent of women in the United States will have undergone a breast biopsy[4]; most of these biopsies show no evidence of cancer.

Hartmann et al. studied a large cohort of women and used current definitions to review all cases, but their results are limited by the retrospectively gathered information on family history (presented for only 53 percent of the women) and the lack of data on breast density and other risk factors. Menopausal status, moreover, was derived mainly from the women's ages. Other variables that underlie both the likelihood of a biopsy and the increase in the rate of detection of breast cancer, such as a high level of concern about breast cancer and consequent frequent examinations when atypia is noted, will require consideration in future studies.

It is unclear whether an atypical histologic appearance is a precursor lesion or a marker of a general tendency to develop breast cancer. Only half of invasive breast cancers arise in the same breast in which atypical hyperplasia was previously diagnosed, suggesting that this lesion is a marker of generalized risk.[1,5]

Additional refinement of risk may come with the identification of molecular markers; in the meantime, reproducibility of findings among pathologists must be improved if we plan to base risk estimates on histologic findings.[5,6] Hartmann et al. provide no data on reproducibility, despite prior studies that have shown major disagreements in the assessment of atypia.[5,6]

How should clinicians communicate the risk of breast cancer and the implications of a benign breast lesion to women? Most of us, who cannot interpret numbers nearly as well as words, have difficulty understanding numerical expressions of risk.[7] In medical schools, courses in statistics usually do not go far enough in teaching statistical or probabilistic thinking, and few teach strategies for effective communication. Hence, most physicians are poorly equipped to discuss risk factors in a way that is readily comprehensible to their patients. This deficiency puts the ideal of informed consent in jeopardy.

Three simple techniques can be helpful.[7-11] First, have numerical risk data on hand while seeing patients; second, communicate risk in a clear way; and third, pay attention to positive and negative framing. Consider a woman who asks about her breast-cancer risk and, like most women, has had no prior breast biopsy. She is white and 45 years

old, had her first menstrual period at the age of 12 and delivered her first child after the age of 30, and has no first-degree relative with breast cancer. According to the Gail risk model, easily obtained on a Web site,[12] her risk of a diagnosis of breast cancer within the next five years is 1.1 percent. Her risk of not receiving a diagnosis of breast cancer within the same period is 98.9 percent. A more transparent formulation is that among 1000 women with these characteristics, 11 would receive a diagnosis of breast cancer within the next five years, whereas 989 would not. The woman should understand that this is a risk of diagnosis, not death, and that treatment has markedly improved over time.

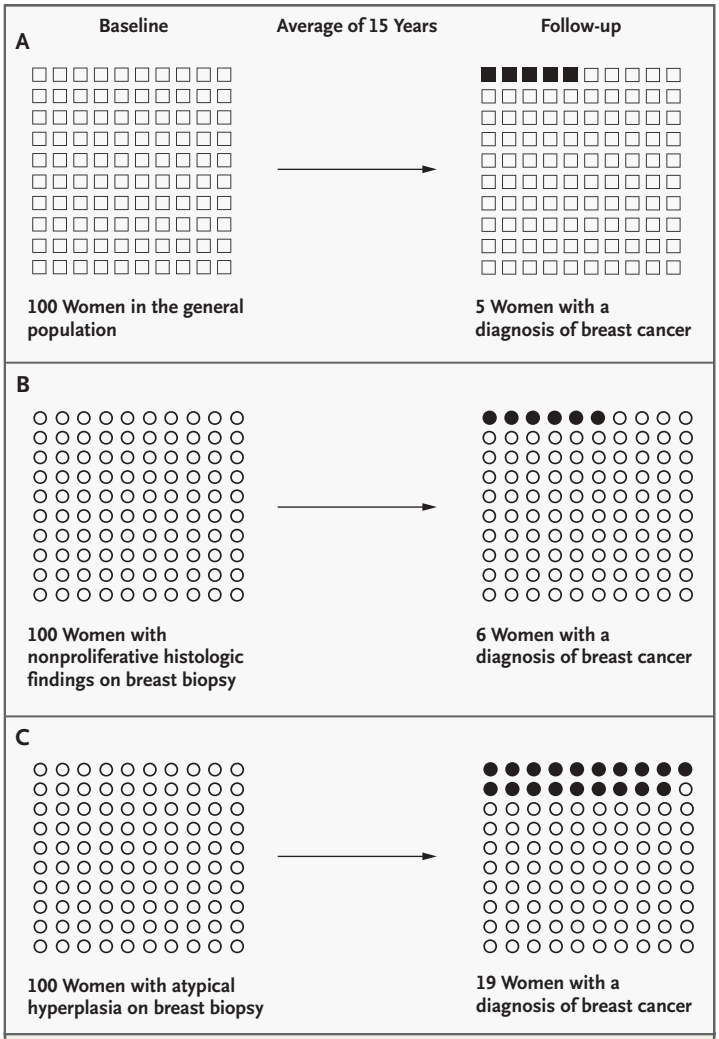


Figure 1. Examples of Outcomes among 100 Women Followed for an Average of 15 Years.

Each box represents a woman in the general population, and each circle represents a woman who has undergone a breast biopsy. The black circles and boxes represent women with a diagnosis of breast cancer after an average of 15 years of follow-up. In Panel A, 100 women in the general population are followed for an average of 15 years, and 5 subsequently receive a diagnosis of breast cancer. In Panel B, 100 women have nonproliferative findings on breast biopsy, and 6 subsequently receive a diagnosis of breast cancer. In Panel C, 100 women have atypical hyperplasia on breast biopsy, and 19 subsequently receive a diagnosis of breast cancer.

How can the results of the study by Hartmann et al. be explained to a woman with a benign breast lesion? They found that among 6061 women with nonproliferative disease, breast cancer developed in 379, as compared with an expected number of 297.7. This difference is reported as a relative risk of 1.27. The result can be communicated in terms of relative risks, which are misunderstood by many physicians and most patients, or absolute risks, which foster insight. Women with nonproliferative hyperplasia had a 27 percent increase in the risk of breast cancer in the ensuing 15 years. This is a relative risk and will most likely be misunderstood. Absolute risks are clearer (for simplicity, numbers are rounded): in the study by Hartmann et al., among 100 women in the general population, breast cancer developed in 5 within an average of 15 years of follow-up (Fig. 1A). Among 100 women with nonproliferative histologic findings, this number increased to approximately 6 (Fig. 1B). Thus, the increase in absolute risk is about 1 in 100. This is a simple way to describe the 27 percent increase in risk reported by Hartmann et al.

Women with proliferative disease but without atypia have an increase in the relative risk of breast cancer of 88 percent. Some women will falsely conclude that breast cancer will develop in 88 percent of such women. A more comprehensible way of communicating the same information is to say that among 100 women with this condition, the number in whom breast cancer will develop increases from 5 to about 10. Women with atypical hyperplasia have an increase in relative risk of 324 percent, equivalent to an increase in absolute risk from about 5 among 100 women in the general population to 19 among 100 women with atypical hyperplasia (compare Fig. 1A and Fig. 1C). The use of relative risks suggests greater effects than truly exist, whereas the use of absolute risks (or equivalent clear forms, such as the number needed to treat or the number needed to screen) prevents this misunderstanding. The use of relative risks should be avoided or employed in combination with more comprehensible forms of communicating risk.

Framing is the presentation of logically equiva-

lent information in different forms. Positive framing emphasizes the absence of disease; negative framing emphasizes the presence of disease. Expressing the absolute risk in a positive frame would lead us to say that among 100 women in the general population, breast cancer will not develop in 95 of them within the next 15 years (Fig. 1A); among 100 women with a biopsy revealing nonproliferative disease, 94 will not receive a diagnosis of breast cancer (Fig. 1B). People are sensitive to framing. Negative framing evokes a willingness to participate in a treatment or a screening, whereas positive framing may not.

Once information about risk is communicated, options for follow-up should be discussed (Table 1).[13,14] The recommended course of action is — and will remain for some time — annual mammographic screening with or without a clinical breast examination. If the woman wants to do more, she can perform breast self-examination, although this is no longer recommended by most expert groups. Annual screening with magnetic resonance imaging is not recommended for women whose only risk factor is benign breast disease. Genetic testing is recommended only for women with risk factors for *BRCA* mutations; it is unlikely to provide useful information for others. The risk of breast cancer among high-risk women can be decreased by chemoprevention and prophylactic surgery, though the potential harms need to be considered.

Informed decisions require that physicians know what the numbers mean and communicate them in ways that patients understand. Improving communication about risk is often treated as a "soft" topic, less important than improving forms of technology. But the best technology offers optimal results only when consumers understand its risks and benefits.[15]

Table 1. Options for Women at Increased Risk for Breast Cancer.

Option	Comment
Surveillance	Surveillance consists of annual mammography with or without clinical breast examination. Annual magnetic resonance imaging or ultrasonography is not recommended for women with benign breast disease.
Genetic testing	This approach is recommended only for women with risk factors for *BRCA* mutations and not for women whose only risk factor is atypia.
Chemoprevention*	Women at increased risk for breast cancer should be counseled about the potential benefits and harms of preventive therapy with a selective estrogen-receptor modulator. Increased risk has been defined as an age of more than 60 years, a 5-year risk of more than 1.66 percent as calculated with the use of the breast-cancer risk tool (available at http://www.cancer.gov/bcrisktool/), or a history of lobular carcinoma in situ. Chemoprevention is not recommended for women at low or average risk for breast cancer.
Prophylactic surgery	Mastectomy, bilateral salpingo-oophorectomy, or both may be an option for women at very high risk for breast cancer (e.g., those with genetic mutations).

* Recommendations for the chemoprevention of breast cancer have been provided by the U.S. Preventive Services Task Force.[13]

From the Department of Medicine, University of Washington School of Medicine, Seattle (J.G.E.); and the Max Planck Institute for Human Development, Berlin (G.G.).

1. Hartmann LC, Sellers TA, Frost MH, et al. Benign breast disease and the risk of breast cancer. N Engl J Med 2005;353:229-37.
2. Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. N Engl J Med 2000;342:564-71.
3. Fletcher SW, Elmore JG. Mammographic screening for breast cancer. N Engl J Med 2003;348:1672-80.
4. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.
5. Schnitt SJ. Benign breast disease and breast cancer risk: morphology and beyond. Am J Surg Pathol 2003;27:836-41.
6. Wells WA, Carney PA, Eliassen MS, Tostenson AN, Greenberg ER. A statewide study of diagnostic agreement in breast pathology. J Natl Cancer Inst 1998;90:142-5.
7. Gigerenzer G. Calculated risks: how to know when numbers deceive you. New York: Simon & Schuster, 2002.
8. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. BMJ 2002;324:827-30.
9. Paling J. Strategies to help patients understand risks. BMJ 2003;327:745-8.
10. Alaszewski A, Horlick-Jones T. How can doctors communicate information about risk more effectively? BMJ 2003;327:728-31.
11. Jekel J, Elmore J, Katz D. Epidemiology, biostatistics and preventive medicine. Philadelphia: Harcourt Health Sciences, 2001.
12. National Cancer Institute. Models for prediction of breast cancer risk: Gail model. (Accessed June 30, 2005, at http://www.cancer.gov/bcrisktool/.)
13. Preventive Services Task Force. Chemoprevention of breast cancer: recommendations and rationale. Ann Intern Med 2002;137:56-8.
14. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. JAMA 2005;293:1245-56.
15. Deyo RA, Patrick DL. Hope or hype: the obsession with medical advances and the high cost of false promises. New York: AMACOM, 2005:335.

Gastric Bypass and Nesidioblastosis — Too Much of a Good Thing for Islets?

David E. Cummings, M.D.

As the obesity pandemic continues to worsen and medical interventions remain only moderately effective, bariatric surgery is at present the only method that reliably results in major, long-term weight loss.[1,2] The most successful procedures durably reduce body weight by about one third and ameliorate nearly all obesity-related complications, most notably — almost miraculously — type 2 diabetes mellitus.[2] Not surprisingly, bariatrics is the fastest-growing surgical subspecialty. The number of bariatric surgeons registered with the American Society for Bariatric Surgery increased by almost 50 percent per year during the past several years, and the number of bariatric operations nearly doubles annually.[3] Still, questions remain with regard to the physiological mechanisms and pathophysiological consequences of bariatric operations, especially Roux-en-Y gastric bypass, the most common bariatric procedure performed in the United States. In this issue of the *Journal*, Service et al. describe a novel complication of Roux-en-Y gastric bypass: pathologic overgrowth of pancreatic beta cells (nesidioblastosis), resulting in life-threatening hyperinsulinemic hypoglycemia.[4] These findings identify a new adverse effect to consider in an increasing patient population and may also shed light on one mechanism by which diabetes resolves after bariatric surgery.

Service et al., experts in the field of hypoglycemic disorders, report that in the past five years, 40 percent of their confirmed cases of nesidioblastosis occurred in persons who had undergone Roux-en-Y gastric bypass surgery, whereas less than 0.1 percent of the general population has had this operation.[3] Their patients presented with repeated episodes of symptoms of profound postprandial neuroglycopenia associated with endogenous hyperinsulinemic hypoglycemia. Nesidioblastosis was definitively confirmed by selective arterial calcium-stimulation testing, alleviation of hypoglycemic symptoms after partial pancreatectomy, and demonstration of diffuse beta-cell hypertrophy and hyperplasia in resected pancreatic tissue. Insulinoma was ruled out with triple-phase spiral computed tomography and intraoperative ultrasonography in all but one patient, who had nesidioblastosis as well as multiple insulinomas. Although a causal link between Roux-en-Y gastric bypass and nesidioblastosis was not established, these observations strongly suggest that gastric bypass can occasionally result in pathologic beta-cell overgrowth and hypoglycemia. This possibility is supported by other case reports of nesidioblastosis[5] and insulinoma[6] after Roux-en-Y gastric bypass.

What mechanisms could cause nesidioblastosis after a procedure that diverts food from 95 percent of the stomach and a few feet of proximal small intestine? One theoretical possibility is that in obese persons with insulin resistance, adaptive beta-cell hypertrophy develops and causes hypoglycemia after insulin sensitivity is improved by the surgically induced weight loss. Service et al. argue against this as a mechanism by showing that the islets in obese control subjects without gastric bypass are of normal size.[4] Moreover, there is no association between non–surgically induced weight loss and endogenous hypoglycemia or nesidioblastosis.

Another plausible explanation, with broader implications, is that nesidioblastosis occasionally arises after Roux-en-Y gastric bypass because of long-term stimulation of beta-cell growth and activity by gut hormones that are perturbed as a result of the altered gastrointestinal transit. A prime candidate to mediate this effect is the incretin hormone glucagon-like peptide 1 (GLP-1). Produced by L cells primarily in the distal intestine, GLP-1 is secreted in response to the rapid passage of food from the stomach into the intestine and direct contact of the hindgut with nutrient chyme.[7] GLP-1 potently increases insulin secretion and, possibly, insulin sensitivity.[8] Moreover, at least in rodents, GLP-1 triggers beta-cell neogenesis and proliferation while inhibiting apoptosis.[9]

Theoretically, operations that expedite the delivery of nutrients to the hindgut should increase GLP-1 secretion, thereby enhancing insulin-mediated glucose disposal. Indeed, large (up to 10-fold) and durable (up to 20-year) elevations of GLP-1 or other nutrient-stimulated L-cell hormones, including peptide YY and enteroglucagon, have been documented after Roux-en-Y gastric bypass, biliopancreatic diversion, and jejunoileal bypass surgery.[10] These bariatric operations create shortcuts to the hindgut for ingested nutrients. They also

powerfully reverse diabetes, which completely resolves in well over 80 percent of cases.[1,2,11,12] Remarkably, diabetes typically remits within days to weeks after these operations — too early for the remission to be explained by weight loss alone, suggesting that other mechanisms, such as modulation of gut hormones, may play a role.

If increased secretion of GLP-1 and possibly other gastrointestinal hormones after gastric bypass provides continuing beta-cell stimulation, could this phenomenon occasionally result in nesidioblastosis? The observations of Service et al. support such a possibility. Accordingly, the authors caution against ascribing postprandial vasomotor symptoms to the dumping syndrome in patients who have undergone gastric bypass without considering organic hyperinsulinism.[4] However, since the Mayo Clinic, a hypoglycemia referral center, observed only six cases of apparent Roux-en-Y gastric bypass–induced nesidioblastosis in five years,[4] this complication hardly represents a public health crisis.

Perhaps more important, the findings of Service et al. hint at a possible risk resulting from long-term medicinal stimulation of GLP-1 signaling, the newest strategy used to treat type 2 diabetes.[13] The GLP-1–receptor agonist exenatide recently received approval by the Food and Drug Administration, and three related compounds — liraglutide, CJC-1131, and ZP10 — are in clinical trials. Agents are also in development that increase endogenous GLP-1 signaling by inhibiting dipeptidyl peptidase IV (DPP-IV), an enzyme that degrades GLP-1. Three such compounds, vildagliptin, sitagliptin, and saxagliptin, are in advanced clinical trials, and many others are in the pipeline. GLP-1–receptor agonists reduce levels of glycosylated hemoglobin as effectively as existing oral agents do while promoting weight loss or preventing weight gain. Moreover, because these compounds may increase beta-cell mass, they might slow or reverse the progressive islet-cell deterioration characteristic of diabetes. Since GLP-1–like agents acutely stimulate insulin secretion only in hyperglycemia, they are heralded as posing little independent risk of hypoglycemia. The findings of Service et al. would suggest that the last of these assertions may not always be true, if long-term overstimulation of GLP-1 signaling can cause pathologic beta-cell hypertrophy and hyperactivity, culminating in hypoglycemia.

Before one laments this potential drawback of medicines that increase GLP-1 signaling, however, some caveats should be considered. First, DPP-IV inhibitors, which hinder degradation of the endogenous hormone, do not stimulate GLP-1 signaling to the extent that high-dose GLP-1–receptor agonists do. Consequently, DPP-IV inhibitors probably decrease glucose somewhat less effectively, but they are also theoretically less prone to cause adverse effects such as nausea and, possibly, nesidioblastosis. Second, even doses of GLP-1–receptor agonists that substantially exceed peak endogenous GLP-1 activity improve diabetes far less than either Roux-en-Y gastric bypass or biliopancreatic diversion. Thus, these bariatric operations probably increase glucose disposal through mechanisms other than just increasing GLP-1. Because postsurgical antidiabetic effects occur before substantial weight loss, alterations in other gut hormones may play a role. For example, the orexigenic foregut hormone ghrelin can exert prodiabetic effects by suppressing insulin secretion, stimulating counter-regulatory hormones, and directly opposing insulin action.[14] Roux-en-Y gastric bypass usually impairs ghrelin secretion,[15] possibly increasing glucose tolerance. Other gut hormones that could theoretically mediate some of the effects of bariatric surgery include peptide YY, oxyntomodulin, and factors yet to be discovered. In short, bariatric surgery may increase glucose disposal through multiple mechanisms, including increased GLP-1 secretion. Thus, the risk of hypoglycemia after gastric bypass should exceed that of the use of medicines that selectively increase GLP-1 signaling.

Finally, should reports of Roux-en-Y gastric bypass–associated nesidioblastosis be considered worrisome or promising? Nesidioblastosis probably represents the pathologic extreme of a phenomenon that would benefit the vast majority of obese patients with diabetes. The possibility that Roux-en-Y gastric bypass stimulates beta-cell–trophic factors should spur research to identify these entities, GLP-1 or otherwise, so that their physiological effects can be harnessed and used pharmacologically to treat diabetes.

From the Department of Medicine, Division of Metabolism, Endocrinology, and Nutrition, University of Washington, and Veterans Affairs Puget Sound Health Care System, Seattle.

1. Sjöström L, Lindroos A-K, Peltonen M, et al. Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. N Engl J Med 2004;351:2683-93.
2. Buchwald H, Avidor Y, Braunwald E, et al. Bariatric surgery: a systematic review and meta-analysis. JAMA 2004;292:1724-37. [Erratum, JAMA 2005;293:1728.]
3. Steinbrook R. Surgery for severe obesity. N Engl J Med 2004; 350:1075-9.

4. Service GJ, Thompson GB, Service FJ, Andrews JC, Collazo-Clavell ML, Lloyd R. Hyperinsulinemic hypoglycemia with nesidioblastosis after gastric-bypass surgery. N Engl J Med 2005;353:249-54.
5. Patti ME, Bonner-Weir S, Goldsmith J, Mun EC, Hanto D, Goldfine AB. Severe hypoglycemia post-gastric bypass requiring parital pancreatectomy: evidence for inappropriate insulin secretion and pancreatic islet hyperplasia. In: Proceedings of the 65th American Diabetes Association Scientific Sessions, San Diego, Calif., June 10–14, 2005:A439. abstract.
6. Zagury L, Moreira RO, Guedes EP, Coutinho WF, Appolinario JC. Insulinoma misdiagnosed as dumping syndrome after bariatric surgery. Obes Surg 2004;14:120-3.
7. Dube PE, Brubaker PL. Nutrient, neural and endocrine control of glucagon-like peptide secretion. Horm Metab Res 2004;36:755-60.
8. D'Alessio D, Vahl TP. Glucagon-like peptide 1: evolution of an incretin into a treatment for diabetes. Am J Physiol Endocrinol Metab 2004;286:E882-E890.
9. Brubaker PL, Drucker DJ. Glucagon-like polypeptides regulate cell proliferation and apoptosis in the pancreas, gut, and central nervous system. Endocrinology 2004;145:2653-9.
10. Cummings DE, Overduin J, Foster-Schubert KE. Gastric bypass for obesity: mechanisms of weight loss and diabetes resolution. J Clin Endocrinol Metab 2004;89:2608-15.
11. Schauer PR, Burguera B, Ikramuddin S, et al. Effect of laparoscopic Roux-en Y gastric bypass on type 2 diabetes mellitus. Ann Surg 2003;238:467-85.
12. Pories WJ, Swanson MS, MacDonald KG, et al. Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus. Ann Surg 1995;222:339-52.
13. Holst JJ, Deacon CF. Glucagon-like peptide 1 and inhibitors of dipeptidyl peptidase IV in the treatment of type 2 diabetes mellitus. Curr Opin Pharmacol 2004;4:589-96.
14. Cummings DE, Foster-Schubert KE, Overduin J. Ghrelin and energy balance: focus on current controversies. Curr Drug Targets 2005;6:153-69.
15. Cummings DE, Weigle DS, Frayo RS, et al. Plasma ghrelin levels after diet-induced weight loss or gastric bypass surgery. N Engl J Med 2002;346:1623-30.

Improving the Quality of Hospital Care in America

Patrick S. Romano, M.D., M.P.H.

Over the past few years, several large studies have shown,[1-4] and the Institute of Medicine has emphasized,[5] that the quality of health care in the United States is not nearly at the level that we should expect from the world's most expensive health care system. Problems with quality are pervasive throughout both outpatient and inpatient settings and may be responsible for thousands of deaths each year.[3,4]

In this issue of the *Journal,* Williams and colleagues,[6] from the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), present time-series data from their "core measures" program, in which hospitals seeking accreditation must collect and submit data on clinical performance according to standardized, evidence-based measure sets. They bring us good news. During the period from late 2002 through early 2004, continuously participating hospitals improved significantly on 15 of the 18 measures for acute myocardial infarction, heart failure, and pneumonia. These hospitals improved significantly on every measure that involved appropriate use of medications or provision of counseling, and the trend was favorable for all three of the measures that did not show statistically significant improvement (i.e., mean time to thrombolysis and inpatient mortality after acute myocardial infarction and blood cultures for pneumonia). Furthermore, hospitals with the poorest performance at baseline improved the most over this two-year period. For five measures, involving the use of aspirin and beta-blockers for acute myocardial infarction and oxygenation assessment for pneumonia, U.S. hospitals are approaching optimal performance.

Williams and colleagues' findings are reassuring but not surprising. In 2003, Jencks and colleagues, from the Centers for Medicare and Medicaid Services (CMS), reported national changes in hospital performance for Medicare beneficiaries on 22 measures of quality,[7] of which 12 were substantively similar to those reported in this issue of the *Journal.* According to data collected by Medicare Quality Improvement Organizations through contracted clinical data abstraction centers, the median state's performance improved on 20 of the 22 measures, by an average of 12 percent, from 1998 and 1999 to 2000 and 2001. For all but one measure, states with poor performance at baseline improved more (in absolute terms) than states with good performance at baseline. Although the specifications of the 12 similar measures have changed slightly, Williams and colleagues present convincing evidence that the improvements documented among Medicare beneficiaries before 2001 have continued through 2004 and have benefited non-Medicare patients as well. Similar improvements have been reported from the Department of Veterans Affairs health care system.[8]

Also in this issue of the *Journal,* Jha and colleagues[9] describe variations in clinical performance during the first half of 2004 among hospitals participating in the CMS Hospital Quality Alliance program, which began as a voluntary public reporting

program with limited participation but grew substantially after the implementation of financial incentives for participation. Although Jha and colleagues report on several of the same measures as Williams and colleagues, the JCAHO and CMS specifications differed slightly until this year, making it difficult to compare results side by side. Nonetheless, the work of Jha and colleagues complements that of Williams and colleagues by identifying settings in which quality remains below average. Specifically, midwestern and northeastern hospitals perform better, on average, than southern and western hospitals in all three clinical domains (i.e., acute myocardial infarction, congestive heart failure, and pneumonia). Northern and central communities (e.g., Boston; Indianapolis; Kansas City, Mo.; and Camden, N.J.) dominate the list of "top-ranked performers," whereas southeastern and southwestern communities (e.g., Little Rock, Ark.; Orlando, Fla.; Miami; San Bernardino, Calif.; and San Diego, Calif.) dominate the list of "bottom-ranked performers." Nonprofit hospitals slightly but consistently outperform for-profit hospitals; teaching hospitals outperform nonteaching hospitals with respect to cardiac care but not pneumonia care.

What are the implications of these two studies for the future of improvements in quality in the U.S. hospital industry? First, they establish that very high levels of adherence with evidence-based guidelines are achievable, with sufficient education of physicians and hospital managers and sufficient attention from outside entities such as JCAHO and CMS. Of course, these high levels of adherence may be somewhat illusory, since physicians and hospitals have become increasingly clever at documenting questionable contraindications to standard therapies, thereby excluding many patients who might have benefited from them.

Second, these studies confirm that variation in hospital performance, at least for three medical conditions, has generally shrunk in recent years. Poorly performing hospitals improved more than highly performing hospitals during the period from 2002 to 2004, and regional disparities appear to be smaller now than they were in 1998 and 1999.[10] Of course, variation in quality persists, and the data reported by Williams et al. and Jha et al. do not address more pernicious disparities related to insurance status and other sociodemographic characteristics.

Third, these studies confirm the long-recognized fact that repeatedly observed behaviors improve over time — a phenomenon known as the "Hawthorne effect," after the location of the facility in which it was described.[11] CMS and JCAHO chose to focus on acute myocardial infarction, congestive heart failure, and pneumonia because these conditions frequently lead to hospitalization and death and because medical therapy for these conditions is clearly effective. However, these conditions do not account for the majority of hospitalizations in the United States. We have no idea whether care for other conditions has deteriorated, even while care for acute myocardial infarction, congestive heart failure, and pneumonia has improved. It is even possible that unmeasured aspects of care for these three conditions have deteriorated. Williams and colleagues' inability to find any decrease in inpatient mortality related to acute myocardial infarction, neonatal mortality, or obstetric lacerations (unpublished data) suggest that the effect of the reported improvements on public health may be modest.

Despite the progress that we have made in hospital care for acute myocardial infarction, congestive heart failure, and pneumonia, we still face major challenges. Medical knowledge continues to expand every year, so practice guidelines and the measures of quality on which they are based require continual updating. Quality measures that are not updated to reflect current research findings will lose their value, and sponsors will lose their credibility.

As hospitals and physicians become more sophisticated in "gaming" quality measures, sponsors must also become more sophisticated in monitoring accuracy. For example, the largest absolute improvements between 2002 and 2004 were documented for smoking-cessation counseling and discharge instructions, which are measures that hospitals can manipulate through check-off forms that nurses complete when they discharge patients with the target conditions. Although any patient documented as receiving aspirin almost certainly received it, we have no such confidence with regard to smoking-cessation counseling. Educational interventions that are effective in a clinical trial may fail abysmally when they are transformed into a check box on a discharge form. Patient and family surveys may be helpful to monitor the delivery of such interventions.

Finally, we must not rest on our laurels and assume that we have solved the problem of quality by improving 15 measures for three conditions in about 1400 to 2000 acute care hospitals. Jha and

colleagues show that performance remains mediocre at hospitals that do not meet the sample-size requirement for public reporting.[9] Brennan and colleagues showed that there is substantial room for improving safety among surgical patients.[3] McGlynn and colleagues showed that patients receive only 50 to 60 percent of indicated interventions across multiple domains of predominantly outpatient care.[1] Indeed, performance improved by a median of 5.4 percent on 24 measures of hospital quality between the 2003 and 2004 editions of the National Healthcare Quality Report but by only 1.4 percent on 49 measures of the quality of ambulatory care.[12] We have barely begun to touch quality-related problems in mental health and substance-abuse care, pediatric care, and home health care. As physicians and health professionals, we have made a little progress, but we still have far to go in closing the "quality chasm" that the Institute of Medicine recognized in 2001.[5]

From the Division of General Medicine and the Center for Health Services Research in Primary Care, University of California–Davis, Sacramento.

1. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.

2. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. 1. The content, quality, and accessibility of care. Ann Intern Med 2003;138:273-87.

3. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. N Engl J Med 1991;324:370-6.

4. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care 2000;38:261-71.

5. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C.: National Academies Press, 2001.

6. Williams SC, Schmaltz SP, Morton DJ, Koss RG, Loeb JM. Quality of care in U.S. hospitals as reflected by standardized measures, 2002–2004. N Engl J Med 2005;353:255-64.

7. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001. JAMA 2003;289:305-12. [Erratum, JAMA 2003;289:2649.]

8. Jha AK, Perlin JB, Kizer KW, Dudley RA. Effect of the transformation of the Veterans Affairs health care system on the quality of care. N Engl J Med 2003;348:2218-27.

9. Jha AK, Li Z, Orav EJ, Epstein AM. Care in U.S. hospitals — the Hospital Quality Alliance program. N Engl J Med 2005;353:265-74.

10. Jencks SF, Cuerdon T, Burwen DR, et al. Quality of medical care delivered to Medicare beneficiaries: a profile at state and national levels. JAMA 2000;284:1670-6.

11. Roethlisberger FJ, Dickson WJ. Management and the worker: technical vs. social organization in an industrial plant. Cambridge, Mass.: Harvard University Press, 1934.

12. 2004 National Healthcare Quality Report (NHQR). Rockville, Md.: Agency for Healthcare Research and Quality, 2004. (AHRQ publication no. 05-0013.)

SOUNDING BOARD

Medicare and Chronic Conditions

Gerard F. Anderson, Ph.D.

When the Medicare program became operational in 1966, its primary orientation was the treatment of acute, episodic illness.[1,2] The design of the program's benefits, coverage policies, payments to providers, and criteria for determining medical necessity were all oriented toward the treatment of acute diseases. Medicare retained this orientation for the next 40 years in spite of the growing number of Americans with chronic conditions.[3,4] The Medicare Prescription Drug Improvement and Modernization Act of 2003 was an important first step in the reorientation of the Medicare program toward the care of patients with chronic disorders. Additional changes, however, will be necessary if the Medicare program is to be truly responsible to its millions of beneficiaries who have chronic conditions, especially those with multiple coexisting illnesses.

BENEFICIARIES WITH FIVE OR MORE CHRONIC CONDITIONS

A total of 83 percent of Medicare beneficiaries have at least one chronic condition. As additional diseases are diagnosed, expenditures and the probability of an adverse outcome increase rapidly. Any policymaker who is considering the modernization of Medicare must recognize that the 23 percent of beneficiaries with five or more chronic conditions account for 68 percent of the program's spending. In addition, the treatment of these beneficiaries is likely to remain a high-cost item until they die, since every year they see an average of 13 physicians and fill an average of 50 prescriptions.[5] They are also the beneficiaries who are most likely to have a preventable hospitalization and have the highest out-of-pocket spending because of gaps in coverage and cost-sharing arrangements.

BEGINNING MODERNIZATION OF MEDICARE

The part of the legislation to modernize Medicare that has received the most attention is the prescription-drug benefit.[6] Coverage of prescription drugs can be viewed as part of a larger initiative to make the Medicare program more responsive to the needs of beneficiaries with chronic conditions (Table 1).

Section 721 created the Chronic Care Improvement Program, which represents an important new initiative to improve the quality of care for beneficiaries with chronic conditions in the Medicare fee-for-service program.[7] It is not a demonstration program but a newly covered service. Initially, a pilot program will offer self-care guidance and support to Medicare beneficiaries who have one or more of three chronic conditions: complex diabetes, congestive heart failure, and chronic obstructive pulmonary disease (COPD). These three diseases were chosen by Congress for multiple reasons, including their high prevalence in the Medicare population and the likelihood that beneficiaries with any one of these chronic conditions has one or more coexisting illnesses. An analysis of Medicare claims data for 2001, for example, shows that 96 percent of beneficiaries with COPD have at least one other coexisting illness, and 68 percent have four or more coexisting illnesses. The objective of Section 721 is to increase adherence to evidence-based care, reduce unnecessary hospital stays and emergency room visits, and help beneficiaries avoid costly and debilitating complications.

The program will be implemented in two phases. A pilot phase will help determine the final design. On December 8, 2004, pilot programs in Maryland, Pennsylvania, Oklahoma, Mississippi, Tennessee, Georgia, the District of Columbia, Florida, Chicago, and Brooklyn and Queens, New York, were selected.[7] These regional programs will be responsible for providing appropriate services to all Medicare beneficiaries who have complex diabetes, congestive heart failure, or COPD. Most of the organizations selected to oversee these programs are disease-management organizations. Payments to the pilot programs will be dependent on improvement in the quality of clinical care, the satisfaction of beneficiaries and providers, and a demonstration

Table 1. Selected Provisions of the Medicare Modernization Act That Address the Needs of Patients with Chronic Conditions.

Section of Act	Title	Focus	Purpose	Current Status
101	National Standards for Electronic Prescriptions	Physicians	To work with clinicians and industry experts to develop national standards	Proposed rule issued February 4, 2005
108	Electronic Prescription Grants	Physicians	To provide grants to implement electronic prescription-drug programs	Under way
231	Specialized Plans for Patients with Special Needs	Managed care	To provide incentives for managed-care plans to enroll patients with complex chronic conditions	Planning stage
721	Chronic Care Improvement	Fee for service	To improve adherence to evidence-based medicine and reduce unnecessary use of care	Funding for pilot programs awarded, mostly to disease-management companies
721*	Care Management for High-Cost Beneficiaries	Fee for service	To involve clinicians in care management	Awards in 2005
723	Strategy	Medicare program	To develop a long-term plan to improve the quality and reduce the cost of care for beneficiaries with chronic conditions	Under way

* Medicare created this program as a companion to Chronic Care Improvement.

of success in lowering costs — all with the use of comparisons with control groups. Phase 2, which is scheduled to begin after 2006, may expand to other geographic regions (or perhaps nationally) programs or program components that have proved to be successful.

GETTING PHYSICIANS INVOLVED

The Medicare program has developed its own companion initiative to Section 721. The focus of the companion initiative is high-cost beneficiaries with chronic conditions who do not have complex diabetes, congestive heart failure, or COPD. Unlike the Chronic Care Improvement Program, which awarded the funding primarily to disease-management organizations, the Care Management for High-Cost Beneficiaries demonstration is targeted primarily at physician groups, hospitals, and integrated delivery systems. One possible reason for this targeting is that Medicare wants to get the clinicians and delivery systems more directly involved in care management, especially for beneficiaries with multiple coexisting illnesses. One congressional study has reported that disease-management programs might not be cost-effective for beneficiaries with multiple coexisting illnesses.[8] The demonstration will require that applicants specify performance standards to improve clinical quality, measure the satisfaction of beneficiaries and providers, and achieve financial savings. Program funding should be awarded later this year.

MANAGED CARE

Section 231 will encourage managed-care organizations to offer specialized plans that serve beneficiaries who have special health care needs. It has been a long-standing concern that managed-care organizations do not have a financial incentive to enroll beneficiaries with multiple serious chronic conditions.[9] Section 231 attempts to address this concern. Beneficiaries who are eligible for these specialized plans will be persons who live in institutions or who qualify for both Medicare and Medicaid; other persons who have chronic conditions or disabilities may be included. On November 8, 2004, the Centers for Medicare and Medicaid Services held a meeting to discuss issues involved with policy and operations. The specifics of this program are also scheduled to be announced later this year.

ELECTRONIC PRESCRIPTIONS

Sections 101 and 108 begin the process that could lead to the integrated electronic medical record. Section 101 requires that the Medicare program work with industry experts to establish national standards for electronic prescriptions, and Section 108 will award grants to physicians to implement

electronic-prescription programs. The legislation envisions a Medicare program in which a doctor can write a prescription on a computer and electronically transmit that prescription to a pharmacy. This is the first step toward a broader objective of creating integrated electronic medical records with shared data repositories.

The Medicare Modernization Act contains numerous other provisions that set the stage for additional transformations in the program. For example, Section 723 mandates that the secretary of the Department of Health and Human Services "develop a plan to improve quality of care and reduce the cost of care for chronically ill Medicare beneficiaries." One of the targets of this report will be beneficiaries with multiple chronic conditions.

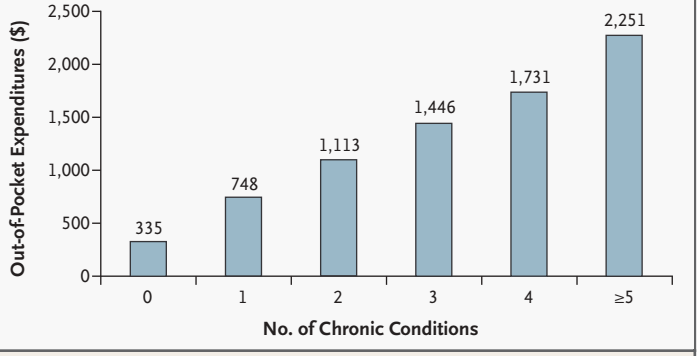


Figure 1. Annual Out-of-Pocket Spending by Medicare Beneficiaries.
Data are from the Medical Expenditure Panel Survey of 2001.[11]

NEXT STEPS

Although the Medicare Modernization Act is an important first step toward reform, additional steps will be needed before the Medicare program is truly oriented toward the treatment of beneficiaries with multiple chronic conditions. The Medicare program cannot do this alone, however. It will also be necessary to change the delivery system, the research infrastructure, clinical education, and methods of financing medical care in order for the health care system to become more responsive to the needs of people with chronic conditions.[10]

One step is to restructure the cost-sharing arrangements in fee-for-service Medicare. Out-of-pocket spending by Medicare beneficiaries increases by an average of nearly $400 with each additional chronic condition (Fig. 1).[12] The current cost-sharing arrangements, such as the 20 percent coinsurance for physician visits or gaps in the prescription-drug benefit, are especially onerous to beneficiaries with multiple chronic conditions because these people are the highest users of medical services.[13] One possible solution is an out-of-pocket maximum. Most private insurers place a limit on the patient's out-of-pocket expenses, and Medicare could adopt a similar approach. Accomplishing this objective while still maintaining budget neutrality could require greater cost sharing by Medicare beneficiaries who have few or no chronic conditions. Alternatively, additional funding could be sought.

A second step is for Medicare to make an additional payment when a standardized electronic medical record is sent to a secure data repository. This would be an expansion of Sections 101 and 108 in the Medicare law and would allow for the creation of integrated electronic medical records, which would be especially helpful for beneficiaries with multiple chronic conditions. The Department of Veterans Affairs already has operational electronic medical records, and countries such as Canada and the United Kingdom are investing billions of dollars to create such systems.

One potential problem is the cost to the Medicare program. Given the tremendous volume of health care visits by Medicare beneficiaries, if Medicare were to pay $5 to a physician, hospital, or other provider to send an electronic medical record to the secure data repository, the cost to the Medicare program would exceed $4 billion annually. However, the Medicare program might be able to reduce costs and improve quality if widespread use of electronic medical records reduced the number of duplicate tests, adverse drug reactions, and unnecessary hospitalizations. To be successful, this program would require the active participation of clinicians — an area in which acceptance so far has been relatively slow.

For Medicare beneficiaries with five or more chronic conditions, who see an average of nine physicians on an outpatient basis and four hospital-based physicians annually, coordination of care is especially important. Both physicians and patients are aware of the problems that can occur when care is not coordinated.[14] A third step in reforming the Medicare program might be to require that the program explicitly pay for care coordination. Under one proposal, each beneficiary with five or more chronic conditions would designate a care coordi-

nator who would be required to communicate with all other clinicians on a periodic basis and help coordinate services.[15] More research will be necessary in order to identify the precise characteristics of beneficiaries who will benefit from care coordination and the specific interventions that will be successful. A major stumbling block could be the minimal training in care coordination that most physicians currently receive.[16] It may also be necessary to restructure the way in which Medicare pays for graduate medical education in order to emphasize training in care coordination in ambulatory settings.

MEDICARE PAYMENT RULES

Important changes in Medicare's payment systems will be needed to pay for some of the proposed improvements in care for chronic conditions.[15,17] Fee-for-service payments will need to be restructured to encourage clinicians to work cooperatively; to encourage additional means of communication, such as e-mail; and to permit doctors to see a group of patients at once and allow other providers to participate in, and be reimbursed for, the care of patients.[15] Current Medicare rules make each of these improvements problematic. One problem is that the cost of processing claims for things such as e-mail communication could be greater than the amount Medicare would pay for the encounter. For some services, it could be difficult to limit the number of encounters between physicians and patients to a medically appropriate number. Patients could send five or more e-mail messages a day to a physician and expect a response if the physician were being reimbursed by Medicare. It is also difficult for the Medicare program to verify that an e-mail communication has occurred.

Current Medicare regulations are very specific about which providers are eligible to be paid and under which circumstances.[18] Nonphysicians are generally not eligible to be paid by the Medicare program unless the service is "incident to" a physician's service, and even then, payment is possible only under certain circumstances. Existing rules preclude payment for services that are commonly furnished in a physician's office or rendered without charge. As a result, explicit payment for patient education, some group visits, and multidisciplinary group conferences will be difficult under existing Medicare rules unless Congress explicitly authorizes payment (e.g., for education about diabetes, as it currently does).

Payments to managed-care plans will need to cover the full expected cost of care for beneficiaries with multiple chronic conditions — something that the current system does not do. Beginning in 2007, Medicare will pay managed-care plans on the basis of a system that is 100 percent risk-adjusted for the types of patients the managed-care plan enrolls. In theory, this risk-adjusted payment would reflect the additional costs of treating a beneficiary with multiple chronic conditions. In reality, the payment will still underestimate the cost of treating a beneficiary who requires expensive care or multiple hospitalizations.[18]

There are several problems to overcome before Medicare can implement any of these recommendations in the next round of program reforms. Some of these proposals are likely to increase the costs of Medicare, at least in the short run. However, spending could be cut by reducing the number of hospitalizations, drug interactions, and duplicate tests. Any savings would need to be demonstrated. The second problem is the potential for fraud and abuse. The concern, as discussed earlier, is how to determine whether services are actually being provided, especially for activities such as e-mail communication. The third problem is how to demonstrate improvement in health outcomes. Both physicians and beneficiaries will need to be convinced that the reforms result in better clinical outcomes. The fourth problem is the unwillingness of some clinicians to participate in the reforms. In some ways, the fourth consideration may be the most important obstacle. Costs can be lowered, fraud and abuse minimized, and outcomes improved only if a high percentage of clinicians perceive that Medicare's new orientation is improving outcomes.

Because of the recent legislation, it can now be said that Medicare is becoming a program for people with chronic conditions. However, we have just begun the journey.

Supported by the Robert Wood Johnson Foundation.

From the Center for Hospital Finance and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore.

1. Corrigan JM, Eden J, Smith BM, eds. Leadership by example: coordinating government roles in improving healthcare quality. Washington, D.C.: National Academies Press, 2003.
2. Williams RD II. The unique needs of Medicare beneficiaries. Brief no. 10. Washington, D.C.: National Academy of Social Insurance, October 2004.
3. Thorpe KE, Florence CS, Joski P. Which medical conditions ac-

count for the rise in health care spending? Bethesda, Md.: Health Affairs, August 25, 2004 (Web exclusive). (Accessed June 30, 2005, at http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.437/DC1.)
4. Wu S, Green A. Projection of chronic illness prevalence and cost inflation. Santa Monica, Calif.: RAND Health, October 2000.
5. Johns Hopkins University. Partnership for solutions: better lives for people with chronic conditions. (Accessed, June 30, 2005, at http://www.partnershipforsolutions.com.)
6. Frank RG. Election 2004: prescription-drug prices. N Engl J Med 2004;351:1375-7.
7. Medicare health support. Baltimore: Centers for Medicare & Medicaid Services, 2005. (Accessed on June 30, 2005, at http://cms.hhs.gov/medicarereform/ccip.)
8. An analysis of the literature on disease management programs: letter to the honorable Don Nickles. Washington, D.C.: Congressional Budget Office, October 13, 2004.
9. Kuttner R. The risk-adjustment debate. N Engl J Med 1998;339:1952-6.
10. Anderson GF, Knickman JR. Changing the chronic care system to meet people's needs. Health Aff (Millwood) 2001;20(6):146-60.
11. 2001 Medical Expenditure Panel Survey. Rockville, Md.: Agency for Healthcare Research and Quality, 2001.
12. Hwang W, Weller W, Ireys H, Anderson GF. Out-of-pocket medical spending for care of chronic conditions. Health Aff (Millwood) 2001;20(6):267-78.
13. Anderson GF, Shea DG, Hussey PS, Keyhani S, Zephyrin L. Doughnut holes and price controls. Bethesda, Md.: Health Affairs, July 21, 2004 (Web exclusive). (Accessed June 30, 2005, at http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.396/DC1.)
14. Anderson GF. Physician, public, and policymaker perspectives on chronic conditions. Arch Intern Med 2003;163:437-42.
15. Berenson RA, Horvath J. Confronting the barriers to chronic care management in Medicare. Bethesda, Md.: Health Affairs, January 22, 2003 (Web exclusive). (Accessed June 30, 2005, at http://content/healthaffairs.org/cgi/content/full/hlthaff.w3.37v1/DC1.)
16. Darer JD, Hwang W, Pham HH, Bass EB, Anderson GF. More training needed in chronic care: a survey of US physicians. Acad Med 2004;79:541-8.
17. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74:511-44.
18. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financ Rev 2004;25(4):119-41.

CLINICAL IMPLICATIONS OF BASIC RESEARCH

A New Strategy to Counter Allergy

Jean-Pierre Kinet, M.D.

Many features of cellular programs are tightly controlled by the balancing actions of activation and inhibitory signals. This principle applies in particular to the cells of the immune system, which are regulated by activating and inhibitory cell-surface receptors. Zhu and colleagues[1] have recently described a clever strategy to repress specific allergic reactions by muting the activity of activation receptors.

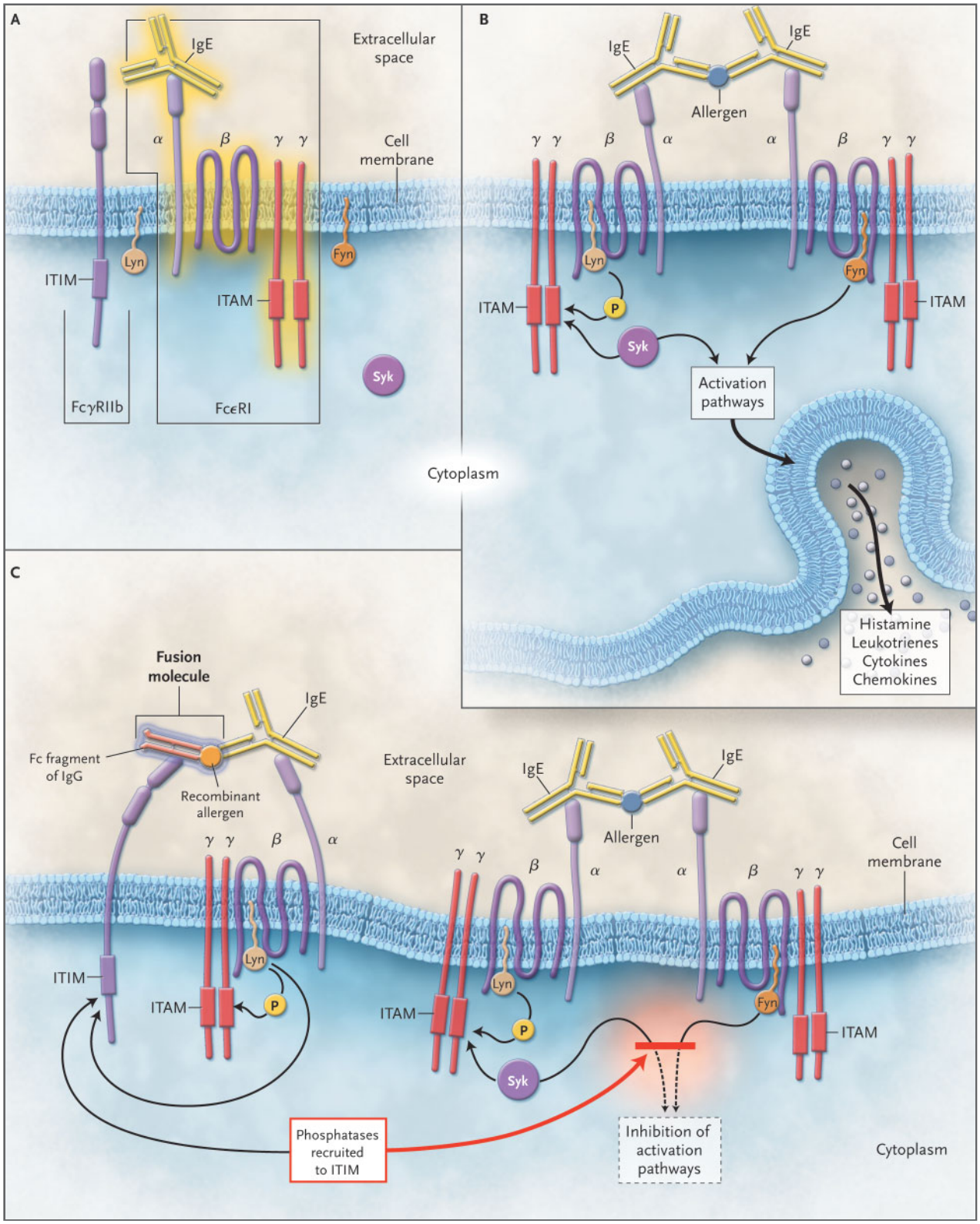
The first step of the allergic response is the immunization phase, in which innocuous allergens trigger B cells to produce IgE and IgG that react specifically to these allergens. These IgE molecules then bind to high-affinity receptors for IgE (FcεRI) on circulating basophils and mast cells in many tissues (Fig. 1). The binding of IgE to FcεRI does not itself induce allergic reactions. A second step is required. Allergens are multivalent: they express multiple epitopes that are recognized by specific IgEs and IgGs. Simultaneous multivalent binding of allergens to several membrane-bound IgEs induces receptor aggregation, which triggers a signaling cascade that leads to the production and release of allergic and inflammatory mediators (histamine, leukotrienes, chemokines, and cytokines) responsible for the symptoms of allergic diseases. This chain of events is central to many allergic diseases, including allergic rhinitis, allergic asthma, urticaria, and systemic anaphylaxis.

How does the aggregation of FcεRI generate this signaling cascade? FcεRI is composed of an IgE-binding unit (the α chain) and a signaling unit (one β and two γ chains) (Fig. 1). The aggregation of FcεRI induces the activation of a tyrosine kinase bound to the β chain; the tyrosine kinase then phosphorylates two tyrosine residues in the γ chains. These tyrosine residues are a central feature of the immunoreceptor tyrosine-based activation motif (ITAM), the canonical activation motif present on numerous receptors in the immune system. After phosphorylation, the ITAMs of the γ chains activate Syk tyrosine kinase, which, through the activation of downstream pathways, induces the release of allergic mediators (Fig. 1).[2,3]

These FcεRI-mediated activation pathways are modulated by inhibitory receptors — the IgG receptor FcγRIIb in particular. The allergen-specific IgGs produced in response to immunization form complexes with allergens, which can, in turn, form a bridge between FcεRI and FcγRIIb. Both receptors are expressed on mast cells and basophils; the Fc fragment of IgG in the immune complex binds to FcγRIIb, whereas the allergen binds to IgE, which is already bound to adjacent FcεRI. The formation of this bridge induces the aggregation of activating FcεRI with inhibitory FcγRIIb, which inhibits the activation pathways activated by FcεRI. This capacity of IgG to counteract IgE is probably the central mechanism behind successful allergen desensitization, although this has yet to be established. However, desensitization does not always work, as illustrated by dramatic failures involving life-threatening anaphylactic shock.

How does FcγRIIb-mediated inhibition work?

Figure 1 (facing page). A New Strategy to Counter Allergy.
Allergens induce the aggregation of membrane-bound IgEs that react specifically to them. The IgEs are bound to the membrane of basophils and mast cells by IgE receptors called FcεRIs (Panel A). Thus, the aggregation of IgE induces the aggregation of FcεRI and, hence, the triggering of activation signals that eventually lead to the release of the mediators (such as histamine) of allergic reactions (Panel B). In persons without allergies, membrane-bound IgG also forms immune complexes with allergens. IgG is tethered to the membrane by binding the Fc fragment of FcγRIIb. When an allergen simultaneously binds IgE and IgG, the activating FcεRI is brought together with the inhibitory FcγRIIb, thereby silencing the FcεRI-mediated activation pathway. Zhu and colleagues[1] report a strategy that takes advantage of the natural capacity of FcγRIIb to inhibit the allergenic activity of FcεRI. They designed a chimeric molecule — a fusion of a cat allergen (Fel d1) and the Fc fragment of human IgG (Panel C) — that abolished allergic reactions to Fel d1 in vitro and in a mouse model. Similar fusion molecules could be designed to counter other types of allergy. Lyn, Fyn, and Syk are protein tyrosine kinases. P denotes phosphorylation.



FcγRIIb contains an immunoreceptor tyrosine-based inhibitory motif (ITIM), a modified version of the activating ITAM present in FcεRI. Like ITAM, ITIM must be phosphorylated by a tyrosine kinase to function. ITIM is phosphorylated when FcγRIIb is brought into proximity to FcεRI resulting in the aggregation of both receptors. Phosphorylated ITIM recruits phosphatases that inactivate key components in the activation pathway.[4,5]

To circumvent the uncertainties of classic allergen desensitization, Zhu and colleagues engineered a chimeric molecule in which the Fc fragment of human IgG1 is fused to the recombinant major cat allergen Fel d1. The fusion molecule thus binds both FcγRIIb and the IgE specific for Fel d1. Because IgE is already bound to FcεRI, the authors predicted that the complex would form a bridge between FcγRIIb and FcεRI, thereby inhibiting allergic reactions. And this is exactly what happened. The fusion protein inhibited allergen-mediated activation of basophils and mast cells in vitro. It also inhibited allergic responses after long-term immunization with Fel d1 in a relevant mouse model.

The potential of this novel approach is not limited to treating allergies to cats. Zhu et al. have provided proof of principle that this strategy may be effective in treating any type of allergy. For example, replacing Fel d1 with the major peanut allergen should help prevent the devastating anaphylactic responses to peanuts seen in many children. The fusion protein could also displace allergens already bound to IgE, in which case it would help terminate ongoing anaphylactic reactions. A potential problem, however, is that the fusion protein could actually induce or exacerbate an allergic reaction, because mast cells and basophils also express FcγRIIa, an ITAM-containing and activating IgG receptor. Classic desensitization may fail in a subgroup of patients because activating FcγRIIa prevails over inhibitory FcγRIIb. If so, it will be essential to identify this subgroup of patients before they receive therapy with fusion molecules, because the fusion molecules could aggregate with the two activating receptors FcεRI and FcγRIIa, thus inducing allergic reactions.

The study by Zhu et al. focuses on the second step of the allergic response, after allergen-specific immunoglobulins have been produced. However, the fusion molecule may have an even stronger effect on countering allergy if it is used to ablate IgE-producing B cells — the source of such immunoglobulins — which also express FcγRIIb. This approach would induce true desensitization.

From the Beth Israel Deaconess Medical Center, Harvard University, Boston.

1. Zhu D, Kepley CL, Zhang K, Terada T, Yamada T, Saxon A. A chimeric human-cat fusion protein blocks cat-induced allergy. Nat Med 2005;11:446-9.
2. Siraganian RP. Mast cell signal transduction from the high-affinity IgE receptor. Curr Opin Immunol 2003;15:639-46.
3. Rivera J. Molecular adapters in FcεRI signaling and the allergic response. Curr Opin Immunol 2002;14:688-93.
4. Ravetch JV, Laniel LL. Immune inhibitory receptors. Science 2000;290:84-9.
5. Daeron M. Fc receptor biology. Annu Rev Immunol 1997;15:203-34.

CORRESPONDENCE

Familial Cancer and *ARLTS1*

TO THE EDITOR: Calin et al. (April 21 issue)[1] report on *ARLTS1*, a novel tumor-suppressor gene with proapoptotic characteristics and a member of the ADP-ribosylation factor family. It is proposed that a protein-truncating mutation (G446A) predisposes patients to cancer on the basis of the higher frequency of this allele among patients with familial cancer than among patients with sporadic tumors. The allele frequency was population-dependent, but when patient groups were compared with the control group, this heterogeneity was not accounted for. Using Fisher's exact test, we arrive at a P value of 0.56 instead of 0.02 when comparing familial and sporadic cases and 0.30 when comparing familial cases with controls. The family used as an example lacks members at risk not carrying the polymorphism, and kindreds carrying the polymorphism remain cancer-free.

The transfection experiments in *ARLTS*-deficient cancer cells indicate limited tumor-suppressive activity, as compared with what can be obtained with, for example, the introduction of wild-type *p53* in *p53*-deficient cancer cells. The importance of the *ARLTS1* G446A polymorphism in familial cancer has not been demonstrated, and the contribution of this study to cancer pathogenesis and biology should be further substantiated with the use of, for example, transgenic experiments.

Sylvia De Brakeleer, M.S.
Erik Teugels, Ph.D.
Jacques De Grève, M.D., Ph.D.

Akademish Ziekenhuis–Vrije Universiteit Brussel
Jette 1090, Belgium
jacquesdegreve@skynet.be

1. Calin GA, Trapasso F, Shimizu M, et al. Familial cancer associated with a polymorphism in *ARLTS1*. N Engl J Med 2005;352:1667-76.

TO THE EDITOR: Calin et al. identified the tumor-suppressor gene *ARLTS1*, a member of the ARF–ARL family, and showed that its G446A (Trp149Stop) variant predisposes patients to familial cancer.[1] As a result, we investigated the influence of this nonsense mutation on the risk of familial breast cancer. Genomic DNA of female index patients who were negative for the *BRCA1* and *BRCA2* mutations and who were from 305 unrelated families with a high risk of breast cancer was analyzed by direct sequencing. There were two groups: families with two or more cases of breast cancer, including at least two cases with an onset before the age of 50 years, and families with one or more cases of breast cancer and at least one case of ovarian cancer[2]; both were recruited by the German Consortium for Hereditary Breast and Ovarian Cancer, from centers in Heidelberg, Cologne, and Munich. DNA from a control group of 530 unrelated women also was analyzed by direct sequencing.[3]

The Trp149Stop mutation was significantly more frequent among women from the high-risk families than among women in the control group (odds ratio, 2.20; 95 percent confidence interval, 1.02 to 4.74; P=0.04), indicating association with familial breast-cancer risk. Beyond this finding, *ARLTS1* Trp149Stop could be detected in three af-

THIS WEEK'S LETTERS

313 **Familial Cancer and *ARLTS1***
314 **Thrombosis of the Cerebral Veins and Sinuses**
315 **The Costs of Institutional Review Boards**
317 **Pesticide-like Poisoning from a Prescription Drug**
318 **Pregnancy after Transplantation of Cryopreserved Ovarian Tissue in a Patient with Ovarian Failure after Chemotherapy**

fected members of a high-risk family, emphasizing its cosegregation with disease.

Bernd Frank, Ph.D.

German Cancer Research Center
69120 Heidelberg, Germany
b.frank@dkfz.de

Rüdiger Klaes, M.D.

University of Heidelberg
69120 Heidelberg, Germany

Barbara Burwinkel, Ph.D.

German Cancer Research Center
69120 Heidelberg, Germany

1. Hashimoto S, Onodera Y, Hashimoto A, et al. Requirement for Arf6 in breast cancer invasive activities. Proc Natl Acad Sci U S A 2004;101:6647-52.
2. Meindl A, German Consortium for Hereditary Breast and Ovarian Cancer. Comprehensive analysis of 989 patients with breast or ovarian cancer provides BRCA1 and BRCA2 mutation profiles and frequencies for the German population. Int J Cancer 2002;97:472-80.
3. Frank B, Hemminki K, Wirtenberger M, et al. The rare ERBB2 variant Ile654Val is associated with an increased familial breast cancer risk. Carcinogenesis 2005;26:643-7.

THE AUTHORS REPLY: De Brakeleer et al. refer to Table 1 of our article, and in particular to the line that specifies that 6 of 109 familial cancers and 8 of 216 sporadic tumors had the G446A mutation. Analysis of these data with Fisher's exact test (the unit used is patients, not alleles) gives a P value of 0.56 for the association, as they have calculated. One of the footnotes in the table clarifies that the analyses and P values are restricted to those patients in whom the gene was directly sequenced and provides a P value of 0.02. This subgroup included 150 patients with sporadic tumors, of whom 3 had the G446A mutation, and 48 patients with familial tumors, of whom 5 had the mutation. The odds ratio is indeed 5.7 for this subgroup, and the P value associated with it is 0.02, as we reported. We should have stated more clearly the effect of restricting the analyses to the subgroup of patients whose gene was directly sequenced.

Our study was not designed to compare the tumor-suppressive effects of the *ARLTS1* gene with that of *TP53* or any other tumor-suppressor gene. We showed that the Trp149Stop variant had a limited effect on apoptosis and tumor suppression; this may explain why the variant is maintained in the general population.

We agree with De Brakeleer and colleagues that the importance of the *ARLTS1* G446A change in establishing a predisposition to familial cancer should be further substantiated. In fact, Frank et al. report that "the Trp149stop mutation was significantly more frequent among women from high-risk families than among women in the control group . . . indicating association with familial breast-cancer risk" and that the mutation "could be detected in three affected members of a high-risk family" and thus supports the role of *ARLTS1* gene in human cancer.

George Adrian Calin, M.D., Ph.D.
Carlo M. Croce, M.D.

Ohio State University Comprehensive Cancer Center
Columbus, OH 43210
carlo.croce@osumc.edu

Donna Neuberg, Sc.D.

Dana–Farber Cancer Institute
Boston, MA 02115

Thrombosis of the Cerebral Veins and Sinuses

TO THE EDITOR: The review by Stam of thrombosis of the cerebral veins and sinuses (April 28 issue)[1] did not include heparin-induced thrombocytopenia as one of the causes of thrombosis. This acquired hypercoagulability state results from the presence of heparin-dependent, platelet-activating IgG antibodies. Cerebral venous or sinus thrombosis occurred in 3 (2.9 percent) of 105 patients and in 7 (1.7 percent) of 408 patients in two studies[2,3] involving the clinical presentation of this adverse drug reaction. If sinus thrombosis is suspected in a patient who is currently using or has recently used heparin, heparin-induced thrombocytopenia must be ruled out, since this condition represents an absolute contraindication to further anticoagulant therapy with unfractionated or low-molecular-weight heparin (the usual treatment for sinus thrombosis). Alternative anticoagulant drugs appropriate for patients with heparin-induced thrombocytopenia include lepirudin, argatroban, and danaparoid.[4] The unfavorable course of cerebral venous and sinus thrombosis associated with heparin-induced thrombocytopenia includes both the progression of cerebral ischemia and the development

of noncerebral thrombosis.[2,5] This suggests that the use of alternative anticoagulants should not be withheld because of the thrombocytopenia and consumption coagulopathy.

Theodore E. Warkentin, M.D.

McMaster University
Hamilton, ON L8N 3Z5, Canada

Andreas Greinacher, M.D.

Ernst-Moritz-Arndt University
D-17487 Greifswald, Germany

Drs. Warkentin and Greinacher report having received lecture fees, consulting fees, or both in relation to the alternative nonheparin anticoagulant drugs lepirudin, argatroban, and danaparoid.

1. Stam J. Thrombosis of the cerebral veins and sinuses. N Engl J Med 2005;352:1791-8.
2. Pohl C, Klockgether T, Greinacher A, Hanfland P, Harbrecht U. Neurological complications in heprin-induced thrombocytopenia. Lancet 1999;353:1678-9.
3. Greinacher A, Farner B, Kroll H, Kohlmann T, Warkentin TE, Eichler P. Clinical features of heparin-induced thrombocytopenia including risk factors for thrombosis: retrospective analysis of 408 patients. Thromb Haemost (in press).
4. Warkentin TE, Greinacher A. Heparin-induced thrombocytopenia: recognition, treatment, and prevention: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004; 126:Suppl:311S-337S. [Erratum, Chest 2005;127:416.]
5. Warkentin TE, Bernstein RA. Delayed-onset heparin-induced thrombocytopenia and cerebral thrombosis after a single administration of unfractionated heparin. N Engl J Med 2003;348:1067-9.

TO THE EDITOR: The review of cerebral vein thrombosis by Stam was comprehensive and informative. However, given the increased use of various therapeutic agents associated with cerebral vein thrombosis, epoetin alfa[1] and tamoxifen[2,3] should be considered in the list of drugs related to this condition. The former causes polycythemia, and the latter a prothrombic state.

Pasquale F. Finelli, M.D.

Hartford Hospital
Hartford, CT 06102
nervline@aol.com

1. Finelli PF, Carley MD. Cerebral venous thrombosis associated with epoetin alfa therapy. Arch Neurol 2000;57:260-2.
2. Masjuan J, Pardo J, Callejo JM, Andres MT, Alvarez-Cermeno JC. Tamoxifen: a new risk factor for cerebral sinus thrombosis. Neurology 2004;62:334-5.
3. Finelli PF, Schauer PK. Cerebral sinus thrombosis with tamoxifen. Neurology 2001;56:1113-4.

TO THE EDITOR: The excellent review by Stam does not mention the important predisposing risk factor of high altitude. Cerebral venous sinus thrombosis has been well documented at high altitudes in the Himalayas and elsewhere.[1] High altitude can precipitate thrombosis because of hypobaric hypoxia and its effect on the coagulation system — in particular, increased factor VIIa activity.[2-4] Dehydration, polycythemia, and vascular spasms may also increase a tendency toward the development of cerebral venous thrombosis at high altitude.[2-4]

Ajit Singh Kashyap, M.D.

Command Hospital (Southern Command)
Pune 411 040, India
kashyapajits@hotmail.com

Kuldip Parkash Anand, M.D.

Command Hospital (Eastern Command)
Kolkata 700 027, India

Surekha Kashyap, M.D.

Command Hospital (Southern Command)
Pune 411 040, India

1. Basnyat B, Cumbo TA, Edelman R. Acute medical problems in the Himalayas outside the setting of altitude sickness. High Alt Med Biol 2000;1:167-74.
2. Bendz B, Rostrup M, Sevre K, Andersen TO, Sandset PM. Association between acute hypobaric hypoxia and activation of coagulation in human beings. Lancet 2000;356:1657-8.
3. Basnyat B, Graham L, Lee S-D, Lim Y. A language barrier, abdominal pain, and double vision. Lancet 2001;357:2022.
4. Kashyap AS, Kashyap S. The clot thickens! in thin air. Arch Intern Med 2002;162:1783.

The Costs of Institutional Review Boards

TO THE EDITOR: The data that Sugarman et al. (April 28 issue)[1] collected in 2002 seriously underestimate the current costs of human-research oversight. Since 2002, more than 75 percent of academic medical centers have begun preparing for, have applied for, or have achieved accreditation of their human-research protection programs (Speers M: personal communication). The self-study required by this process inevitably reveals deficiencies, and correction requires the investment of significant additional resources. Indeed, the accreditation process in itself may play an important role in eliminating the inadequate support for these programs that has been reported.[2]

Most research universities are unable to recover these costs (or those of related, unfunded federal mandates) from federal research sponsors because of the cap on administrative indirect costs imposed by the Office of Management and Budget in 1993. The cap is unfair and perverse: unfair, because it applies only to universities (and not to teaching hospitals, nonprofit research institutes, or for-profit research organizations), and perverse, because rather than promoting investment in strengthened institutional oversight, it is a negative incentive that most severely penalizes those institutions willing to go the extra mile in protecting human research subjects.

Howard B. Dickler, M.D.
David Korn, M.D.

Association of American Medical Colleges
Washington, DC 20037
hdickler@aamc.org

1. Sugarman J, Getz K, Speckman JL, Byrne MM, Gerson J, Emanuel EJ. The cost of institutional review boards in academic medical centers. N Engl J Med 2005;352:1825-7.
2. Department of Health and Human Services. Institutional review boards: a time for reform. Washington, D.C.: Office of Inspector General, June 1998. (OEI-01-97-00193).

TO THE EDITOR: Sugarman et al. emphasize the financial burdens brought about by institutional review boards (IRBs), which are an important mechanism to protect human research participants. Whereas IRBs in the United States may be underfunded, resources for IRBs in the developing world are acutely scarce — a deficiency that we recently encountered firsthand in India. In many countries where active research involving human subjects is occurring, costs have come to present a major barrier to effective ethical review.[1] Particularly because human subjects in developing nations are increasingly becoming the focus of studies originating in the United States and because ethical review by both the sponsoring institution and a local committee is important for ethical coherency,[2] the academic community in the developed world is partially responsible for promoting a worldwide ethical-review infrastructure. This effort may, in turn, mean that institutions in developing nations will need to be educated about maintaining efficient IRBs or, in some cases, that independent, systemic financial support will need to be provided. Science and ethics are integrally linked; it is essential that effective IRBs be supported wherever research involving human subjects occurs.

Suneel Bhat

Princeton University
Princeton, NJ 08544
sbhat@princeton.edu

Timmanna T. Hegde, M.B., B.S.

Indian Medical Association
Kumta 581343, India

1. Hyder AA, Wali SA, Khan AN, Teoh NB, Kass NE, Dawson L. Ethical review of health research: a perspective from developing country researchers. J Med Ethics 2004;30:68-72.
2. Shapiro HT, Meslin EM. Ethical issues in the design and conduct of clinical trials in developing countries. N Engl J Med 2001; 345:139-42.

THE AUTHORS REPLY: Determining the true costs of operating IRBs is a critical first step in identifying ways of reimbursing academic medical centers for this important activity. Dickler and Korn claim that in our letter we "seriously underestimate" the current costs of operating IRBs in academic medical centers. Perhaps we do, since our estimates represent a snapshot of costs at a single point in time; the costs of many things change, and there can be intermittent needs for increased spending. Indeed, Dickler and Korn suggest that we underestimate costs because of the substantial expense associated with voluntary accreditation and the concomitant recognition of the need to enhance IRB systems within institutions. In addition, for our index year, academic medical centers expended considerable effort in preparing for compliance with the Health Insurance Portability and Accountability Act, an effort that should be minimized in later years. Thus, it seems naive simply to assume that the costs of IRBs must be higher now. Rather, it would be helpful if actual data on this point were provided. Nevertheless, our work offers a benchmark to help guide academic medical centers, government sponsors, regulators, and others to determine what resources IRBs require. We welcome future efforts that measure the costs of IRBs longitudinally.

Regardless of the disagreement about the precision of our benchmark, Dickler and Korn raise an important issue regarding the ability of academic medical centers to recover the costs of IRBs. Data such as those we present in our letter should be helpful in bringing this and related reimbursement issues into focus.

Bhat and Hegde claim that IRB work in developing nations is often underfunded. Again, it would be helpful to have actual data on the current levels

of funding and the costs of this activity, which is aimed at being a key mechanism for ensuring the protection of research participants in these settings. Until there are better data regarding the true costs of IRB oversight in developing countries, our benchmarks may provide a means for estimating the time needed for oversight and thus help international organizations and sponsors establish appropriate levels of support for it.

Jeremy Sugarman, M.D., M.P.H.
Johns Hopkins University
Baltimore, MD 21202

Ezekiel J. Emanuel, M.D., Ph.D.
National Institutes of Health
Bethesda, MD 20892

for the Consortium to Examine Clinical Research Ethics

Pesticide-like Poisoning from a Prescription Drug

TO THE EDITOR: Acetylcholinesterase inhibitors have long been known for their use as pesticides. Since the 1990s, acetylcholinesterase-inhibiting pharmaceuticals (donepezil, tacrine, galantamine, and rivastigmine) have been used to treat Alzheimer's disease; they are currently being investigated as a treatment for dementia associated with Parkinson's disease in adults[1] and with Tourette's syndrome and autistic and attention deficit–hyperactivity disorders in children.[2,3] Wider use of this class of medications for a broader variety of disorders increases the possibility of pesticide-like poisoning from a prescribed medication. We report a case of such poisoning.

A healthy 11-month-old girl (weight, 7.5 kg) presented to a pediatric hospital with rapid onset of general weakness during a period of several hours. The infant had been discharged from the same hospital three days earlier after undergoing intravenous rehydration for gastroenteritis. On arrival at the second visit, she was hypotonic, hyporeflexic, and had miosis and a weak cry. She had had no diarrhea, wet diapers, or lacrimation. Neurology was consulted; a workup was initiated to rule out botulism and Guillain–Barré syndrome. The levels of electrolytes, blood urea nitrogen, creatinine, and glucose and the complete blood count were normal, as were the results of liver-function tests, electrocardiography, urinalysis, computed tomography of the head, and lumbar-puncture studies. There was no family history of myasthenia gravis or neuromuscular disorders. The family did not keep plants in the home, and they did not use rodenticides. Since it was winter in New England, there was no outdoor exposure to insecticides. Further questioning revealed that the patient's mother had found the child chewing a capsule containing rivastigmine, a drug taken by the grandmother, earlier in the day, but because of history-taking and language problems (the family was first-generation Vietnamese), the connection was not initially made between the drug exposure and the child's symptoms.

Rivastigmine is a centrally acting carbamate derivative and a reversible inhibitor of acetylcholinesterase that produces a self-limited cholinergic constellation of signs and symptoms consistent with a specific drug class or xenobiotic poisoning that rarely lasts more than 24 hours. Muscarinic effects may be treated with atropine and nicotinic effects with oximes such as pralidoxime. However, muscarinic effects are often absent in children,[4] as was the case with our patient, and the administration of oximes does not appear to speed recovery from nicotinic effects in such cases.[5] Our patient improved with supportive care alone, and by 24 hours after she had ingested the drug, she was sitting without assistance; by 48 hours, she had regained her normal strength and was discharged.

Our patient had notable, predominantly nicotinic cholinergic effects after exposure to a nonpesticide carbamate acetylcholinesterase inhibitor intended for the treatment of memory-impaired adults. As the use of this class of drugs becomes more widespread, we want to alert clinicians to consider such exposure when evaluating weakness of rapid onset or any case of pesticide-like poisoning in which pesticide exposure is unlikely.

Melisa W. Lai, M.D.
Children's Hospital
Boston, MA 02115
melisa.lai@alumni.brown.edu

Matthew Moen, Pharm.D.
Massachusetts College of Pharmacy and Health Sciences
Boston, MA 02115

Michele Burns Ewald, M.D.
Children's Hospital
Boston, MA 02115

1. Emre M, Aarsland D, Albanese A, et al. Rivastigmine for dementia associated with Parkinson's disease. N Engl J Med 2004;351:2509-18.
2. Chez MG, Aimonovitch M, Buchanan T, Mrazek S, Tremb RJ. Treating autistic spectrum disorders in children: utility of the cholinesterase inhibitor rivastigmine tartrate. J Child Neurol 2004;19:165-9.
3. Biederman J, Spencer T. Non-stimulant treatments for ADHD. Eur Child Adolesc Psychiatry 2000;9:Suppl 1:I-51–I-59.
4. Lifshitz M, Shahak E, Sofer S. Carbamate and organophosphate poisoning in young children. Pediatr Emerg Care 1999;15:102-3.
5. Lifshitz M, Rotenberg M, Sofer S, Tamiri T, Shahak E, Almog S. Carbamate poisoning and oxime treatment in children: a clinical and laboratory study. Pediatrics 1994;93:652-5.

Pregnancy after Transplantation of Cryopreserved Ovarian Tissue in a Patient with Ovarian Failure after Chemotherapy

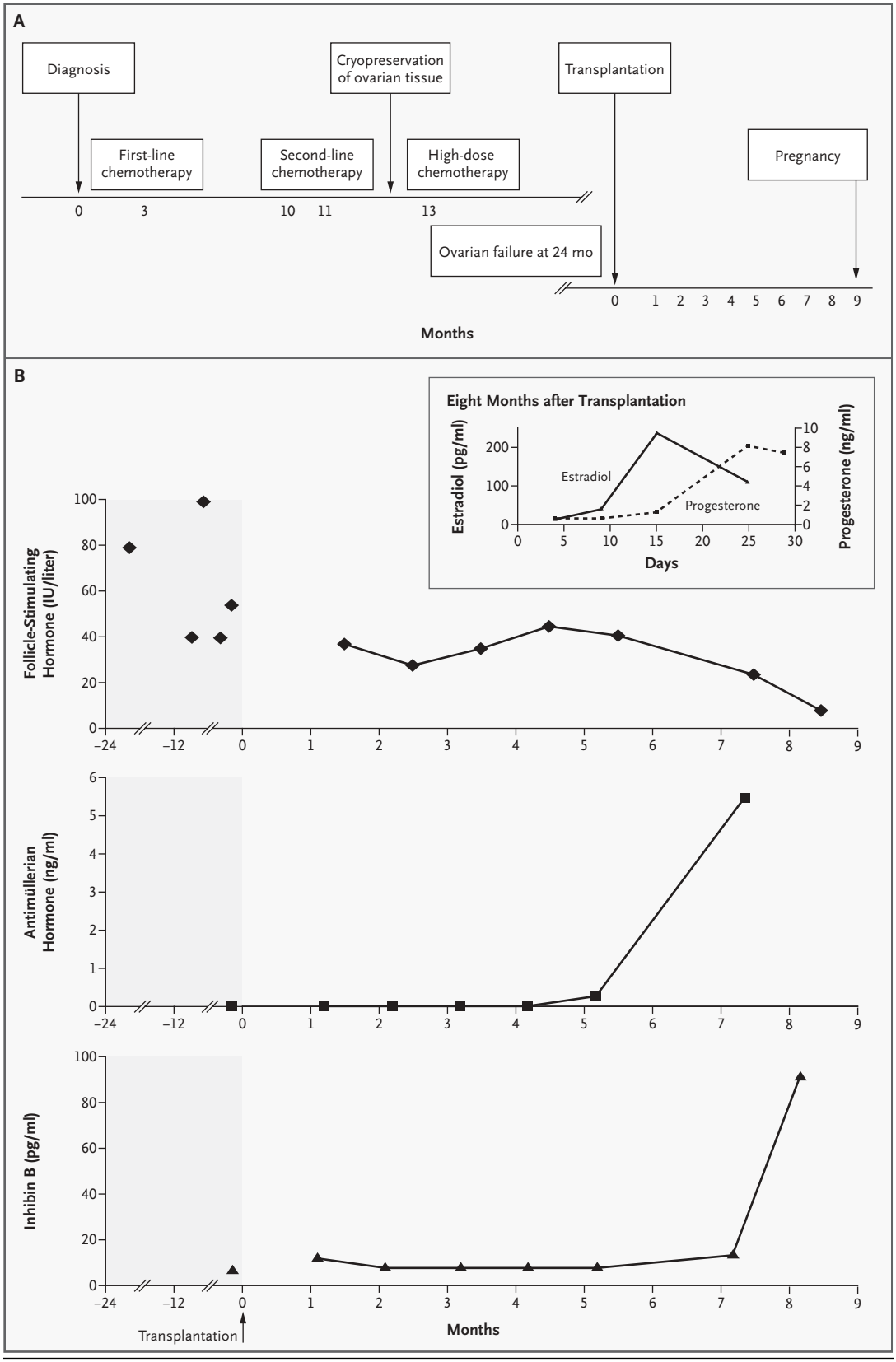
TO THE EDITOR: Premenopausal women who undergo high-dose chemotherapy have a very high risk of ovarian failure.[1] Cryopreservation of ovarian tissue with subsequent autotransplantation has effectively preserved fertility in an animal model,[2] but its efficacy in humans has been uncertain. Eggs that were aspirated from cryopreserved ovarian tissue transplanted in heterotopic sites did not result in a pregnancy.[3] A live birth was reported after transplantation of cryopreserved ovarian tissue in a woman who had undergone treatment for Hodgkin's disease; however, since the woman had ovulated before transplantation, it is uncertain whether the egg came from the native ovary or the transplanted ovary.[4] A recent report described a live birth after transplantation of fresh ovarian tissue from a fertile woman to her sterile monozygotic twin, but this approach does not involve preservation of fertility and hence is not applicable to women facing sterilizing chemotherapy.[5]

We describe a live birth after in vitro fertilization following the transplantation of thawed cryopreserved ovarian cortical tissue into the ovaries of a 28-year-old woman who had ovarian failure after high-dose chemotherapy for non-Hodgkin's lymphoma. Ovarian tissue (containing many primordial follicles) was harvested after administration of a second-line conventional chemotherapy regimen, before treatment with high-dose chemotherapy (Fig. 1A). The patient's menses ceased after the high-dose chemotherapy. During the ensuing 24 months, the amenorrhea persisted, and laboratory testing consistently revealed high levels of follicle-stimulating hormone and luteinizing hormone (40 to 104 IU per liter) and undetectable levels of antimüllerian hormone and inhibin B — findings consistent with ovarian failure (Fig. 1B).

At 24 months, the patient remained free of disease and requested autotransplantation of the

Figure 1 (facing page). Time Line of Treatment (Panel A) and Hormone Levels (Panel B).

After non-Hodgkin's lymphoma was diagnosed, a first-line chemotherapy regimen (etoposide, doxorubicin, cyclophosphamide, vincristine, bleomycin, and corticosteroids [VACOP-B]) was administered (Panel A), but a relapse occurred six months later; during that period the patient had monthly menstrual cycles. A second-line regimen (mesna, ifosfamide, mitoxantrone, etoposide, cytarabine, cisplatin, and corticosteroids [MINE–ESHAP]) was administered, followed by high-dose chemotherapy (carmustine, 300 mg per square meter of body-surface area; etoposide, 1600 mg per square meter; cytarabine, 1600 mg square meter; and melphalan 140 mg per square meter [BEAM]) with autologous stem-cell support. Ovarian tissue was harvested for cryopreservation before high-dose chemotherapy. Transplantation of thawed cryopreserved ovarian tissue was performed after 24 months of persistent ovarian failure. During the ninth month after transplantation, in vitro fertilization resulted in a pregnancy. Panel B shows basal blood levels of follicle-stimulating hormone, antimüllerian hormone, and inhibin B before and after transplantation. Six months after transplantation, the levels of follicle-stimulating hormone began to decrease gradually; normal levels were attained in the ninth month. The levels of antimüllerian hormone (based on three to five measurements every cycle) were undetectable until the sixth month after transplantation. High levels of antimüllerian hormone were measured during the eighth month after transplantation — a finding compatible with good ovarian reserve. Basal levels of inhibin B were undetectable after transplantation but became high during the ninth month. The levels of follicle-stimulating hormone were measured by a chemiluminescent immunometric method (Immulite 2000, Diagnostic Products); inhibin B levels were measured with the use of two-site enzyme-linked immunosorbent assays (Serotec); and levels of antimüllerian hormone were measured with the use of an ultrasensitive two-site enzyme-linked immunosorbent assay (Diagnostic Systems Laboratories). The inset shows hormone secretion during the eighth month after transplantation. Spontaneous menstruation was followed by a rise in estradiol levels, and the rise in progesterone levels after the 15th cycle day indicated luteinization.



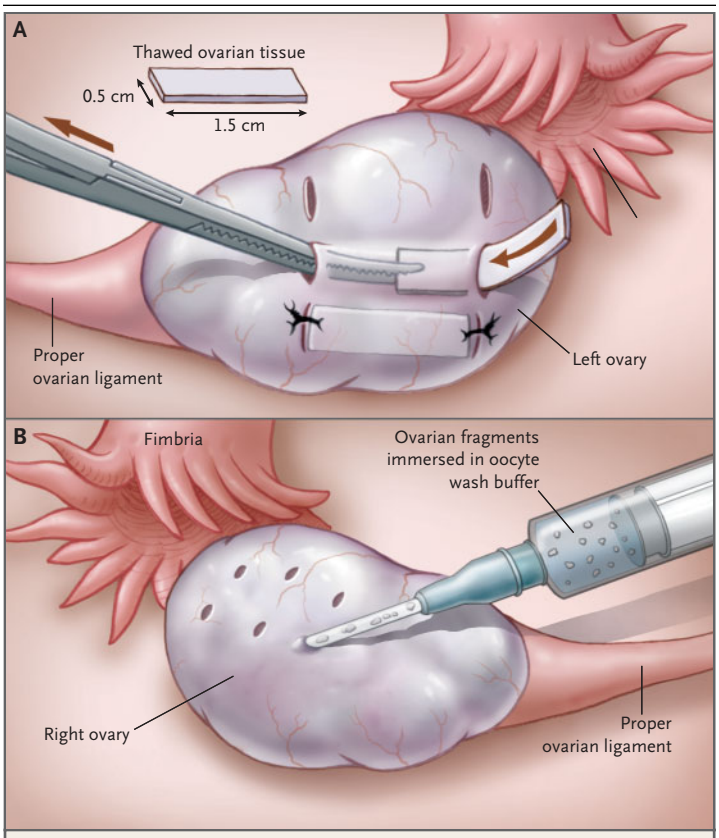


Figure 2. Surgical Technique.

Three pairs of 5-mm transverse incisions were made in the left ovary through the tunica albuginea (Panel A). With blunt dissection, cavities were formed beneath the cortex for each of the three strips. Each piece of thawed ovarian tissue (1.5 by 0.5 cm in area and 0.1 to 0.2 cm in thickness) was gently placed in a cavity, and the incisions were closed with 4/0 Vicryl sutures. In the smaller, right ovary, tiny ovarian fragments immersed in oocyte wash buffer were injected beneath the cortex (Panel B). Only the ovarian strips placed in the left ovary resumed function.

ovarian tissue in an attempt to restore fertility. After approval from the institutional review board and the patient's written informed consent had been obtained, a laparotomy was performed; strips of thawed ovarian tissue were transplanted to the left ovary, and small fragments were injected into the right ovary (Fig. 2). Eight months after transplantation, the patient spontaneously menstruated. Basal levels of antimüllerian hormone (which previously was undetectable) were found to be high, a finding consistent with the presence of active, early-stage, growing follicles.[6] This change was followed by a rise in inhibin B levels to the levels reported in ovulatory women (Fig. 1B). Ultrasonography revealed a preovulatory follicle in the left ovary. The time from transplantation to recovery was compatible with the time needed for the growth and maturation of primordial follicles.[7]

Nine months after transplantation, the patient had a second spontaneous menstrual period. The level of follicle-stimulating hormone was 7.9 IU per liter, the level of luteinizing hormone 6.8 IU per liter, the level of estradiol 118 pg per milliliter, and the level of progesterone 0.5 ng per milliliter. A decision was made to perform in vitro fertilization. After a modified natural cycle,[8] a single mature egg with a large cumulus was retrieved. The egg was fertilized in vitro with sperm from the patient's husband, and two days later, a four-cell embryo was transferred to the uterus. Serum testing for human chorionic gonadotropin was positive 12 days after the embryo transfer. Repeated ultrasonography during the pregnancy showed normal fetal growth and development. At 38 weeks 5 days of gestation, a healthy-appearing female infant weighing 3000 g was delivered by cesarean section. The Apgar scores were 9 at one minute and 10 at five minutes.

Transplantation of ovarian tissue is associated with a theoretical risk of grafting malignant cells. Tissue was harvested in this patient after therapy, with no evidence of disease, and conventional histologic analyses showed no cancer cells.

Although we cannot rule out the possibility that the egg was derived from the native ovary, we consider this possibility very unlikely, given the consistent evidence of ovarian failure after high-dose chemotherapy and the timing of restoration of ovarian function after transplantation. The hormone levels provided strong evidence of the success of transplantation, despite its being performed after initial chemotherapy, rather than of the activity of a few residual follicles. Our results indicate that fertility preservation with cryopreservation and orthotopic transplantation of ovarian tissue can be successfully performed in humans.

Dror Meirow, M.D.
Jacob Levron, M.D.

Chaim Sheba Medical Center
Tel Hashomer, Israel
meirow@post.tau.ac.il

Talia Eldar-Geva, M.D., Ph.D.

Shaare-Zedek Medical Center
Jerusalem, Israel

Izhar Hardan, M.D.
Eduard Fridman, M.D.
Yaron Zalel, M.D.
Eyal Schiff, M.D.
Jehoshua Dor, M.D.
Chaim Sheba Medical Center
Tel Hashomer, Israel

1. Meirow D, Nugent D. The effects of radiotherapy and chemotherapy on female reproduction. Hum Reprod Update 2001;7:535-43.
2. Gosden RG, Baird DT, Wade JC, Webb R. Restoration of fertility to oophorectomized sheep by ovarian autografts stored at –196°C. Hum Reprod 1994;9:597-603.
3. Donnez J, Dolmans MM, Demylle D, et al. Livebirth after orthotopic transplantation of cryopreserved ovarian tissue. Lancet 2004; 364:1405-10. [Erratum, Lancet 2004;364:2020.]
4. Oktay K, Buyuk E, Veeck L, et al. Embryo development after heterotopic transplantation of cryopreserved ovarian tissue. Lancet 2004;363:837-40.
5. Silber SJ, Lenahan KM, Levine DJ, et al. Ovarian transplantation between monozygotic twins discordant for premature ovarian failure. N Engl J Med 2005;353:58-63.
6. Fanchin R, Schonauer LM, Righini C, Guibourdenche J, Frydman R, Taieb J. Serum anti-Mullerian hormone is more strongly related to ovarian follicular status than serum inhibin B, estradiol, FSH and LH on day 3. Hum Reprod 2003;18:323-7.
7. Gougeon A. Regulation of ovarian follicular development in primates: facts and hypotheses. Endocr Revs 1996;17:121-55.
8. Elizur SE, Aslan D, Shulman A, Weisz B, Bider D, Dor J. Modified natural cycle using GnRH antagonist can be an optional treatment in poor responders undergoing IVF. J Assist Reprod Genetics 2005;22:75-9.

This letter was published on June 27, 2005, at www.nejm.org.

INSTRUCTIONS FOR LETTERS TO THE EDITOR

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

Our Web address: **http://authors.nejm.org**

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

BOOK REVIEWS

PATIENT SAFETY: ACHIEVING A NEW STANDARD FOR CARE

(Quality Chasm Series.) Edited by Philip Aspden, Janet M. Corrigan, Julie Wolcott, and Shari M. Erickson. 528 pp. Washington, D.C., National Academies Press, 2004. $44.95. ISBN 0-309-09077-6.

ACCOUNTABILITY: PATIENT SAFETY AND POLICY REFORM

(Hastings Center Studies in Ethics.) Edited by Virginia A. Sharpe. 276 pp. Washington, D.C., Georgetown University Press, 2004. $49.95. ISBN 1-58901-023-X.

THE BELIEF THAT SYSTEMS, MORE THAN individuals, are responsible for medical errors permeates current efforts to improve the quality of medical care and patient safety. Error reduction thus requires that systems be reformed, which necessitates far more information than is currently available. *Patient Safety: Achieving a New Standard for Care* articulates a vision of "a new health care delivery system . . . that both prevents errors and learns from them when they occur," to be achieved through a national infrastructure of health information. The technical complexity alone that would be required to develop such an infrastructure would guarantee full employment for a generation of information-management technocrats and computer geeks. But the authors of *Patient Safety* foresee continuous, active use of the health information infrastructure to improve patient safety, and this would vastly expand the complexity of the undertaking.

The use of a comprehensive, well-functioning information system to enhance the processes and outcomes of medical care also encompasses a human component, one that goes beyond sophisticated scientific and technological expertise. *Patient Safety* acknowledges those human factors, noting the need for cultural change if a new health care system that minimizes errors is to emerge. The essays in *Accountability: Patient Safety and Policy Reform* explore human and cultural elements in depth, offering thought-provoking commentaries on the rapidly burgeoning patient-safety movement. The goal of *Accountability* is "to analyze the values and ethical issues at stake in policy proposals on patient safety and the reduction of medical error." At the least, cultural considerations span legal, ethical, philosophical, and psychological spheres, raising fundamental questions as to the appropriateness and achievability of improving both information systems and the quality of health care.

The two books share some heritage and subject matter, but little else. The genesis for each included the landmark Institute of Medicine report *To Err Is Human: Building a Safer Health System* (National Academies Press, 2000) and related publications, but these works go in different directions, which probably reflects their respective authorship. *Patient Safety* was produced by the Committee on Data Standards for Patient Safety and the Board on Health Care Services, both of the Institute of Medicine, whose members' expertise is predominantly in the fields of clinical care, informatics, and other forms of technology. The authors of the chapters in *Accountability* include relatives of persons harmed by medical errors, ethicists, lawyers, and health professionals.

Taken together, the two books provide contrasting perspectives on data-driven improvements in patient safety. *Patient Safety* is an exhaustive review of the extraordinary requirements of an infrastructure to improve care in our complex and incomplete assortment of health care delivery arrangements, and the book documents the laudable efforts of public agencies and the private sector to move in that direction. By showing us the extent of cultural adaptation that would be necessary, even if the technical components can be developed, *Accountability* forces the question of whether a visionary new health care system is a realistic goal or will remain no more than a vision toward which we should strive.

Patient Safety compiles and discusses the technical and operational requirements for the establishment of a national health information infrastructure and its use in the improvement of health care quality and safety. The current flow of health care information invokes an image of myriad isolated populations, each with different and incomplete knowledge bases, speaking different languages over incompatible cell phones. *Patient Safety* argues that improvements in care require data on virtually the full spectrum of health care components, which encompass nearly every step in every health care encounter, and more. Such information must be col-

lected and maintained in systems that can talk to one another, with universally accepted standards in terms of format and terminology.

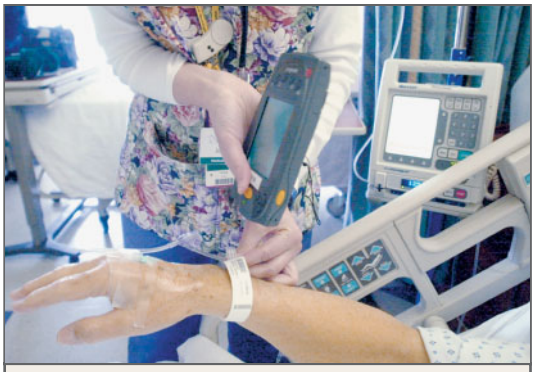
Furthermore, the actual improvement of patient safety requires that the wealth of information be used actively, for the ongoing reengineering of the health care system. The information must be linked to a series of computerized models and other measures that support the development and application of an enhanced evidence base for medical practice. The system must allow for the identification and analysis not only of adverse events that produce harm through errors of either commission or omission, but also of the larger universe of near misses, in which potential harm is averted.

Perhaps as an understatement of the remaining practical and technical impediments, the *Patient Safety* authoring committee asserts that the barriers that have prevented the implementation of "a health information infrastructure that supports learning and accountability systems for patient safety" are not primarily technological ones: "Rather, the lack of technology implementation and the failure to use common data standards have been the principal barriers." The solution includes a fostering of leadership, the provision of financial incentives and technical assistance, and the enforcement of safeguards for privacy and security.

Accountability sheds light on where all this information would have to come from and why it is not currently available. The book begins with three stories of persons who suffered catastrophic harm from medical errors, as told by relatives who describe the consequences for the patients and their families. These narratives describe how errors were handled by health care providers — that they often involved defensive posturing, concealment, and insensitivity to the needs and feelings of both victims and survivors. The subsequent scholarly essays provide insight into these unfortunate responses to medical tragedies, discussing deep-seated human and cultural characteristics that may not easily be overcome. Although the authors accept that there is a need for systematic reforms, several of them raise the concern that the vast amount of information required to populate the new infrastructure might never be reported or generate the anticipated improvements in medical care.

A central tenet of the patient-safety movement, as reflected in *Patient Safety*, is that health care professionals and other health workers would behave differently and be willing to share information if the "blame and shame" character of our current system were replaced by a supportive, nonpunitive environment that rewarded the full reporting of adverse events and near misses. In *Accountability*, E. Haavi Morreim, a professor of human values and ethics, questions "whether the proposed protections for error-reporters will actually yield all the hoped-for information," citing "ego, embarrassment, peer ostracism, loss of reputation, . . . skepticism that their reports will actually lead to useful change, or a desire to avoid the inconvenience of explaining what happened" as additional barriers to admitting mistakes. Physician and ethicist Edmund D. Pellegrino cautions us to recognize "the associated dangers of complacency and dulling of the moral sensibilities of the humans in the system when either a 'blame-free' approach or a 'blame-the-system' approach is adopted." As a result, "the interaction of a system of error prevention with preservation of a sense of individual responsibility and culpability is essential." Legal scholar and physician William M. Sage explores the complex role that reputation plays in physicians' behavior, concluding that "self-regulatory models that rely on professional peer review and informal sanctions to surface, analyze, and prevent error are inadequate substitutes for public surveillance and accountability." Ethicist Nancy Berlinger reinforces the human facet, reminding us that the adverse consequences of medical error "happen to patients," so that systems that promote "confession, repentance, and forgiveness" that are limited to the professional side of medical errors come up short by failing to ensure justice for patients.

Early on, the *Patient Safety* committee states, "Americans should be able to count on receiving health care that is safe." In the last chapter of *Accountability*, Morreim argues that "probably the most



A Nurse Uses a Handheld Scanner to Read a Bar Code on a Patient's Wrist before Giving Medication.

AP Photo/Wenatchee World. Photograph by Mike Bonnicksen.

important element in any system for addressing errors would be to ensure universal access to a reasonable level of health care for all citizens" and further asserts that differential access to insured care distorts the management of error and skews the compensation system. Morreim's statement suggests a paraphrase from *Patient Safety*: Americans should be able to count on receiving health care, and that will help ensure their care is safe. This should further caution us that we will come up short if we achieve a delivery system that minimizes error but serves only a portion of our population.

Peter P. Budetti, M.D., J.D.

University of Oklahoma College of Public Health
Oklahoma City, OK 73104
peter-budetti@ouhsc.edu

MEDICAL ERRORS AND MEDICAL NARCISSISM

By John D. Banja. 229 pp. Sudbury, Mass., Jones and Bartlett, 2005. $49.95. ISBN 0-7637-8361-7.

Why do physicians have such a difficult time talking to their patients about errors made in the course of patient care? This is the main question posed in John Banja's carefully written and useful book. The answer, the author suggests, is primarily "medical narcissism," a muted version of the narcissistic personality that, in many ways, also helps physicians do their difficult work.

Banja suggests that medical narcissists may find the disclosure of an error to be too much of a challenge to their self-image of competence, control, and "treatment-oriented focus." Hence, they have a tendency to rationalize the error as unavoidable, unimportant, or unnecessary to reveal because it will not change the outcome.

Banja argues for a policy of full disclosure to patients as a moral responsibility of the physician and a moral right of the patient. He notes that in 1981, the American Medical Association endorsed the position that errors should be "truthfully and honestly disclosed regardless of their legal consequences." He also suggests that full disclosure can rebuild the shattered doctor–patient relationship after an error has been made and can heal the hurt that both parties feel.

The author touches on the negative role of the tort system and explores the potential effect of tort reform and such alternatives as no-fault insurance and enterprise liability. It is encouraging to see Banja's recognition that current ideas of tort reform reflect the needs of the insurance industry more than they do the needs of either patients or doctors and will do little to promote full disclosure or reduce error. Banja fails to note that any authentic replacement of the tort system has to include rigorous efforts on the part of the medical profession itself to reduce medical errors and improve patient safety. The Swedish no-fault health insurance program rigidly separates compensation for patients from punishment of physicians. Hence, patients are compensated on the basis of degree of harm, not the fault of the doctor.

Because the book focuses on psychological arguments for full disclosure and the medical narcissism that stands in the way, Banja misses another absolutely crucial argument for full disclosure — the need for authentic information about error and patterns of error. The sharing of complete information is the only way that the medical profession can make progress on the reduction of errors. Full disclosure to the patient is a worthy goal, but it must be combined with full disclosure to the institution where the error took place.

One of many ways to improve the dissemination of data is to reshape the dynamics of the morbidity and mortality conference and to make the incident report an acceptable mechanism for physicians. The conference can become not just a mechanism for admitting error but a vehicle for collecting information, and the incident report should be seen as a mechanism for all clinicians, not just for nurses.

It would also improve the understanding of patterns of error (and the language of full disclosure) if physicians would examine more closely what they call "unavoidable," "anticipated," or "known" adverse events. A close examination of empirical studies shows that many such events are more preventable than is commonly perceived.

Full disclosure to patients — and to the health care delivery system — will give the medical profession the tools to maintain the trust that is essential in the doctor–patient relationship and to understand and reduce the patterns of medical error. A combination of these efforts will enhance the ability of physicians to serve their patients according to the standards that both parties ardently seek.

Marilynn M. Rosenthal, Ph.D.

University of Michigan Medical School
Ann Arbor, MI 48104
mmrosent@umich.edu

important element in any system for addressing errors would be to ensure universal access to a reasonable level of health care for all citizens" and further asserts that differential access to insured care distorts the management of error and skews the compensation system. Morreim's statement suggests a paraphrase from *Patient Safety*: Americans should be able to count on receiving health care, and that will help ensure their care is safe. This should further caution us that we will come up short if we achieve a delivery system that minimizes error but serves only a portion of our population.

Peter P. Budetti, M.D., J.D.

University of Oklahoma College of Public Health
Oklahoma City, OK 73104
peter-budetti@ouhsc.edu

MEDICAL ERRORS AND MEDICAL NARCISSISM

By John D. Banja. 229 pp. Sudbury, Mass., Jones and Bartlett, 2005. $49.95. ISBN 0-7637-8361-7.

Why do physicians have such a difficult time talking to their patients about errors made in the course of patient care? This is the main question posed in John Banja's carefully written and useful book. The answer, the author suggests, is primarily "medical narcissism," a muted version of the narcissistic personality that, in many ways, also helps physicians do their difficult work.

Banja suggests that medical narcissists may find the disclosure of an error to be too much of a challenge to their self-image of competence, control, and "treatment-oriented focus." Hence, they have a tendency to rationalize the error as unavoidable, unimportant, or unnecessary to reveal because it will not change the outcome.

Banja argues for a policy of full disclosure to patients as a moral responsibility of the physician and a moral right of the patient. He notes that in 1981, the American Medical Association endorsed the position that errors should be "truthfully and honestly disclosed regardless of their legal consequences." He also suggests that full disclosure can rebuild the shattered doctor–patient relationship after an error has been made and can heal the hurt that both parties feel.

The author touches on the negative role of the tort system and explores the potential effect of tort reform and such alternatives as no-fault insurance and enterprise liability. It is encouraging to see Banja's recognition that current ideas of tort reform reflect the needs of the insurance industry more than they do the needs of either patients or doctors and will do little to promote full disclosure or reduce error. Banja fails to note that any authentic replacement of the tort system has to include rigorous efforts on the part of the medical profession itself to reduce medical errors and improve patient safety. The Swedish no-fault health insurance program rigidly separates compensation for patients from punishment of physicians. Hence, patients are compensated on the basis of degree of harm, not the fault of the doctor.

Because the book focuses on psychological arguments for full disclosure and the medical narcissism that stands in the way, Banja misses another absolutely crucial argument for full disclosure — the need for authentic information about error and patterns of error. The sharing of complete information is the only way that the medical profession can make progress on the reduction of errors. Full disclosure to the patient is a worthy goal, but it must be combined with full disclosure to the institution where the error took place.

One of many ways to improve the dissemination of data is to reshape the dynamics of the morbidity and mortality conference and to make the incident report an acceptable mechanism for physicians. The conference can become not just a mechanism for admitting error but a vehicle for collecting information, and the incident report should be seen as a mechanism for all clinicians, not just for nurses.

It would also improve the understanding of patterns of error (and the language of full disclosure) if physicians would examine more closely what they call "unavoidable," "anticipated," or "known" adverse events. A close examination of empirical studies shows that many such events are more preventable than is commonly perceived.

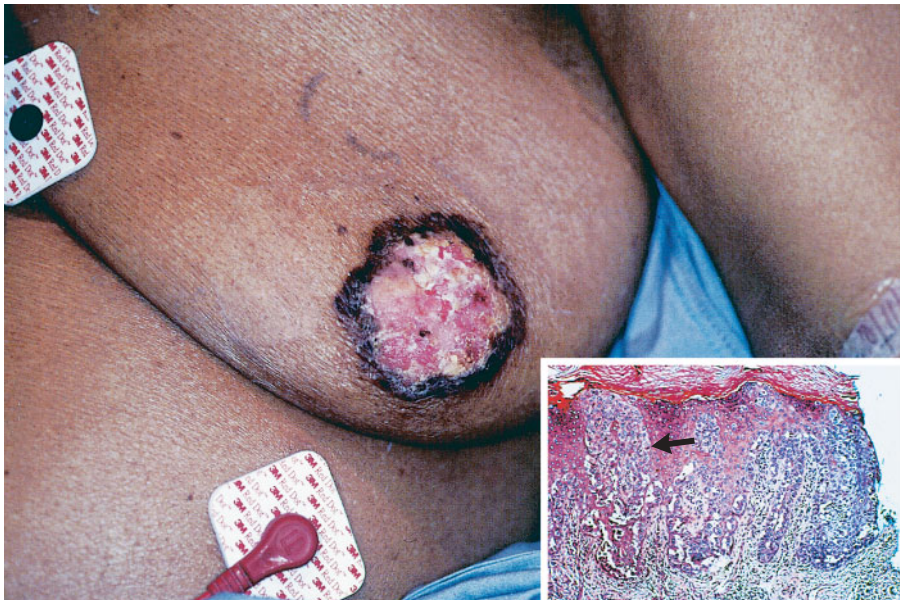
Full disclosure to patients — and to the health care delivery system — will give the medical profession the tools to maintain the trust that is essential in the doctor–patient relationship and to understand and reduce the patterns of medical error. A combination of these efforts will enhance the ability of physicians to serve their patients according to the standards that both parties ardently seek.

Marilynn M. Rosenthal, Ph.D.

University of Michigan Medical School
Ann Arbor, MI 48104
mmrosent@umich.edu

IMAGES IN CLINICAL MEDICINE

Paget's Disease of the Breast



A 70-year-old woman presented with an erythematous, scaly plaque with a hyperpigmented border that had replaced the areola and completely effaced the nipple of the left breast. No breast mass or lymphadenopathy was detected. A skin biopsy demonstrated large round cells, with sizable nuclei and abundant, pale-staining cytoplasm, permeating singly and in groups throughout the epidermis (inset, arrow). Immunohistochemical staining identified carcinoembryonic antigen, confirming the diagnosis of Paget's disease. Mammography and ultrasonography demonstrated no underlying abnormality in the left breast. A needle-core biopsy of a hypoechoic, ill-defined nodule at the areolar margin of the right breast showed benign, nonproliferative fibrocystic changes without atypia. The patient declined further treatment and was lost to follow-up. Paget's disease is often associated with underlying in situ or invasive carcinoma of the breast. Careful physical examination for a palpable breast mass or lymphadenopathy is an essential part of the evaluation.

Lindy Peta Fox, M.D.

Yale University School of Medicine
New Haven, CT 06520

Marc E. Grossman, M.D.

Columbia University Medical Center
New York, NY 10032