Volume 356 — May 24, 2007 — Number 21, pp. 2125-2228

THIS WEEK IN THE JOURNAL

Article Summaries

PERSPECTIVE

The Partial Death of Abortion Rights
R. A. Charo

The Intimidation of American Physicians — Banning Partial-Birth Abortion
M. F. Greene

ORIGINAL ARTICLES

Aspirin and the Risk of Colorectal Cancer in Relation to the Expression of COX-2
A. T. Chan, S. Ogino, and C. S. Fuchs

Burch Colposuspension versus Fascial Sling to Reduce Urinary Stress Incontinence
M. E. Albo and Others

Advanced Life Support for Out-of-Hospital Respiratory Distress
I. G. Stiell and Others

SPECIAL ARTICLES

Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants
C. S. Phibbs and Others

CLINICAL THERAPEUTICS

Bariatric Surgery for Morbid Obesity
E. J. DeMaria

VIDEOS IN CLINICAL MEDICINE

Central Venous Catheterization
A. S. Graham and Others
IMAGES IN CLINICAL MEDICINE

The Sign of Leser–Trélat
S. Kilickap and B. Yalcin

A Complication of Central Venous Catheterization
S. Nanda and L. Strockoz-Scaff

CASE RECORDS OF THE MASSACHUSETTS GENERAL HOSPITAL

Case 16-2007 — A 61-Year-Old Man with a Mediastinal Mass
C. D. Wright and Others

EDITORIALS

Government in Medicine
J. M. Drazen Aspirin and Colon Cancer — Targeting Prevention?
S. D. Markowitz

Shades of Dry — Curing Urinary Stress Incontinence
K. Strohbehn

HEALTH LAW, ETHICS, AND HUMAN RIGHTS

The Supreme Court and Abortion Rights
G. J. Annas

CLINICAL IMPLICATIONS OF BASIC RESEARCH

A Healthy Tan?
G. Barsh and L. D. Attardi

CORRESPONDENCE

Prevention of Death in COPD

Posaconazole Prophylaxis in Hematologic Cancer

Treatment of Symptomatic Uterine Fibroids

New Treatments for Diabetes

Monoclonal Gammopathy of Undetermined Significance

Pulmonary-Valve Endocarditis
**BOOK REVIEWS**

Natriuretic Peptides: The Hormones of the Heart

Angiogenesis: From Basic Science to Clinical Applications

Congestive Heart Failure

**CORRECTIONS**

Orotracheal Intubation
On April 18, 2007, the Supreme Court signaled a significant change in abortion jurisprudence. It held in *Gonzales v. Carhart* that a federal statute outlawing the use of “partial-birth abortion” is constitutional, even though many members of the medical community believe that the procedure should be available when it is the safest option for terminating a pregnancy. No exception was made for protecting a woman’s health; only a threat to a woman’s life would excuse the use of the procedure. Absent that excuse, a physician who knowingly performs an intact dilation and extraction (D&X) is subject to 2 years in prison, a fine of up to $250,000, and monetary damages for psychological injury to the husband or parents of the pregnant woman.

Ever since the 1973 decision in *Rae v. Wade*, it has been understood that states may regulate pre-viability abortion and outlaw post-viability abortion completely, provided that the rules protect both the life and the health of the pregnant woman. And in the 1992 decision in *Planned Parenthood v. Casey*, the Court reaffirmed this principle, requiring a health exception in bans of post-viability abortion and stating that regulations regarding pre-viability abortion may not impose an “undue burden” on women. This balance reflects the view that a woman’s interest in preserving her own health should be protected more strongly than any state interest in preserving intrauterine life. The latest decision of the Supreme Court alters this balance and holds that requiring women to submit to an unnecessarily dangerous version of an abortion procedure (in cases in which D&X is deemed by a physician to be the safest option) is neither an undue burden on them nor a dereliction of the state’s duty to guard women’s health and personal autonomy. The decision thus opens the door to revisiting any number of state and federal efforts to restrict access to abortion services.

The Supreme Court considered this very same issue 7 years ago in *Stenberg v. Carhart*, when it struck down a similar Nebraska state statute because it did not contain an exception to protect a woman’s health and because its definition of the prohibited procedure was so vague that it could reasonably have been interpreted by doctors to include not only the D&X procedure but other more...
common abortion methods as well. Writing for the 5-to-4 majority in Gonzales v. Carhart, however, Justice Anthony Kennedy distinguished the federal statute from the Nebraska one, noting that it included a more precise definition of the prohibited acts.

A so-called partial-birth abortion, or D&X, involves dilating the cervix, partially extracting the fetus, puncturing the skull while it remains in the uterus, and removing the brain tissue through suction, thus allowing for easy removal of the otherwise intact fetus through the birth canal. In cases in which the procedure is performed, it is usually done late in the second trimester of pregnancy, though in some cases it is used during the third trimester. D&X procedures are rare; in 2000, only 2200 were performed by 31 providers, accounting for 0.17% of all abortions in the United States that year. The more common abortion procedures are suction curettage (used in the first trimester) and dilation and evacuation (D&E), which is the most common procedure in the second trimester. D&E requires dismembering the fetus within the uterus, which poses risks of uterine damage or perforation from surgical instruments and sharp remnants of fetal bone.

Congress passed two statutes banning D&X procedures in the 1990s and another in 2003 that added language to define the prohibited acts more specifically. All three versions allowed for an exception to the ban only in cases in which a woman’s life was in danger. Supporters of the Partial-Birth Abortion Ban Act of 2003 argued that a health exception would most likely be interpreted so broadly by doctors that it would render the legislation meaningless. Similar legislation was passed by state legislatures, and by the late 1990s, 31 states had enacted laws prohibiting partial-birth abortion; only 5 of the bans contain any kind of health exceptions (see map).

When Congress passed its latest ban, it included a lengthy section of “factual findings,” asserting that a “moral, medical and ethical consensus” exists that partial-birth abortion is “a gruesome and inhuman procedure
that is never medically necessary and should be prohibited." The legislation further asserted that D&X is "never necessary to preserve the health of the woman."

It is unclear, however, what degree of deference the Court should grant to such findings when Congress is acting as a source of scientific and medical authority. Legislation often must be passed despite the presence of scientific uncertainty, and much environmental-protection legislation, for example, could be challenged if complete scientific consensus were required before restrictions on industrial pollution could be upheld. But this case is singular in that the Court upheld congressional findings even in instances in which multiple state trial courts had found these same assertions to be based on nonexpert testimony and, in several instances, factually erroneous. The Court then argued that since medical opinion is divided about D&X, Congress has the authority to invade the doctor–patient relationship and substitute blanket legislative judgment for individualized medical judgment concerning the best care for a particular patient. Although regulation of the drugs and devices marketed for use in medical care has long been accepted, legislative restriction of doctors' individual medical judgments is far more contentious. Where governmental involvement in medical decision making is warranted, it is best handled through dispassionate, evidence-based expert reviews. As Kassirer has written, "The data upon which many important medical decisions are based are often contradictory and still in evolution. Legislators do not have the context nor the capacity to weigh medical evidence adequately."

And indeed, the tradition has been to allow the medical profession to define for itself the meaning of "medically indicated."

The decision in Gonzales v. Carhart poses a threat to physicians who perform the still-legal D&E procedure and to those who determine that a D&X is indeed needed to save the life of a pregnant woman. Although the Court emphasized that a physician cannot be convicted unless he or she intentionally violated the statute, questions as to whether a woman's life was in danger and whether the physician's intent was to perform a D&X (as opposed to a D&E) are matters of interpretation. Even if physicians ultimately expect to be exonerated, the mere prospect of being investigated by a possibly hostile prosecutor may well have a chilling effect on their decision making. Certainly, that was the effect on many physicians after the 1975 prosecution of Dr. Kenneth Edelin, who was indicted for manslaughter for performing a second-trimester abortion 2 years after Roe v. Wade.

As to the substantive issue of medical necessity, Greene and Ecker have raised troubling pragmatic questions about what would count as a risk to a woman's life, rather than merely to her health, and where the threshold of "necessity" lies: "Would a procedure that averts a 50 percent risk of death be adequate to qualify as 'necessary to save the life of the mother?'"

In an impassioned dissenting opinion delivered from the bench, Justice Ruth Bader Ginsburg recalled the statement in the 1992 Casey decision that "liberty finds no refuge in a jurisprudence of doubt." The 5-to-4 decision in Gonzales v. Carhart, occasioned by a change in the makeup of the Supreme Court, illustrates how fragile are the constitutional interpretations by which reproductive rights are guaranteed. The prospect of yet further revisions occasioned by future appointments to the Court adds yet another measure of uncertainty beyond the uncertainties that physicians will face when choosing how best to terminate a dangerous pregnancy. And throughout the country, in light of this decision, states will be determining whether their previously unconstitutional bans on partial-birth abortion have now been brought back into effect, many of them incorporating language and proscriptions different from and broader than those of the federal law.

But the greatest uncertainty of all concerns the continued viability of any right to abortion in all but imminently life-threatening situations. The federal statute makes no distinction between pre-viability and post-viability abortions and bans the D&X procedure in both situations, even in cases in which physicians believe that the alternatives are more dangerous to a woman's health. The prospect that a woman's health might be endangered by limiting access to D&X procedures is deemed insufficient to qualify as an "undue burden." Justice Kennedy's majority opinion in Gonzales v. Carhart endorses this conclusion, stating that it is "legitimate" because "a fetus is a living organism within the womb, whether or not it is viable outside the womb" and that "choosing not to prohibit a bru-
The Intimidation of American Physicians — Banning Partial-Birth Abortion

Michael F. Greene, M.D.

A Dutch oncologist was describing to an audience of American physicians in Amsterdam the circumstances under which euthanasia was performed in the Netherlands at a time when the practice was illegal yet widely used. Each act of euthanasia was reported, after the fact, to the local prosecutor, who investigated the case and routinely declined to prosecute any treating physician who had acted transparently and in the best interest of the terminally ill patient. The American physicians were incredulous that their Dutch colleagues were willing to place themselves at risk for criminal prosecution by providing care that might, on later review, be determined to have violated criminal law. The Americans had no confidence that their own judicial system would judge them fairly under similar circumstances, even if they had acted in good faith and in the patient’s best interest.

This lack of confidence that the U.S. judicial system would treat them fairly has cast a pall over those who practice reproductive medicine as they consider the recent decision by the Supreme Court, in Gonzales v. Carhart, to uphold the Partial-Birth Abortion Ban Act of 2003. The ruling creates an intimidating environment surrounding pregnancy terminations at more advanced gestational ages. The decision to pursue a second-trimester abortion is never taken lightly and usually results only after anguished discussions among the patient, her loved ones, and her health care providers. Once the decision has been made to perform a second-trimester surgical abortion, the last thing a provider needs is to have to worry that the procedure could potentially evolve into a criminal act if a fetus in breech presentation should slip out intact through a partially dilated cervix. But this is exactly the situation created by the partial-birth abortion bill.

Defenders of the law point out that its scienter requirement means that physicians can be prosecuted only if it can be demonstrated that the provider “deliberately and intentionally” delivered a living fetus and performed an “overt act” to kill it. But this protection seems fragile to practitioners. In the situation just described, how would the vital status of the partially delivered fetus be determined, and by whom? The only way to complete the delivery through the incompletely dilated cervix may be to reduce the size of the after-coming head. Would any procedure to accomplish that goal be seen as facilitating the
delivery? Or as intentionally killing the fetus? Once the prosecutor knocks on the door, the onus will be on the physician to show that there was no intent to perform a banned procedure. Lack of confidence in the judicial system, physicians may choose to avoid performing second-trimester surgical abortions, thus restricting access to them, perhaps even if the mother's life is in jeopardy.

In the same way that it might be difficult to discern the intent of a physician during the conduct of a pregnancy-termination procedure, it is difficult to know the true intent of the 108th Congress when it passed the partial-birth abortion bill in 2003. Was the intent, as the law claims, simply to ban “a gruesome and inhumane procedure that is never medically necessary”? Or was this law the carefully calculated first step in a larger strategy for the gradual erosion of access to abortion services?

No aspect of medicine seems to attract as much popular and political attention as reproductive medicine. In recent years, our government has restricted women’s options for preventing conception and now for coping with pregnancies that threaten their health or are simply unplanned and undesired. Both health care providers and patients should be alarmed by the current degree of intrusion by our government into the practice of medicine and even more so by the apparent trajectory that it seems poised to follow in the near future.
Aspirin and Colorectal Cancer in Relation to the Expression of COX-2

This study of 636 cases of colorectal cancer culled from two large cohorts of participants who reported data on aspirin use showed that regular use of aspirin reduced the risk of having a colorectal cancer that expressed high amounts of COX-2, an enzyme affected by aspirin.

Burch Colposuspension versus Fascial Sling to Reduce Urinary Stress Incontinence

This multicenter, randomized clinical trial compared two surgical procedures — the Burch colposuspension and the autologous fascial pubovaginal sling — in women with urinary stress incontinence. Success rates (in terms of overall urinary-incontinence measures and stress-incontinence measures specifically) were higher at 2 years for the sling group, but this group also had greater morbidity. These findings inform decision making with respect to surgical treatment of stress incontinence and underscore the importance of surgical randomized trials.

Advanced Life Support for Out-of-Hospital Respiratory Distress

Patients with out-of-hospital respiratory distress who received treatment from emergency medical services personnel with advanced-life-support training had a lower in-hospital mortality than those who received treatment from providers without this training. The difference may be in part attributable to advanced-life-support interventions. Whether the data are sufficient to justify broad implementation of such training is unclear.

NICU Volume and Mortality among Very-Low-Birth-Weight Infants

As compared with mortality in hospitals with high-level, high-volume neonatal intensive care units (NICUs), the mortality among very-low-birth-weight infants was higher at NICUs with lower levels of care and lower patient volumes. Although these data cannot prove cause and effect, the results suggest that increased regional consolidation of perinatal care is feasible and might reduce mortality among very-low-birth-weight infants.

Bariatric Surgery for Morbid Obesity

A 44-year-old morbidly obese woman inquires about bariatric surgery, a treatment option for patients with a body-mass index of 40 or more (or of 35 or more when there are coexisting medical conditions). Bariatric surgery has been shown to result in substantial weight loss and resolution of associated conditions. A successful clinical outcome requires an experienced multidisciplinary management team and an informed patient who will follow a plan of long-term management and self-care.

A Man with a Mediastinal Mass

A 61-year-old man was referred to the thoracic oncology service for management of a thymoma. Six weeks earlier, a mediastinal mass, 4 cm in diameter, had been detected on a computed tomographic scan obtained because of chest pain. A positron-emission tomographic scan showed increased tracer uptake in the area of the lesion. Biopsy specimens showed a World Health Organization type B1 thymoma. Physical examination was normal. A decision on management was made.

Tanning and p53

Exposure of skin to ultraviolet light results in the synthesis of pro-opiomelanocortin (the precursor of melanin). A recent study implicates p53 in this process.
Aspirin and the Risk of Colorectal Cancer in Relation to the Expression of COX-2

Andrew T. Chan, M.D., M.P.H., Shuji Ogino, M.D., Ph.D., and Charles S. Fuchs, M.D., M.P.H.

ABSTRACT

BACKGROUND
Regular use of aspirin reduces the risk of a colorectal neoplasm, but the mechanism by which aspirin affects carcinogenesis in the colon is not well understood.

METHODS
We estimated cyclooxygenase-2 (COX-2) expression by immunohistochemical assay of sections from paraffin-embedded colorectal-cancer specimens from two large cohorts of participants who provided data on aspirin use from a questionnaire every 2 years. We applied Cox regression to a competing-risks analysis to compare the effects of aspirin use on the relative risk of colorectal cancer in relation to the expression of COX-2 in the tumor.

RESULTS
During 2,446,431 person-years of follow-up of 82,911 women and 47,363 men, we found 636 incident colorectal cancers that were accessible for determination of COX-2 expression. Of the tumors, 423 (67%) had moderate or strong COX-2 expression. The effect of aspirin use differed significantly in relation to COX-2 expression (P for heterogeneity=0.02). Regular aspirin use conferred a significant reduction in the risk of colorectal cancers that overexpressed COX-2 (multivariate relative risk, 0.64; 95% confidence interval [CI], 0.52 to 0.78), whereas regular aspirin use had no influence on tumors with weak or absent expression of COX-2 (multivariate relative risk, 0.96; 95% CI, 0.73 to 1.26). The age-standardized incidence rate for cancers that overexpressed COX-2 was 37 per 100,000 person-years among regular aspirin users, as compared with 56 per 100,000 person-years among those who did not use aspirin regularly; in contrast, the rate for cancers with weak or absent COX-2 expression was 27 per 100,000 person-years among regular aspirin users, as compared with 28 per 100,000 person-years among nonregular aspirin users.

CONCLUSIONS
Regular use of aspirin appears to reduce the risk of colorectal cancers that overexpress COX-2 but not the risk of colorectal cancers with weak or absent expression of COX-2.
OBSERVATIONAL STUDIES AND RANDOMIZED INTERVENTION TRIALS HAVE FOUND THAT REGULAR USE OF ASPIRIN REDUCES THE RISK OF COLORECTAL NEOPLASMS. The mechanism by which aspirin influences the risk of colorectal cancer is not well understood. Aspirin inhibits cyclooxygenase, which catalyzes the rate-limiting step in the metabolic conversion of arachidonic acid to prostaglandins and related eicosanoids. One form of cyclooxygenase, termed cyclooxygenase-2 (COX-2), promotes inflammation and cell proliferation, and colorectal cancers often overexpress this enzyme. Randomized trials have demonstrated that selective inhibitors of COX-2 reduce the risk of recurrent adenoma in participants at high risk. However, aspirin and nonsteroidal antiinflammatory drugs (NSAIDs) decrease proliferation and increase apoptosis of colorectal cancer cell lines that have no detectable cyclooxygenase activity. Aspirin has other effects that are unrelated to cyclooxygenase, including inhibition of nuclear factor-κB, induction of apoptosis by activation of p38 kinase, and catabolism of polyamines. If aspirin exerts its effect on the formation of adenomas and cancers by inhibiting COX-2 or its downstream effectors, then the use of aspirin should preferentially reduce the risk of tumors for which growth depends on COX-2 function.

We investigated whether the influence of aspirin on the risk of colorectal cancer varied with the expression of COX-2 in the tumor. For this purpose, we used tumor specimens from two large prospective studies in which an association was found between the regular use of aspirin and a reduced risk of colorectal cancer.

STUDY POPULATION

The Nurses’ Health Study (NHS) was initiated in 1976 when 121,701 U.S. female registered nurses, 30 through 55 years of age, completed a mailed questionnaire. The Health Professionals Follow-up Study (HPFS) was established in 1986 with a parallel cohort of 51,529 U.S. male dentists, optometrists, osteopaths, podiatrists, pharmacists, and veterinarians who were 40 through 75 years of age at entry. The cohorts had a follow-up rate of 92%. We mailed questionnaires every 2 years to obtain updated information and identify new cases of cancer. In 1980, the NHS questionnaire was expanded to include a validated assessment of diet and aspirin use; a similar instrument had been used in the 1986 HPFS questionnaire. The overall response rate for these questionnaires was 90%. The institutional review boards at Brigham and Women’s Hospital and the Harvard School of Public Health approved this study; completion of the questionnaire was considered to imply informed consent.

ASSESSMENT OF MEDICATION USE

Assessment of aspirin use in both the NHS and the HPFS has been described in detail previously. Briefly, in 1980 we asked NHS participants whether they used aspirin, the number of pills taken each week, and the number of years of use. We updated this information every 2 years except for 1986 with specific questions on the number of aspirin tablets taken per week (in categories of number taken). In the 1986 HPFS questionnaire and in questionnaires every 2 years thereafter, we inquired whether the men in the study used aspirin two or more times per week. Beginning in 1992, we also asked these men the average number of tablets taken per week (in categories of number taken). In both cohorts, we specifically inquired about standard-dose (325-mg) aspirin tablets. However, to reflect secular trends in consumption of low-dose (baby) aspirin, the questionnaires after 1992 asked participants to convert four baby aspirin tablets to one adult standard-dose tablet when responding. We did not collect consistent data on use of non-aspirin NSAIDs in either cohort. We did not evaluate the use of selective inhibitors of COX-2 that were introduced in the United States in 1999.

ASSESSMENT OF CASES

We requested permission to acquire medical records and pathology reports from participants who reported colorectal cancer on our biennial questionnaire. We identified deaths from the National Death Index and from next of kin. A study physician who was unaware of information on the participants’ intake of aspirin reviewed medical and pathological records to extract information on histologic types and anatomic locations of the cancers. A single study pathologist reviewed all of the cases that were retrieved for determination of COX-2 expression.
department of treating hospitals, available pathological specimens from participants whom we confirmed had received a diagnosis of colorectal cancer. We obtained specimens from 648 cases (76%) over 16 years of follow-up in the HPFS and 662 cases (58%) over 22 years of follow-up in the NHS. We limited our analysis of the expression of COX-2 to unstained paraffin blocks with amounts of tumor tissue and adjacent mucosa that were sufficient for immunohistochemical analysis (636 cases: 368 from the NHS and 268 from the HPFS). The baseline characteristics of participants with colorectal cancer whose tumors we analyzed were similar to those of participants whose tumors we did not analyze (mean age, 54.4 vs. 54.9 years; nonwhite, 3% vs. 3%; former or current smoker, 60% vs. 58%; mean body-mass index [the weight in kilograms divided by the square of the height in meters], 25.5 vs. 25.3; mean metabolic equivalent [MET; i.e., exercise intensity] score per week, 16.6 vs. 15.3; current multivitamin use, 36% vs. 33%; previous lower gastrointestinal endoscopy, 13% vs. 12%; folate intake, 401 vs. 408 μg per day; calcium intake, 798 vs. 786 mg per day; alcohol intake, 9.9 vs. 10.2 g per day; P>0.25 for all comparisons).

We used a previously described method for COX-2 immunostaining on whole tissue sections or on microassays constituted from specimens.\(^{25,26}\)

We incubated deparaffinized tissue sections with citrate buffer (BioGenex) in a microwave oven for 15 minutes, then cooled them for 40 minutes. Tissue sections were then incubated with 3% hydrogen peroxide for 20 minutes, avidin block for 15 minutes, and biotin block for 15 minutes. A mouse anti–COX-2 antibody (Cayman Chemical) diluted 1:300 in phosphate-buffered saline was applied, and the slides were maintained overnight at 4°C. Next, we applied an antimouse horse IgG antibody (Vector Laboratories) for 20 minutes, followed by an avidin–biotin complex conjugate (Vector Laboratories). The immunohistochemical reaction was revealed by diaminobenzidine and methyl-green counterstain. For each assay run, we included a cancer tissue with COX-2 overexpression and a negative control (normal colonic tissue). We also examined a specimen with known COX-2 overexpression treated with phosphate-buffered saline, but not the anti-COX-2 antibody, as a control for nonspecific binding of the immunohistochemical marker.

A pathologist who was unaware of any data concerning the participants interpreted COX-2 expression, using a standardized grading system (absent, weak, moderate, or strong). The pathologist classified staining of tumor cells as “absent” if COX-2 expression was at the same level of intensity as that in adjacent normal colonic epithelium; weak, moderate, or strong staining indicated progressively increasing degrees of overexpression. A random sample of 108 cancers was reread by a second pathologist unaware of data on the participants; the concordance between the two pathologists was 0.92 (κ = 0.62, P<0.001).\(^{26}\) If the intensity of immunostaining was moderate or strong, tumors were classified as cancers with COX-2 overexpression (COX-2–positive). If immunostaining intensity was weak or absent, tumors were classified as negative for COX-2 overexpression (COX-2–negative) (Fig. 1).

**Statistical Analysis**

At baseline, we excluded men and women who did not complete the baseline dietary questionnaire or medication questions, who recorded implausible dietary or aspirin data (e.g., responded “yes” to use but then recorded use of zero aspirin tablets per week), or reported a baseline history of cancer (except nonmelanoma skin cancer). We also excluded participants with inflammatory bowel disease, familial polyposis, or hereditary nonpolyposis colorectal cancer. After these exclusions, 82,911 women and 47,363 men were eligible for analysis. They accrued follow-up time beginning on the month of return of their baseline questionnaire and ending on the month of diagnosis of colorectal cancer, death from other causes, or June 2002, whichever came first. Participants in whom colorectal cancer was adjudicated as the cause of death by study physicians after review of medical records were included as cases. Data from cases of colorectal cancer for which we were unable to assay COX-2 expression in the tumor were censored from the analysis at the date of diagnosis.

To reduce within-person variation and to better estimate long-term intake, we used the cumulative average intake of aspirin as reported on all available questionnaires up to the start of each 2-year follow-up interval.\(^{27}\) In the NHS, women who reported taking two or more standard (325-mg) aspirin tablets per week were classified as regular users, whereas those who reported less aspirin use were classified as nonusers. We also estimated the duration of the regular use of aspirin by NHS participants from the number of years of regular use...
reported in 1980, with updating of this variable every 2 years.\textsuperscript{19,20,23} In the HPFS, men who reported using aspirin at least two times per week were classified as regular users, whereas those who did not report aspirin use at least two times per week were classified as nonusers.\textsuperscript{18} For our main analyses, we pooled data from both cohorts. We conducted tests of heterogeneity using the Q statistic before pooling.\textsuperscript{28} In this analysis, we observed no heterogeneity between the cohorts regarding the association of regular aspirin use and risk of colorectal cancer ($P=0.33$ for Cochran’s Q test).

We calculated the incidence rates of colorectal cancer for participants in a specific category of aspirin use by dividing the number of incident cases by the number of person-years of follow-up. We calculated relative risks by dividing the incidence rate in one category by the incidence rate in the reference category; we used Cox proportional-hazards modeling to control for multiple variables simultaneously and to calculate 95% confidence intervals. Multivariate relative risks are adjusted for age, sex, smoking before the age of 30 years (0, 1 to 4, 5 to 10, 11 to 15, or more than 15 pack-years); body-mass index (in quintiles); regular vigorous exercise (in quintiles of MET task score per week); colorectal cancer in a parent or sibling (yes or no); previous lower gastrointestinal endoscopy (yes or no); history of polyps (yes or no); current multivitamin use (yes or no); consumption of beef, pork, or lamb as a main dish (0 to 3 times per month, 1 time per week, 2 to 4 times per week, or 5 or more times per week); alcohol consumption (0, 0.1 to 4.9, 5.0 to 14.9, or 15 g or more per day); and energy-adjusted folate and calcium intake (in quintiles). For the analyses restricted to women, the multivariate models additionally adjusted for postmenopausal hormone-replacement therapy (premenopausal women and postmenopausal women who never used, formerly used, or were...
Aspirin and Colorectal Cancer in Relation to COX-2 Expression

RESULTS

Among the 82,911 eligible women and 47,363 eligible men at the time of the baseline questionnaire, participants reporting the regular use of aspirin (taking two or more standard aspirin tablets per week or using aspirin at least two times per week) were older, were slightly less physically active, were more likely to be current or former smokers, were more likely to use multivitamins regularly, were more likely to use postmenopausal hormones (among women), consumed more alcohol, and consumed more folate (P<0.001 for all unadjusted comparisons) (Table 1). On the basis of the updated data on aspirin use we obtained on biennial questionnaires, we observed that patterns of aspirin use changed over time. Averaged over the total number of person-years of follow-up (2,446,431), the difference in mean age between regular aspirin users and nonusers gradually increased (59.0 years for regular aspirin users and 56.5 years for nonusers, P<0.001). For this analysis, we identified 636 incident cases of colorectal cancer among users and nonusers of aspirin that were available for analysis of COX-2 expression (Table 1). Among these 636 tumors, 423 (67%) had moderate or strong COX-2 expression (i.e., were COX-2–positive), whereas 213 (33%) had weak or absent COX-2 expression (i.e., were COX-2–negative).

As in our previous studies,18–20 we observed in both cohorts a significantly lower risk of colorectal cancer among regular aspirin users than among participants who did not regularly use aspirin, after controlling for other known or suspected colorectal cancer risk factors (Table 2). For the combined cohorts, regular aspirin use was associated with a multivariate relative risk of colorectal cancer of 0.73 (95% confidence interval [CI], 0.62 to 0.86) for cases with tissue available for COX-2 analysis. Inclusion of all documented cases of colorectal cancer that were found during follow-up, irrespective of the availability of tissue for analysis of COX-2 expression, did not materially alter these results (multivariate relative risk, 0.76; 95% CI, 0.70 to 0.84 for 1142 cases in the NHS and 853 cases in the HPFS; P<0.001). Regular aspirin use was associated with a multivariate relative risk of 0.78 (95% CI, 0.69 to 0.87; P<0.001) among participants with colorectal cancer in whom COX-2 expression was not assayed.

We evaluated the influence of aspirin on the risk of colorectal cancer according to overexpression of COX-2 in the tumor (Table 2). The benefit of regular aspirin use appeared to be confined to cancers with COX-2 overexpression (multivariate relative risk, 0.64; 95% CI, 0.52 to 0.78). In contrast, aspirin use did not appear to be associated with the risk of colorectal cancer with weak or absent COX-2 expression (multivariate relative risk, 0.96; 95% CI, 0.73 to 1.26). A test for heterogeneity for the association of regular aspirin use with COX-2–positive or COX-2–negative tumors found a statistically significant association (P for heterogeneity=0.02).

Table 3 shows the association between the duration of regular use of aspirin and the risk of a COX-2–positive or COX-2–negative colorectal cancer. A statistically significant reduction in the number of cases of COX-2–positive cancer was not evident until aspirin had been used regularly for more than 10 years (multivariate relative risk, 0.59; 95% CI, 0.42 to 0.82; P for trend, <0.001); there was no statistically significant reduction in the number of COX-2–negative cancers with increasing duration of aspirin use (P for trend=0.25).

Table 4 shows the association between the dose of aspirin (tablets per week) and the risk of colorectal cancer. A reduction in the multivariate relative risk of colorectal cancer became apparent with an intake of more than five tablets per week, and the relative risk was further reduced as the...
Table 1. Baseline Characteristics of the Study Cohort.†

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Women</th>
<th></th>
<th>Men</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonregular Users of Aspirin</td>
<td>Regular Users of Aspirin</td>
<td>Nonregular Users of Aspirin</td>
<td>Regular Users of Aspirin</td>
<td>Nonregular Users of Aspirin</td>
<td>Regular Users of Aspirin</td>
</tr>
<tr>
<td>Median age (yr)</td>
<td>46</td>
<td>47</td>
<td>53</td>
<td>56</td>
<td>48</td>
<td>49</td>
</tr>
<tr>
<td>Race (%)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonwhite</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>White</td>
<td>97</td>
<td>98</td>
<td>94</td>
<td>96</td>
<td>96</td>
<td>98</td>
</tr>
<tr>
<td>Former or current smoker %</td>
<td>56</td>
<td>58</td>
<td>49</td>
<td>58</td>
<td>53</td>
<td>58</td>
</tr>
<tr>
<td>No. of pack-years‡</td>
<td>20.4</td>
<td>20.8</td>
<td>24.2</td>
<td>25.6</td>
<td>21.8</td>
<td>22.5</td>
</tr>
<tr>
<td>Body-mass index§</td>
<td>24.1</td>
<td>24.6</td>
<td>25.5</td>
<td>25.7</td>
<td>24.6</td>
<td>24.9</td>
</tr>
<tr>
<td>Physical activity (METs/wk)¶</td>
<td>14.3</td>
<td>13.5</td>
<td>21.0</td>
<td>20.8</td>
<td>17.3</td>
<td>16.0</td>
</tr>
<tr>
<td>Postmenopausal women (%)‖</td>
<td>43</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never used hormones (%)</td>
<td>62</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formerly used hormones (%)</td>
<td>19</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently using hormones (%)</td>
<td>19</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current multivitamin use (%)</td>
<td>31</td>
<td>38</td>
<td>39</td>
<td>49</td>
<td>34</td>
<td>42</td>
</tr>
<tr>
<td>Previous lower gastrointestinal endoscopy (%)</td>
<td>2</td>
<td>2</td>
<td>25</td>
<td>27</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Colorectal cancer in a parent or sibling (%)</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Dietary intake**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beef, pork, or lamb as a main dish (servings/wk)</td>
<td>2.5</td>
<td>2.6</td>
<td>1.8</td>
<td>1.8</td>
<td>2.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Folate (µg/day)</td>
<td>360</td>
<td>371</td>
<td>473</td>
<td>498</td>
<td>404</td>
<td>411</td>
</tr>
<tr>
<td>Alcohol (g/day)</td>
<td>6.1</td>
<td>6.8</td>
<td>10.8</td>
<td>12.6</td>
<td>7.9</td>
<td>8.7</td>
</tr>
<tr>
<td>Calcium (mg/day)</td>
<td>736</td>
<td>726</td>
<td>890</td>
<td>917</td>
<td>795</td>
<td>787</td>
</tr>
</tbody>
</table>

* Data are from the baseline questionnaire administered in 1980 to women enrolled in the Nurses’ Health Study (NHS) and in 1986 to men enrolled in the Health Professionals Follow-up Study (HPFS). In the NHS, regular aspirin use was based on previously described categorization as the consumption of at least two 325-mg tablets per week. Nonregular use was defined as the consumption of fewer than two tablets per week. In the HPFS, regular aspirin use was based on previously described categorization as the consumption of aspirin at least two times per week. Nonregular use was defined as the consumption of aspirin fewer than two times per week. All values, other than age, have been directly standardized according to the age distribution of the cohort. Values are means unless otherwise indicated.

† Race was self-assessed by the participants on questionnaires.

‡ Pack-years were calculated for former and current smokers only.

§ The body-mass index is the weight in kilograms divided by the square of the height in meters.

¶ MET denotes metabolic equivalent, calculated according to the frequency of a range of physical activities (e.g., jogging) in 1986 for both women and men.

‖ Hormones are defined as postmenopausal estrogen or estrogen plus progesterone preparations. The percentages of those never using, formerly using, and currently using hormones were calculated for postmenopausal women only.

** Nutrient values (folate and calcium) represent the mean of energy-adjusted intakes.
Table 2. Relative Risk of Colorectal Cancer in Relation to COX-2 Expression and Regular Aspirin Use.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Women</th>
<th>Men</th>
<th>Total†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonregular Users of Aspirin</td>
<td>Regular Users of Aspirin</td>
<td>Nonregular Users of Aspirin</td>
</tr>
<tr>
<td>All colorectal cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>225/1,037,122</td>
<td>143/704,361</td>
<td>161/402,337</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>0.80 (0.65–0.99)</td>
<td>1.0</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0</td>
<td>0.80 (0.65–0.99)</td>
<td>1.0</td>
</tr>
<tr>
<td>COX-2–positive cancer§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>154/1,037,181</td>
<td>88/704,404</td>
<td>117/402,372</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>0.72 (0.55–0.94)</td>
<td>1.0</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0</td>
<td>0.72 (0.56–0.94)</td>
<td>1.0</td>
</tr>
<tr>
<td>COX-2–negative cancer‖</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>71/1,037,259</td>
<td>55/704,435</td>
<td>44/402,442</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>0.98 (0.69–1.39)</td>
<td>1.0</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0</td>
<td>0.98 (0.69–1.40)</td>
<td>1.0</td>
</tr>
</tbody>
</table>

* The women were participants in the Nurses’ Health Study (NHS)19,20 enrolled in 1980 and followed until 2002. The men were participants in the Health Professionals Follow-up Study (HPFS)18 enrolled in 1986 and followed until 2002. In the NHS, regular aspirin use was based on previously described categorization as the consumption of at least two 325-mg tablets per week. Nonregular use was defined as the consumption of fewer than two tablets per week. In the HPFS, regular aspirin use was based on previously described categorization as the consumption of aspirin at least two times per week. Nonregular use was defined as the consumption of aspirin fewer than two times per week. Relative risks are for regular users as compared with nonregular users.

† Pooled data are from NHS and HPFS (P = 0.33 with the use of Cochran’s Q test for heterogeneity).

‡ Multivariate relative risks are adjusted for age, sex, smoking before 30 years of age (0 pack-years, 1 to 4 pack-years, 5 to 10 pack-years, 11 to 15 pack-years, or >15 pack-years); body-mass index (in quintiles); regular vigorous exercise (in quintiles of metabolic equivalent [MET] task score per week); colorectal cancer in a parent or sibling (yes or no); previous endoscopy (yes or no); history of polyps (yes or no); current multivitamin use (yes or no); consumption of beef, pork, or lamb as a main dish (0 to 3 times per month, once a week, 2 to 4 times per week, or ≥5 times per week); alcohol consumption (0 g, 0.1 to 4.9 g, 5.0 to 14.9 g, or ≥15.0 g per day); and energy-adjusted folate and calcium intake (in quintiles). For the analyses restricted to women, the multivariate models were also adjusted for postmenopausal hormone-replacement therapy (premenopausal women and postmenopausal women who never used, formerly used, or were currently using hormone-replacement therapy) and did not include sex. For the analyses restricted to men, the multivariate models did not include sex.

§ Cancers with immunohistochemical COX-2 staining of moderate to strong intensity are classified as COX-2–positive cancers.

¶ The P value for heterogeneity of the association of regular aspirin use with COX-2–positive cancer and of regular aspirin use with COX-2–negative cancer is 0.02 (χ² = 5.7, 1 degree of freedom).

‖ Cancers with no immunohistochemical COX-2 staining or with staining of weak intensity are classified as COX-2–negative cancers.
Table 3. Relative Risk of Colorectal Cancer in Relation to COX-2 Expression and Duration of Regular Aspirin Use.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Years of Regular Aspirin Use</th>
<th>P Value†‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1–5</td>
</tr>
<tr>
<td>All colorectal cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>244/1,012,439</td>
<td>139/481,563</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>0.98 (0.80–1.12)</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)</td>
<td>1.0</td>
<td>0.92 (0.74–1.13)</td>
</tr>
<tr>
<td>COX-2–positive cancer§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>169/1,012,502</td>
<td>94/481,604</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>0.96 (0.74–1.24)</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)</td>
<td>1.0</td>
<td>0.90 (0.69–1.16)</td>
</tr>
<tr>
<td>COX-2–negative cancer¶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. of person-yr</td>
<td>75/1,012,593</td>
<td>45/481,647</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0</td>
<td>1.03 (0.71–1.50)</td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)</td>
<td>1.0</td>
<td>0.96 (0.66–1.40)</td>
</tr>
</tbody>
</table>

* In the Nurses’ Health Study (NHS), regular aspirin use was based on previously described categorization as the consumption of at least two 325-mg tablets per week. Nonregular use was defined as the consumption of fewer than two tablets per week. In the Health Professionals Follow-up Study (HPFS), regular aspirin use was based on previously described categorization as the consumption of aspirin at least two times per week. Nonregular use was defined as the consumption of aspirin fewer than two times per week. Relative risks are for regular users as compared with nonregular users.

† P values are for linear trend.

‡ Multivariate relative risks are adjusted for age, sex, smoking before 30 years of age (0 pack-years, 1 to 4 pack-years, 5 to 10 pack-years, 11 to 15 pack-years, or >15 pack-years); body-mass index (in quintiles); regular vigorous exercise (in quintiles of metabolic equivalent [MET] task score per week); colorectal cancer in a parent or sibling (yes or no); previous endoscopy (yes or no); history of polyps (yes or no); current multivitamin use (yes or no); consumption of beef, pork, or lamb as a main dish (0 to 3 times per month, once a week, 2 to 4 times per week, or ≥5 times per week); alcohol consumption (0 g, 0.1 to 4.9 g, 5.0 to 14.9 g, or ≥15.0 g per day); and energy-adjusted folate and calcium intake (in quintiles). For the analyses restricted to women, the multivariate models were also adjusted for postmenopausal hormone-replacement therapy (premenopausal women and postmenopausal women who never used, formerly used, or were currently using hormone-replacement therapy) and did not include sex. For the analyses restricted to men, the multivariate models did not include sex.

§ Cancers with immunohistochemical COX-2 staining of moderate to strong intensity are classified as COX-2–positive cancers.

¶ Cancers with no immunohistochemical COX-2 staining or with staining of weak intensity are classified as COX-2–negative cancers.
Aspirin and Colorectal Cancer in Relation to COX-2 Expression

Aspirin dose is classified according to the number of standard 325-mg tablets taken per week. In the Health Professionals Follow-up Study (HPFS), data on the number of tablets of aspirin taken per week were not collected until 1992. Thus, this analysis includes 368 incident cases from the Nurses’ Health Study (NHS) from 1980 through 2002 and 184 incident cases from the HPFS from 1992 through 2002. Relative risks are calculated for each dose category as compared with no aspirin use.

<table>
<thead>
<tr>
<th>Cancer</th>
<th>No. of Aspirin Tablets per Week</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>All colorectal cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>126/523,876</td>
<td>0.002</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.75–1.18)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.74–1.17)</td>
<td>0.001</td>
</tr>
<tr>
<td>COX-2–positive cancer§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>82/523,912</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.78–1.36)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.77–1.34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>COX-2–negative cancer¶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>44/523,950</td>
<td>0.58</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.52–1.16)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.52–1.14)</td>
<td>0.52</td>
</tr>
</tbody>
</table>

‡ Multivariate relative risks are adjusted for age, sex, smoking before 30 years of age (0 pack-years, 1 to 4 pack-years, 5 to 10 pack-years, 11 to 15 pack-years, or >15 pack-years); body-mass index (in quintiles); regular vigorous exercise (in quintiles of metabolic equivalent [MET] task score per week); colorectal cancer in a parent or sibling (yes or no); previous endoscopy (yes or no); history of polyps (yes or no); current multivitamin use (yes or no); consumption of beef, pork, or lamb as a main dish (0 to 3 times per month, once a week, 2 to 4 times per week, or ≥5 times per week); alcohol consumption (0 g, 0.1 to 4.9 g, 5.0 to 14.9 g, or ≥15.0 g per day); and energy-adjusted folate and calcium intake (in quintiles). For the analyses restricted to women, the multivariate models additionally adjusted for postmenopausal hormone-replacement therapy (premenopausal women and postmenopausal women who never used, formerly used, or were currently using hormone-replacement therapy) and did not include sex. For the analyses restricted to men, the multivariate models did not include sex.

§ Cancers with immunohistochemical COX-2 staining of moderate to strong intensity are classified as COX-2–positive cancers.

¶ Cancers with no immunohistochemical COX-2 staining or with staining of weak intensity are classified as COX-2–negative cancers.

We found that participants in two large prospective cohorts who reported long-term, regular aspirin use (at least two standard tablets per week or use of aspirin at least two times per week) had a significant reduction in the relative risk of colorec-

Discussion

The age-standardized incidence rate of COX-2–positive tumors was 37 per 100,000 person-years for regular aspirin users as compared with 56 per 100,000 person-years for nonusers. In contrast, the age-standardized incidence rate of COX-2–negative tumors was 27 per 100,000 person-years for regular aspirin users as compared with 28 per 100,000 person-years for nonusers.

Table 4. Relative Risk of Colorectal Cancer in Relation to COX-2 Expression and Aspirin Dose.*

<table>
<thead>
<tr>
<th>Cancer</th>
<th>No. of Aspirin Tablets per Week</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>All colorectal cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>126/523,876</td>
<td>0.002</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.75–1.18)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.74–1.17)</td>
<td>0.001</td>
</tr>
<tr>
<td>COX-2–positive cancer§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>82/523,912</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.78–1.36)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.77–1.34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>COX-2–negative cancer¶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of cases/no. person-yr</td>
<td>44/523,950</td>
<td>0.58</td>
</tr>
<tr>
<td>Age-adjusted relative risk (95% CI)</td>
<td>1.0 (0.52–1.16)</td>
<td></td>
</tr>
<tr>
<td>Multivariate relative risk (95% CI)‡</td>
<td>1.0 (0.52–1.14)</td>
<td>0.52</td>
</tr>
</tbody>
</table>

* Aspirin dose is classified according to the number of standard 325-mg tablets taken per week. In the Health Professionals Follow-up Study (HPFS), data on the number of tablets of aspirin taken per week were not collected until 1992. Thus, this analysis includes 368 incident cases from the Nurses’ Health Study (NHS) from 1980 through 2002 and 184 incident cases from the HPFS from 1992 through 2002. Relative risks are calculated for each dose category as compared with no aspirin use.

† P values are for linear trend.

‡ Multivariate relative risks are adjusted for age, sex, smoking before 30 years of age (0 pack-years, 1 to 4 pack-years, 5 to 10 pack-years, 11 to 15 pack-years, or >15 pack-years); body-mass index (in quintiles); regular vigorous exercise (in quintiles of metabolic equivalent [MET] task score per week); colorectal cancer in a parent or sibling (yes or no); previous endoscopy (yes or no); history of polyps (yes or no); current multivitamin use (yes or no); consumption of beef, pork, or lamb as a main dish (0 to 3 times per month, once a week, 2 to 4 times per week, or ≥5 times per week); alcohol consumption (0 g, 0.1 to 4.9 g, 5.0 to 14.9 g, or ≥15.0 g per day); and energy-adjusted folate and calcium intake (in quintiles). For the analyses restricted to women, the multivariate models additionally adjusted for postmenopausal hormone-replacement therapy (premenopausal women and postmenopausal women who never used, formerly used, or were currently using hormone-replacement therapy) and did not include sex. For the analyses restricted to men, the multivariate models did not include sex.

§ Cancers with immunohistochemical COX-2 staining of moderate to strong intensity are classified as COX-2–positive cancers.

¶ Cancers with no immunohistochemical COX-2 staining or with staining of weak intensity are classified as COX-2–negative cancers.
tential cancer, as compared with participants who reported less use of aspirin. This reduction in risk was due almost entirely to a reduction in the number of COX-2–positive colorectal cancers among regular aspirin users. In contrast, the difference in the incidence of COX-2–negative colorectal cancers between regular users and nonusers of aspirin was not statistically significant. A reduction of the risk of COX-2–positive tumors was also found with increasing aspirin dose and increasing duration of use. No significant dose or duration effect was observed for COX-2–negative tumors.

These results are consistent with a cross-sectional case–control study in which the association of NSAIDs with a reduced risk of adenomatous polyps was strongest in cases with high expression of COX-2 messenger RNA. Our finding that the association with a reduced risk of colorectal cancer is most apparent at intakes of more than five aspirin tablets per day is consistent with the results of studies in which higher doses of aspirin were required to inhibit COX-2 than to inhibit COX-1. Moreover, the association with a reduction in cancer risk was found only after several years of aspirin use, a finding that suggests an effect of aspirin on the early stages of colorectal adenoma or cancer.

COX-2 is progressively overexpressed during the stepwise sequence from adenoma to carcinoma, and randomized, placebo-controlled trials have shown that selective COX-2 inhibitors prevent recurrence of adenoma among patients with a history of adenoma or familial polyposis. Our data suggest that the anticancer benefit of aspirin is mediated, at least in part, by inhibition of COX-2. Experimental studies have shown that aspirin and NSAIDs, especially at high doses, have a range of other potentially antineoplastic actions; these results indicate that further work is needed to elucidate the effects of these agents and COX-2 (or its downstream effectors) on the development of colorectal cancer.

Our study has several strengths. First, because we collected detailed, updated information on aspirin use during 22 years of follow-up, we were able to evaluate the effects of long-term use across a broad range of intakes. Second, we obtained aspirin data prospectively, before the diagnosis of colorectal cancer. Any errors in recall would have tended to attenuate rather than exaggerate true associations. Third, because the participants were all health professionals, the accuracy of self-reported aspirin use is likely to be high. Fourth, we collected detailed data on potential confounders and had a high follow-up rate (92% of participants returned questionnaires or were identified as deceased). Finally, our results were remarkably consistent between the two independent cohorts.

There are several limitations to our study. The study was observational, and aspirin use was self-selected; thus, among aspirin users and nonusers, there were small, albeit statistically significant, differences in risk factors for colorectal cancer, including smoking history, physical activity, and use of multivitamins. However, adjustment for these characteristics, as well as for a wide range of other potential risk factors, had minimal influence on our findings, suggesting little potential for residual or uncontrolled confounding. We did not have sufficient data on use of nonaspirin NSAIDs in both cohorts to reliably examine their use in relation to COX-2–expressing tumors. However, it is improbable that use of nonaspirin NSAIDs influenced our findings with aspirin; our previous study did not demonstrate any modification by NSAID use of the reduced risk of colorectal cancer observed with aspirin.

We were unable to obtain tumor tissue from all cases of confirmed colorectal cancer detected in the two cohorts, but the risk factors in these cases did not appreciably differ from those in cases with tumor tissue available.

Another limitation of our study is the lack of a widely accepted, standardized classification scheme for COX-2 expression in colorectal cancer. Nevertheless, previous studies have demonstrated that the results of Western and Northern blot analyses correlate well with immunohistochemical grading of COX-2. With our classification of COX-2 expression, we found, in our overall population, a proportion of tumors that overexpressed COX-2 similar to those found by other investigators. In validation studies of the central, blinded review of tumor specimens, we observed substantial interobserver agreement (92%). Any significant misclassification of COX-2 overexpression would be expected to bias our results toward finding no significant differences in the effect of aspirin on tumors in relation to COX-2 expression.

Our results support the importance of continued investigation into COX-2 and related pathways for the development of new treatments and the potential use of COX-2 as a molecular marker for...
tailoring chemoprevention in participants with a history of colorectal cancer.

Supported by grants from the National Cancer Institute (CA87969, CA55075), the National Institutes of Health, the Marshall S. Kates Memorial Fund, the Bennett Family Fund for Targeted Therapies Research, and the Entertainment Industry Foundation National Colorectal Cancer Research Alliance. Dr. Chan is a recipient of the American Gastroenterological Association/Foundation for Digestive Health and Nutrition Research Scholar Award and a career development award from the National Cancer Institute (CA10741). The National Cancer Institute, the National Institutes of Health, the Marshall S. Kates Memorial Fund, the Bennett Family Fund for Targeted Therapies Research, the Entertainment Industry Foundation National Colorectal Cancer Research Alliance, the American Gastroenterological Association, and the Foundation for Digestive Health and Nutrition had no role in the collection, management, analysis, or interpretation of the data and had no role in the preparation, review, or approval of the manuscript.

Dr. Chan reports receiving a career development award from the Glaxo Institute for Digestive Health for an unrelated study. No other potential conflict of interest relevant to this article was reported.

We thank Mr. Gregory Kirkner for assistance in accessing tumor specimens, Ms. Mami Cantor for assistance in immunohistochemical assays, Dr. Reiko Dehari for interobserver variability studies, Dr. Eva Schernhammer for statistical expertise, and Dr. Edward Giovannucci and Dr. Walter Willett for scientific insights.

REFERENCES

32. He TC, Chan TA, Vogelstein B, Kinzler KW. PPARdelta is an APC-regulated target
of nonsteroidal anti-inflammatory drugs.


Copyright © 2007 Massachusetts Medical Society.
Burch Colposuspension versus Fascial Sling to Reduce Urinary Stress Incontinence


From the University of California, San Diego, San Diego (M.E.A., C.N., S.M.); the University of Alabama at Birmingham, Birmingham (H.E.R., L.K.L., R.E.V.); Loyola University Medical Center, Maywood, IL (L.B., M.F., K.K.); the University of Utah Health Sciences Center, Salt Lake City (P.N.); the University of Texas Health Sciences Center, San Antonio (S.R.K.); the University of Texas Southwestern, Dallas (P.E.Z., G.E.L.); the University of Maryland, Baltimore (T.C.C., H.W.J.); Magee Women’s Hospital, University of Pittsburgh, Pittsburgh (H.Z., W.L., P.M.); Beaumont Hospital Medical Center, Royal Oak, MI (A.C.D., L.S.); New England Research Institutes, Watertown, MA (S.T., A.M.S., K.J.D.); Oakwood Hospital, Dearborn, MI (V.M.); the National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD (J.W.K., L.M.N.); and the University of Virginia Health Systems, Charlottesville (W.S.). Address reprint requests to Dr. Albo at the Division of Urology, University of California, San Diego Medical Center, 200 W. Arbor Dr., San Diego, CA 92103-8897.

The members of the Urinary Incontinence Treatment Network are listed in the Appendix.

BACKGROUND

Many surgical procedures are available for women with urinary stress incontinence, yet few randomized clinical trials have been conducted to provide a basis for treatment recommendations.

METHODS

We performed a multicenter, randomized clinical trial comparing two procedures — the pubovaginal sling, using autologous rectus fascia, and the Burch colposuspension — among women with stress incontinence. Women were eligible for the study if they had predominant symptoms associated with the condition, a positive stress test, and urethral hypermobility. The primary outcomes were success in terms of overall urinary-incontinence measures, which required a negative pad test, no urinary incontinence (as recorded in a 3-day diary), a negative cough and Valsalva stress test, no self-reported symptoms, and no retreatment for the condition, and success in terms of measures of stress incontinence specifically, which required only the latter three criteria. We also assessed postoperative urge incontinence, voiding dysfunction, and adverse events.

RESULTS

A total of 655 women were randomly assigned to study groups: 326 to undergo the sling procedure and 329 to undergo the Burch procedure; 520 women (79%) completed the outcome assessment. At 24 months, success rates were higher for women who underwent the sling procedure than for those who underwent the Burch procedure, for both the overall category of success (47% vs. 38%, P=0.01) and the category specific to stress incontinence (66% vs. 49%, P<0.001). However, more women who underwent the sling procedure had urinary tract infections, difficulty voiding, and postoperative urge incontinence.

CONCLUSIONS

The autologous fascial sling results in a higher rate of successful treatment of stress incontinence but also greater morbidity than the Burch colposuspension. (ClinicalTrials.gov number, NCT00064662.)
U RINARY INCONTINENCE AFFECTS AN estimated 15 to 50% of women, resulting in a significant medical, social, and economic burden. In 1995 dollars, the annual direct costs of incontinence in the United States were estimated to be more than $16 billion. Among women with incontinence, 50 to 80% are identified as having stress incontinence, or involuntary leakage of urine resulting from physical exertion or sneezing and coughing. Although the initial treatment of stress incontinence is often nonsurgical (behavioral therapy, pelvic-floor exercises, or incontinence devices), surgical treatment is considered for patients who are bothered by persistent symptoms. An estimated 4 to 10% of women in the United States undergo surgery intended to restore continence, and this rate has increased steadily during the past 20 years.

Many surgical procedures have been described for women with stress incontinence, yet few randomized clinical trials have been conducted to provide a basis for treatment recommendations. The fascial-sling procedure and Burch colposuspension are two well-established procedures with reported cure rates of 70 to 85% at 5 to 8 years. In the Burch modified colposuspension, the anterior vaginal wall is suspended (at the level of the bladder neck) with permanent sutures tied to the iliopubic ligament (Fig. 1A). In the autologous sling procedure, a harvested strip of rectus fascia is placed transvaginally at the level of the proximal urethra. The fascial strip is secured superiorly to the rectus fascia with permanent sutures (Fig. 1B). Although it has been suggested that the sling procedure may result in higher cure rates, this advantage may be offset by increased obstructive complications, such as voiding dysfunction and urge incontinence. We conducted a multicenter, randomized surgical trial, the Stress Incontinence Surgical Treatment Efficacy Trial, to compare the efficacy and safety of the sling and Burch procedures 24 months after surgery.

Methods

Patients and Study Design
Women who were planning to undergo stress-incontinence surgery were invited to participate in the trial. Eligibility requirements included documented pure or predominant symptoms of stress incontinence for at least 3 months and a positive standardized urinary stress test.

Details of the study methods have been published previously. All study procedures were approved by the institutional review board at each participating clinical center, and written consent was obtained from all women before enrollment. Randomization was performed in the operating room after anesthesia induction with the use of a permuted-block randomization schedule with stratification according to clinical site. The patients were aware of study-group assignments postoperatively. An independent data and safety monitoring board oversaw the progress, interim results, and safety of the study.

Formal interim time-to-event analyses of the primary outcome of overall success were planned for three time points, with the use of an O’Brien–Fleming stopping boundary, and were conducted when 19%, 44%, and 76% of failures had occurred. Although the test statistic at the third analysis crossed the stopping boundary in favor of the sling procedure, the protocol did not require stopping the trial when the boundary was crossed, and the data and safety monitoring board recommended that the study continue. No adjustment was made for these analyses.

Definitions of clinical terms, urodynamic nomenclature, and methods of evaluation of patients were uniform across all sites and in accordance with recommendations from the standardization committees of the International Continence Society. Key elements of the two surgical procedures were standardized among all participating surgeons and included the use of preoperative antibiotics, skin-incision length, number and type of Burch sutures, fascial-sling length and width, and cystoscopic evaluation of the bladder. Because these procedures are frequently performed in conjunction with surgery for pelvic prolapse, abdominal and vaginal approaches for both pelvic prolapse repair and hysterectomy were permitted. However, surgeons were required to declare before randomization which concomitant procedures would be performed.

The two primary outcomes were composite measures of success in terms of overall urinary-incontinence measures and of stress-incontinence measures specifically. Overall treatment success was defined as no self-reported symptoms of urinary incontinence, an increase of less than 15 g in pad weight during a 24-hour pad test, no incontinence episodes recorded in a 3-day diary, a negative urinary stress test (no leakage noted on...
Figure 1. Burch Modified Colposuspension and Autologous Sling Procedure.

In the Burch procedure (Panel A), permanent sutures are placed in the anterior vaginal wall at the level of the bladder neck and proximal urethra and are then sutured to the iliopectineal ligament. In the autologous sling procedure (Panel B), a strip of rectus fascia is harvested, and permanent sutures are placed at its two ends. The sling is placed beneath the proximal urethra through a vaginal incision. The two ends of the sling are passed behind the pubic bone to the anterior abdominal wall, where they are secured, either to each other or to the rectus fascia.
examination during cough and Valsalva maneuvers at a standardized bladder volume of 300 ml), and no retreatment for urinary incontinence (including behavioral, pharmacologic, and surgical therapies). Since the study surgeries are intended to correct symptoms of stress incontinence without necessarily improving concomitant urge incontinence and the voiding diary and pad test do not differentiate between urge-incontinence and stress-incontinence events, the definition of success specific to stress incontinence was limited to no self-reported symptoms of stress incontinence, a negative stress test, and no retreatment for stress incontinence.

Data were collected preoperatively and postoperatively at 6 weeks and at 3, 6, 12, 18, and 24 months by means of interview and clinical examination. Baseline measures included sociodemographic characteristics; risk factors for urinary incontinence, including a high body-mass index, a history of vaginal childbirth, and previous surgery for urinary incontinence; quality of life specific to urinary incontinence;27; clinical characteristics of urinary incontinence, including current behavioral or pharmacologic therapy, self-reported urinary-incontinence symptoms on a validated questionnaire distinguishing stress leakage from urge leakage,18 quantity of urine leakage on a pad test,19 and the number of incontinence episodes as recorded in a 3-day voiding diary;19; findings on physical examination, including urethral hypermobility as measured by the Q-tip test21 and pelvic-organ prolapse;22; and urodynamic evaluation, including the presence of urodynamic stress incontinence and detrusor-overactivity incontinence.

The principal investigator at each site reported adverse events to the adverse-events committee, which comprised four investigators who were unaware of site-specific information. In certain cases, the descriptive details of the adverse event may have made it possible to discern the randomized surgical procedure. All adverse events were assigned a severity code according to a modified version of the classification system developed by Dindo and colleagues.23 This system, which has been validated for use among surgical patients, classifies the severity of an event into one of four levels on the basis of the clinical measures taken to treat that event.

Postoperative urge incontinence was defined as treatment of clinically diagnosed new-onset or persistent urge incontinence after the 6-week follow-up visit. Adequacy of voiding was assessed and categorized dichotomously at hospital discharge and again 6 weeks after surgery. Voiding dysfunction was defined by the need for surgical revision to facilitate bladder emptying or the use of any type of catheter after the 6-week visit.

Patient satisfaction was assessed at 24 months with the question “How satisfied or dissatisfied are you with the result of bladder surgery related to urine leakage?” Patients rated their overall satisfaction, choosing one of five options that ranged from “completely satisfied” to “completely dissatisfied.” Patients who answered that they were either “completely satisfied” or “mostly satisfied” were classified as being satisfied with the outcome.

**STATISTICAL ANALYSIS**

We calculated that 260 women per group would provide a power of 80% to detect a 12% difference between study groups (60% vs. 72%) with the use of a two-sided alternative hypothesis at a significance level of 5%. To allow for attrition and missed visits, we recruited a total of 655 women. Treatment success was assessed at follow-up visits every 6 months. If a treatment failed between scheduled visits, it was considered to have failed at the next visit. Data for women whose treatment was not known to have failed but who had not completed all assessments at the 24-month visit were censored at the last visit on which all failure assessments were complete.

For both outcome measures, we compared the success rates in the two groups at 24 months with the use of time-to-event methods for interval censored data.24 We used Kaplan–Meier product-limit analysis to estimate the success rates at 24 months in the two groups and compared the treatment-failure distributions in the two groups, controlling for stratification by clinical site, with the use of the log-rank test. To determine whether concomitant surgery might have had an effect on the results, we tested the interaction between treatment group and concomitant surgery with the use of the Weibull accelerated failure-time model. All analyses were carried out with SAS statistical software, version 9.2 (SAS Institute).

**RESULTS**

**PATIENTS**

From February 2002 to June 2004, we screened 2405 women for trial eligibility (Fig. 2). Of these women, 556 were ineligible, 1193 declined to
participate or withdrew consent, and 1 died before randomization. A total of 655 women were randomly assigned to a study procedure: 326 to undergo the sling procedure and 329 to undergo the Burch procedure. One woman did not undergo the assigned treatment (Burch procedure), and four women were found to be ineligible after randomization (one assigned to the sling procedure and three assigned to the Burch procedure). A total of 520 women (79%) — 265 in the sling group (81%) and 255 in the Burch group (78%) — either were assessed for treatment success at the 24-month visit or were deemed to have had a treatment failure before that visit.

Women in the two surgical groups were similar in demographic, anthropometric, clinical, and urodynamic-study characteristics (Table 1). The frequency of previous surgery for urinary incontinence was similar in the two groups (13% in the sling group and 15% in the Burch group). The rates of concomitant surgery for pelvic prolapse (including anterior and posterior vaginal repairs, apical suspension procedures, and hysterectomy) were also similar in the two groups (55% in the sling group and 48% in the Burch group). The sling and Burch groups had similar estimated blood loss during the procedure (229 ml and 238 ml, respectively) and similar operative times (136 minutes and 138 minutes, respectively).

Women in the sling group had 24-month cumulative rates of success that were significantly higher than those in the Burch group, with overall success rates of 47% versus 38% (P=0.01), and rates of success specific to stress incontinence of 66% versus 49% (P<0.001) by the log-rank test of equality of distributions with adjustment for site (Fig. 3). There was no clinically or statistically significant interaction effect of concomitant surgery and treatment group on either outcome (P=0.74 for overall success, and P=0.84 for success specific to stress incontinence).

The rate of occurrence of each component of the composite measure of success, as a percentage of patients with complete follow-up assessments, differed according to the treatment group (Fig. 4). These differences reflected the fact that the sling group had lower rates of reported symptoms related to stress incontinence, positive stress tests, and retreatment of stress incontinence than did the Burch group.

There was no significant difference between the sling and Burch groups in the percentage of patients who had serious adverse events (13% and 10%, respectively; P=0.20) (Table 2). However, surgical procedures to reduce voiding symptoms or improve urinary retention were performed exclusively in the sling group, in which 19 patients underwent 20 such procedures. Adverse events were more common in the sling group than in the Burch group (63% vs. 47%, P<0.001), with 415 events among 206 women in the sling group, as compared with 305 events among 156 women in the Burch group. This difference was due primarily to urinary tract infections; 157 women in the sling group (48%) had 305 events and 105 women in the Burch group (32%) had 203 events. When urinary tract infections were excluded, the rates of adverse events were similar in the two groups.
The distribution of time to return to normal voiding differed significantly between the two groups (P<0.001). At the time of hospital discharge, fewer patients in the sling group than in the Burch group had voiding with a residual volume of less than 100 ml (44% vs. 58%), and the difference persisted at 6 weeks (86% vs. 97%). Voiding dysfunction was more common in the sling group than in the Burch group (14% vs. 2%, P<0.001). More patients were treated for postoperative urge incontinence in the sling group than in the Burch group (87 patients [27%] vs. 65 patients [20%], P=0.04). The difference in urge incontinence was due to differences in the proportion of patients treated for persistent urge incontinence (79 patients in the sling group [24%] vs. 59 patients in the Burch group [18%]) rather than to differences in the proportion with new-onset urge incontinence (11 patients [3%] in both groups).

Treatment-satisfaction rates for the 480 subjects who answered the satisfaction question at 24 months were significantly higher in the sling group than in the Burch group (87 patients [27%] vs. 65 patients [20%], P=0.04). The difference in urge incontinence was due to differences in the proportion of patients treated for persistent urge incontinence (79 patients in the sling group [24%] vs. 59 patients in the Burch group [18%]) rather than to differences in the proportion with new-onset urge incontinence (11 patients [3%] in both groups).

Table 1. Selected Characteristics of the Patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Burch Procedure (N=329)</th>
<th>Sling Procedure (N=326)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td>52.2±10.5</td>
<td>51.6±10.1</td>
<td>0.47</td>
</tr>
<tr>
<td>Racial or ethnic group (%)†</td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>75</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic other</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Marital status (%)</td>
<td></td>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td>Married or living with partner</td>
<td>69</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Not married</td>
<td>31</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Education (%)</td>
<td></td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>High school or less</td>
<td>33</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Some training after high school</td>
<td>40</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>College degree or more</td>
<td>27</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Household income (%)</td>
<td></td>
<td></td>
<td>0.65</td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>21</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>$20,000–49,999</td>
<td>29</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>$50,000–79,999</td>
<td>21</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>≥$80,000</td>
<td>29</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td><strong>Risk factors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body-mass index</td>
<td>29.7±6.1</td>
<td>30.3±6.1</td>
<td>0.26</td>
</tr>
<tr>
<td>No. of vaginal deliveries (%)</td>
<td></td>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1–2</td>
<td>46</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>46</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Previous incontinence surgery (%)</td>
<td>15</td>
<td>13</td>
<td>0.46</td>
</tr>
<tr>
<td>Smoking status (%)</td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>Never smoked</td>
<td>59</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Former smoker</td>
<td>29</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>12</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Hormone-replacement therapy (%)</td>
<td></td>
<td></td>
<td>0.66</td>
</tr>
<tr>
<td>Yes</td>
<td>35</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>36</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>No, premenopausal</td>
<td>29</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>
### Table 1. (Continued.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Burch Procedure (N = 329)</th>
<th>Sling Procedure (N = 326)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of life‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score on Urogenital Distress Inventory</td>
<td>150.3±49.9</td>
<td>151.6±47.4</td>
<td>0.73</td>
</tr>
<tr>
<td>Total score on Incontinence Impact Questionnaire</td>
<td>173.2±99.2</td>
<td>169.7±103.4</td>
<td>0.66</td>
</tr>
<tr>
<td>Pad test weight (g)</td>
<td>42.4±61.2</td>
<td>44.7±94.3</td>
<td>0.71</td>
</tr>
<tr>
<td>Incontinence episodes per day (no.)</td>
<td>3.3±3.1</td>
<td>3.1±2.9</td>
<td>0.52</td>
</tr>
<tr>
<td>Urinary-incontinence symptom score§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress score</td>
<td>19.5±4.5</td>
<td>19.2±4.7</td>
<td>0.37</td>
</tr>
<tr>
<td>Urge score</td>
<td>6.6±3.9</td>
<td>6.3±3.9</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>Prolapse stage (%)¶</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 or 1</td>
<td>26</td>
<td>24</td>
<td>0.60</td>
</tr>
<tr>
<td>2</td>
<td>59</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>3 or 4</td>
<td>15</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Q-tip test (degree)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resting angle</td>
<td>15.6±17.1</td>
<td>15.2±18.3</td>
<td>0.77</td>
</tr>
<tr>
<td>Straining angle</td>
<td>61.1±19.3</td>
<td>59.3±17.3</td>
<td>0.23</td>
</tr>
<tr>
<td>Difference between straining angle and resting angle</td>
<td>45.5±19.1</td>
<td>44.1±17.3</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Urodynamic studies (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stress incontinence</td>
<td></td>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td>Yes</td>
<td>89</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Invalid study</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Valsalva leak point pressure‖</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤60 cm of H₂O</td>
<td>4</td>
<td>3</td>
<td>0.46</td>
</tr>
<tr>
<td>Change of ≤60 cm of H₂O</td>
<td>22</td>
<td>20</td>
<td>0.54</td>
</tr>
<tr>
<td>Detrusor overactivity</td>
<td>11</td>
<td>7</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Surgical characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concomitant surgery (%)**</td>
<td></td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>None</td>
<td>44</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Prolapse surgery with repair of anterior vaginal wall (with or without other repair)</td>
<td>17</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Prolapse surgery without repair of anterior vaginal wall (including posterior wall and apex)</td>
<td>31</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Other nonprolapse surgery††</td>
<td>8</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. Body-mass index is the weight in kilograms divided by the square of the height in meters. Percentages may not total 100 because of rounding.
† Racial or ethnic group was reported by the patients.
‡ Scores on the Urogenital Distress Inventory range from 0 to 300, with higher scores indicating greater distress. Scores on the Incontinence Impact Questionnaire range from 0 to 400, with higher scores indicating greater impact.¹⁷
§ Symptom scores are the sum of responses to nine questions regarding stress symptoms (with scores ranging from 0 to 27 and higher scores indicating greater severity) and six questions regarding urge symptoms (with scores ranging from 0 to 18 and higher scores indicating greater severity) adapted from the Medical, Epidemiological, and Social Aspects of Aging questionnaire.¹⁸
¶ Prolapse staging is based on the methods of the Pelvic Organ Prolapse Quantification system.²²
‖ Valsalva leak point pressure refers to the vesical pressure at the time of leakage. The change in the Valsalva leak point pressure is the vesical pressure at the time of leakage minus the baseline vesical pressure.
** Concomitant prolapse repairs included repair of the anterior vaginal wall (anterior colporrhaphy and paravaginal repair), posterior colporrhaphy, apical suspension procedures (sacrospinous ligament suspension, uterosacral ligament suspension, and sacrocolpopexy), enterocele repair, and hysterectomy.
†† Other concomitant surgeries included anal-sphincter repair, tubal ligation, and abdominoplasty.
group than in the Burch group (86% vs. 78%, \( P = 0.02 \)).

**Discussion**

At 24 months, the pubovaginal fascial sling had significantly higher rates of success — both overall and specific to stress incontinence — than did the Burch colposuspension in women with predominant stress incontinence. These findings were not modified by performance of concomitant surgery for pelvic-organ prolapse. In addition, the frequency of surgical retreatment for stress incontinence was greater in the Burch group than in the sling group. Success rates declined steadily over the 2-year follow-up period, which confirmed previous observations \(^{25,26} \) and underscored the need for long-term follow-up in these patients.

However, the higher success rates in the sling group were offset by higher rates of urinary tract infection, urge incontinence, voiding dysfunction, and the need for surgical revision to improve voiding. The increased efficacy and greater morbidity of the sling procedure confirm and quantify the results of previous systematic reviews\(^ {27-29} \) and may explain some of the reluctance in the past to use this procedure as a primary surgical treatment for stress incontinence.\(^ {14} \)

Our large, randomized surgical trial comparing the fascial-sling procedure with the Burch procedure had a robust 24-month follow-up with the use of standardized definitions, procedures, and methods of evaluation to assess a variety of outcome measures and comprehensive postoperative morbidity. The absence of such information to date has precluded rigorous assessment of surgical outcomes for this condition.\(^ {30,31} \) Reported success rates of surgery have varied widely.\(^ {27,28} \) Factors contributing to this variation have included the lack of standardized outcome measures, differences in the baseline characteristics of the study populations, and the length of follow-up.\(^ {32,33} \)

Success rates that are based on reporting by patients are consistently lower than those based on physician-reported measures.\(^ {34,35} \) Current research guidelines emphasize the importance of evaluating treatment efficacy with composite out-

<table>
<thead>
<tr>
<th>Event</th>
<th>Burch Procedure (N = 329)</th>
<th>Sling Procedure (N = 326)</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Serious adverse events‡</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients with event</td>
<td>32 (10)</td>
<td>42 (13)</td>
<td>0.20</td>
</tr>
<tr>
<td>Total events</td>
<td>39</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Genitourinary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ureteral injury</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Ureterovaginal fistula</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Incidental vaginotomy</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Incidental cystotomy</td>
<td>10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Erosion of suture into bladder</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Recurrent cystitis, leading to diagnostic cystoscopy</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Pyelonephritis</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Catheter complication</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Voiding dysfunction leading to surgical revision</td>
<td>0</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Pelvic pain</td>
<td>0</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>Bleeding</td>
<td>3</td>
<td>1</td>
<td>0.62</td>
</tr>
<tr>
<td>Wound complication requiring surgical intervention</td>
<td>13</td>
<td>11</td>
<td>0.83</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>1</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td>Respiratory distress requiring intubation</td>
<td>0</td>
<td>1</td>
<td>0.50</td>
</tr>
<tr>
<td>Laryngospasm requiring reintubation</td>
<td>0</td>
<td>1</td>
<td>0.50</td>
</tr>
</tbody>
</table>
come measures that include both subjective and objective efficacy measures as well as an assessment of morbidity.6-8 Success rates in our trial were low, as compared with those in previous studies.9,10 This finding may be related to our use of composite outcome measures, resulting in a stricter definition of success. The substantial variation in failure rates among studies using single-component measures supports the use of composite outcome measures and highlights the lack of concordance among several traditional measures.

Table 2. (Continued.)

<table>
<thead>
<tr>
<th>Event</th>
<th>Burch Procedure (N = 329)</th>
<th>Sling Procedure (N = 326)</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with event</td>
<td>156 (47)</td>
<td>206 (63)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total events</td>
<td>305</td>
<td>415</td>
<td></td>
</tr>
<tr>
<td>Genitourinary</td>
<td>203</td>
<td>305</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cystitis</td>
<td>202</td>
<td>299</td>
<td></td>
</tr>
<tr>
<td>Pyelonephritis</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Vascular or hematologic</td>
<td>5</td>
<td>9</td>
<td>0.29</td>
</tr>
<tr>
<td>Deep-vein thrombosis</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Bleeding</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Wound complication not requiring surgical intervention</td>
<td>69</td>
<td>71</td>
<td>0.69</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>7</td>
<td>8</td>
<td>0.80</td>
</tr>
<tr>
<td>Pulmonary</td>
<td>10</td>
<td>9</td>
<td>1.00</td>
</tr>
<tr>
<td>Neurologic</td>
<td>6</td>
<td>5</td>
<td>1.00</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>0</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>Allergic (hives, itching)</td>
<td>0</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>Constitutional</td>
<td>3</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>Dermatologic (rash, erythema)</td>
<td>2</td>
<td>4</td>
<td>0.45</td>
</tr>
</tbody>
</table>

*The severity grade was determined by using a slightly modified version of the Dindo classification system, which is based on the level of therapy required to treat an event: grade I, no pharmacologic, surgical, or radiologic intervention (allowed therapeutic regimens include antiemetics, antipyretics, analgesics, diuretics, electrolytes, and physiotherapy); grade II, pharmacologic treatment with drugs other than those allowed for grade I complications (including antibiotics, blood transfusions, and total parenteral nutrition); grade III, surgical, endoscopic, or radiologic intervention; grade IV, life-threatening complication requiring intensive care management; and grade V, death. Serious adverse events were defined as a severity of grade III, grade IV, or grade V; no grade V events occurred in either group.

†P values were calculated with the use of Fisher’s exact test.

‡Catheter complications included clot retention requiring cystoscopy (sling group, 1 patient) or a suprapubic tube stitched in place (Burch group, 1 patient). Wound complications requiring surgical intervention included incisional hernia (Burch, 5 patients; sling, 3), seroma or hematoma (Burch, 2; sling, 3), infection (Burch, 2; sling, 2), abscess (Burch, 1; sling, 1), and vaginal wound revision (Burch, 3; sling, 2). Gastrointestinal complications included 1 rectal injury (in the sling group) and 1 episode of constipation requiring surgical disimpaction (in the Burch group).

§Cystitis was defined as culture-proven bladder infection or, in the absence of a culture, clinical suspicion of a bladder infection that resulted in treatment. Wound complications not requiring surgical intervention included 2 sling exposures (visualization of the sling material in the vagina), incisional hernia (Burch group, 2; sling group, 1), superficial wound-edge separation (Burch, 10; sling, 5), seroma or hematoma (Burch, 13; sling, 11), infection (Burch, 31; sling, 21), and granulation tissue or stitch granulomas (Burch, 13; sling, 31). Gastrointestinal events included ileus (Burch, 5; sling, 2) and other complications (anal fissure, constipation, prolapsed hemorrhoids, nausea and vomiting, abdominal pain, rectal bleeding, and pseudomembranous colitis) (Burch, 2; sling, 6). Pulmonary events included atelectasis (Burch, 4; sling, 6), pneumonia (Burch, 2; sling, 1), pulmonary edema (Burch, 1; sling, 1), and other complications (anesthesia airway difficulty resulting in rescheduling of surgery, oversedation, upper respiratory infection) (Burch, 3; sling, 1). Neurologic complications included sciatica (Burch, 1; sling, 1), numbness or weakness or pain temporarily related to surgery (Burch, 4; sling, 3), and vertigo or vestibular neuritis (Burch, 1; sling, 1). In the sling group, cardiovascular events included bradycardia treated in the recovery room (1) and junctional rhythm ruled out for myocardial infarction (1). In the Burch group, constitutional events included fever of unknown origin (2) and hypokalemia (1).
Our finding that the two procedures had similar success rates as measured by pad tests and voiding diaries may reflect the higher number of patients with symptoms of urge incontinence in the sling group, since these two measures cannot differentiate stress incontinence from urge incontinence. It is likely that we underestimated the rate of postoperative urge incontinence, since our definition was restricted to patients who received treatment for this condition. This factor may explain in part why only 3% of the patients in our trial had new-onset urge incontinence, a rate that is at the low end of the range reported by others.29,39

The higher rate of urinary tract infections reported in the sling group, as compared with the Burch group, may be related to a delayed return to adequate voiding and a prolonged need for catheterization in the sling group. Our definition of urinary tract infection did not require a positive urine culture, and it is possible that some patients received empirical antibiotic therapy for symptoms alone, leading us to overestimate the true incidence of postoperative urinary tract infection in either or both groups. For instance, the higher rate of urge incontinence identified in the sling group may have led to more false diagnoses of urinary tract infection in that group.

All the patients in our study received care in tertiary care centers, and the study population included only women with urethral hypermobility and pure or stress-predominant incontinence. Whether the results can be generalized to other groups of women is uncertain. Both the patients and the health care providers were aware of study-group assignments, and it is possible that bias affected the measurement of some outcomes.

Just over half the women underwent concomitant surgery for pelvic-organ prolapse, a proportion consistent with that in other studies.8 Although we did not find any material differences in success rates on the basis of concomitant surgery, such procedures could potentially influence the number of adverse events identified in both groups.

The sling group also had higher satisfaction rates than did the Burch group, a difference that was consistent with the success rates. However, satisfaction rates were higher in both groups than were success rates, indicating that satisfaction was influenced by factors beyond the resolution of incontinence symptoms. Further analyses are needed to assess the relationships among the satisfaction reported by patients, improvement in the quality of life, and outcome measures described here.

New surgical procedures for stress incontinence continue to be introduced into clinical practice without evaluation of their efficacy and safety in well-designed, randomized clinical trials.27,28 There has been a recent transition from
the fascial sling and Burch procedure to the newer midurethral synthetic sling. A previous randomized surgical trial comparing the midurethral sling with the Burch procedure showed similar efficacy of the two procedures, although that study has been criticized for being underpowered. No randomized trials have compared the midurethral sling with the autologous fascial sling. The relative frequency with which these procedures are performed in the United States is difficult to assess because they have identical procedural codes. Rigorous comparative trials are needed to assess the efficacy and safety of these novel surgical techniques as compared with the efficacy and safety of the procedures studied in our trial.

The number of women undergoing surgical therapy for stress incontinence is increasing, and this trend is likely to continue as the population ages. Our data show that the pubovaginal fascial sling has significantly higher efficacy than the Burch abdominal colposuspension at 24 months in women with predominant stress incontinence, but such success comes at the cost of higher complications. Clinicians should discuss such trade-offs when making recommendations to patients regarding the optimal procedure and should emphasize that complete resolution of incontinence symptoms after surgery is unlikely.

Supported by cooperative agreements (U01 DK58225, U01 DK58229, U01 DK58234, U01 DK58231, U01 DK60379, U01 DK60380, U01 DK60393, U01 DK60395, U01 DK60397, and U01 DK60401) with the National Institute of Diabetes and Digestive and Kidney Diseases and by the National Institute of Child Health and Human Development and Office of Research in Women’s Health of the National Institutes of Health.

Dr. Albo reports receiving consulting fees from Pfizer and Astellas and grant support from Pfizer; Dr. Richter, receiving lecture fees from Pfizer, Esprit Pharmaceutical, and Indevus Pharmaceuticals and grant support from Pfizer; Dr. Brubaker, receiving consulting fees from Pfizer and Q-Med and grant support from Q-Med, Allergan, and Pfizer; Dr. Norton, receiving consulting fees from Eli Lilly; Dr. Kraus, receiving consulting fees from Lilly ICOS and Pfizer, lecture fees from Novartis, Astellas, Pfizer, and Ortho-McNeil, and grant support from Pfizer and GlaxoSmithKline; Dr. Chai, receiving consulting fees from Astellas and Pfizer and grant support from Pfizer; Dr. Zyczynski, receiving consulting fees from Ethicon and grant support from Novartis and Johnson & Johnson; Dr. Diokno, receiving lecture fees from Astellas, Medtronic, Janssen-Cilag, and Ortho-McNeil, and grant support from GlaxoSmithKline, Allergan, and Allergan, and holding a patent on a modified vaginal speculum marketed by Mentor; Dr. Tennstedt, receiving grant support from Lilly ICOS; Dr. Lloyd, receiving consulting fees from Novartis, GlaxoSmithKline, Pfizer, and AMS, lecture fees from Lilly ICOS, Ortho-McNeil, Bayer, and Boehringer Ingelheim, and grant support from Ortho-McNeil, MediciNova, and Allergan; Dr. FitzGerald, receiving consulting fees from GlaxoSmithKline, lecture fees from Medtronic, and grant support from Pfizer; Dr. Lemack, receiving consulting fees from Pfizer and Allergan and lecture fees from Pfizer and Astellas and having equity interest in Pfizer; Dr. Johnson, receiving lecture fees from Johnson & Johnson; Dr. Stoddard, having equity in Johnson & Johnson, Bristol-Myers Squibb, Elan, and Procter & Gamble; Dr. Meneely, receiving grant support from Boston Scientific and having equity interest in Pfizer; Dr. Varner, receiving lecture fees from Pfizer; Dr. Moalli,
receiving lecture fees from Boston Scientific that were contributed to a fellowship research fund; Dr. Steers, receiving consulting fees from Dynogen, Sanofi-Aventis, Astellas, and Allergan, lecture fees from Watson, Astellas, and GlaxoSmithKline Beecham, and grant support from Pfizer and having equity interest in Abbott and Johnson & Johnson. No other potential conflict of interest relevant to this article was reported.

We thank Kathleen Cannon at the New England Research Institutes for her tireless organizational efforts on behalf of the Urinary Incontinence Treatment Network, and Bette Jo Garrett and Patricia Lane at the University of California at San Diego for their editorial assistance.

APPENDIX


REFERENCES


Copyright © 2007 Massachusetts Medical Society.

---

**FULL TEXT OF ALL JOURNAL ARTICLES ON THE WORLD WIDE WEB**

Access to the complete text of the _Journal_ on the Internet is free to all subscribers. To use this Web site, subscribers should go to the _Journal’s_ home page (www.nejm.org) and register by entering their names and subscriber numbers as they appear on their mailing labels. After this one-time registration, subscribers can use their passwords to log on for electronic access to the entire _Journal_ from any computer that is connected to the Internet. Features include a library of all issues since January 1993 and abstracts since January 1975, a full-text search capacity, and a personal archive for saving articles and search results of interest. All articles can be printed in a format that is virtually identical to that of the typeset pages. Beginning 6 months after publication, the full text of all Original Articles and Special Articles is available free to nonsubscribers who have completed a brief registration.
Advanced Life Support for Out-of-Hospital Respiratory Distress


ABSTRACT

BACKGROUND
Respiratory distress is a common symptom of patients transported to hospitals by emergency medical services (EMS) personnel. The benefit of advanced life support for such patients has not been established.

METHODS
The Ontario Prehospital Advanced Life Support (OPALS) Study was a controlled clinical trial that was conducted in 15 cities before and after the implementation of a program to provide advanced life support for patients with out-of-hospital respiratory distress. Paramedics were trained in standard advanced life support, including endotracheal intubation and the administration of intravenous drugs.

RESULTS
The clinical characteristics of the 8138 patients in the two phases of the study were similar. During the first phase, no patients were treated by paramedics trained in advanced life support; during the second phase, 56.6% of patients received this treatment. Endotracheal intubation was performed in 1.4% of the patients, and intravenous drugs were administered to 15.0% during the second phase. This phase of the study was also marked by a substantial increase in the use of nebulized salbutamol and sublingual nitroglycerin for the relief of symptoms. The rate of death among all patients decreased significantly, from 14.3% to 12.4% (absolute difference, 1.9%; 95% confidence interval [CI], 0.4 to 3.4; \( P = 0.01 \)) from the basic-life-support phase to the advanced-life-support phase (adjusted odds ratio, 1.3; 95% CI, 1.1 to 1.5).

CONCLUSIONS
The addition of a specific regimen of out-of-hospital advanced-life-support interventions to an existing EMS system that provides basic life support was associated with a decrease in the rate of death of 1.9 percentage points among patients with respiratory distress.
EACH YEAR, EMERGENCY MEDICAL SERVICES (EMS) personnel in the United States transport 2 million patients with respiratory distress to hospitals by ambulance. Respiratory distress is the second most common symptom of adults transported by ambulance and is associated with a relatively high overall mortality before hospital discharge of 18%. Among the most common causes of respiratory distress in this setting are congestive heart failure, pneumonia, chronic obstructive pulmonary disease, and asthma.

In many cities in the United States and Canada, out-of-hospital care for critically ill and injured patients is provided by paramedics who are trained in advanced-life-support measures. Advanced life support includes endotracheal intubation and intravenous drug therapy. In contrast, paramedics who are trained in basic-life-support measures administer oxygen, bag–valve–mask ventilation, and in some cases nebulized bronchodilators and sublingual nitroglycerin, but they do not perform endotracheal intubation or administer intravenous drugs.

The benefit of advanced life support for patients with respiratory distress has not been established. There are few controlled clinical trials of out-of-hospital advanced life support and respiratory distress and, consequently, there is very little evidence regarding the optimal therapy for patients before they arrive at the hospital. To our knowledge, no studies have shown improved survival for patients with respiratory distress who receive advanced life support before they arrive at the hospital, and there is some evidence that inappropriate drug therapy in this setting may increase the rate of death.

In Ontario, a Canadian jurisdiction of 12 million people, the provincial government has funded the Ontario Prehospital Advanced Life Support (OPALS) Study, a large, multicenter, controlled clinical trial. This multiphase study evaluated specific programs in several cities to determine the incremental benefit to survival and morbidity associated with out-of-hospital advanced life support for four major groups of critically ill and injured patients (those with cardiac arrest, major trauma, respiratory distress, and chest pain). We have shown that advanced-life-support programs have no significant effect on the outcomes of patients with cardiac arrest. The objective of the current study, the OPALS Respiratory Distress Study, was to assess the incremental benefit with respect to morbidity and mortality that results from the implementation of an advanced-life-support program for the evaluation and management of respiratory distress before patients arrive at the hospital.

METHODS

DESIGN

Detailed methods for the OPALS Respiratory Distress Study have been described previously. We performed a prospective “before-and-after” controlled trial (before and after advanced-life-support programs were instituted) among all eligible patients with respiratory distress seen during two distinct phases of the study: the basic-life-support phase (6 months) and the advanced-life-support phase (6 months). The study was funded by peer-reviewed grants from the Emergency Health Services Branch of the Ontario Ministry of Health and Long-Term Care and the Canadian Health Services Research Foundation.

SETTING AND POPULATION

The study was conducted in 18 urban communities throughout Ontario under the medical direction of 11 base-hospital programs. The aggregate population was 2.5 million people, with the populations of individual cities ranging from 20,000 to 750,000. One community had a population of less than 30,000, five had populations of 30,000 to 99,999, four had populations of 100,000 to 199,999, four had populations of 200,000 to 500,000, and one had a population of more than 500,000. Each community was served by a Central Ambulance Communications Center, which provided the study with electronic and synchronized dispatch information regarding all patients transported by ambulance during the study. Out-of-hospital care was documented with the use of the standard Ontario Ambulance Call Report form, which included specific data regarding the call code, time of events, medications administered, and procedures performed.

The study population included all patients 16 years of age and older whose primary symptom was shortness of breath, including those who were assessed by EMS personnel but not transported to the hospital. Excluded were patients with full cardiac arrest before the arrival of EMS personnel, patients whose primary symptom was chest pain or any other nonrespiratory symptom, and patients with respiratory distress due to trauma, a postictal...
state, or another nonrespiratory illness, according to information available to paramedics at the time of the initial assessment of the patient in the field. The study received full approval of the Ottawa Hospital Research Ethics Board, and the requirement for informed consent was waived.

**INTERVENTION**

During the basic-life-support phase, each community provided tiered EMS, with firefighters responding first, followed by “primary care” paramedics. These paramedics had previously graduated from a 10-month program at a community college and were trained to provide all basic-life-support measures, including oxygen, bag–valve–mask ventilation, and automated external defibrillation. All paramedics also had several years of experience (median, 5 years).

The study intervention consisted of an advanced-life-support program in which primary care paramedics were trained to perform endotracheal intubation, insert intravenous lines, and administer intravenous medications. After this training, they were called “advanced-care” paramedics. The Emergency Medical Technician Level III training program of the Canadian Medical Association involved 6 weeks of didactic instruction, 6 weeks of clinical instruction, and 12 weeks of preceptorship training in the field. To qualify for the advanced-life-support phase of the OPALS Study, each community had to meet four criteria with regard to patients with cardiac arrest. First, EMS technicians had to achieve a rapid-defibrillation response interval of 8 minutes or less for 90% of patients. Second, paramedics trained to provide advanced care had to respond for 95% of patients. Third, paramedics trained to provide advanced care had to respond to the scene within 11 minutes for 80% of patients. Finally, paramedics trained to provide advanced care had to successfully perform endotracheal intubation for 90% of patients. These criteria were monitored regularly, and data collection for the advanced-life-support phase of the study in each community did not begin until the criteria were met. The three communities that did not meet the standards were excluded from the study.

During the advanced-life-support phase, the decision to dispatch a crew trained to provide advanced life support was made by the dispatcher on the basis of information provided during the initial emergency call and the availability of a team that could provide this type of support at the time of the call. Medications administered to patients with respiratory distress during this phase included intravenous furosemide and morphine as well as nebulized salbutamol and sublingual nitroglycerin. In some instances, patients also received nebulized salbutamol and sublingual nitroglycerin during the basic-life-support phase as part of a “symptom relief” program. This program was gradually introduced to primary care paramedics throughout Ontario during the end of the basic-life-support phase of the OPALS Study.

**OUTCOME MEASURES**

The primary outcome measure was mortality, defined as the rate of death before hospital discharge regardless of the duration of admission. Secondary outcome measures included intubation in the emergency department, evidence of aspiration, admission to a hospital, the length of stay in the hospital, the patient’s destination after discharge, and the patient’s functional status according to a five-point cerebral-performance category scale. An additional end point was paramedic coding of the patient’s status as being improved, unchanged, or worsened on the patient’s arrival in the emergency department. Study data provided by each base-hospital program included ambulance call reports, dispatch reports, and a review of hospital records. Trained analysts determined the final discharge diagnoses on the basis of hospital records. For a few patients for whom hospital records were not available, data regarding survival to 30 days after the day of study enrollment was ascertained by a review of records from the Ontario Death Registry.

**STATISTICAL ANALYSIS**

For comparisons of mortality, the minimal sample size was estimated to be 4630 patients in the basic-life-support phase and 4630 patients in the advanced-life-support phase, on the basis of a type I error of 0.05, a type II error of 0.20, a baseline mortality of 17%, and a clinically important difference of 2%. We therefore defined the 6-month duration of each phase of the study based on the expectation that we would be able to enroll at least this number of patients in that time interval.

The primary outcome measure of death before hospital discharge was assessed with chi-square analysis. Ninety-five percent confidence intervals were calculated for the absolute difference in mor-
tality between phases. Stepwise logistic-regression analysis was performed to control for possible confounding variables. These variables included age, sex, initial respiratory rate, initial pulse rate, EMS priority return code (a measure of urgency assigned by the on-site paramedic after initial assessment of the patient’s condition), treatment administered, and final diagnosis. Comparisons of the rates of death between the two phases were made for the following subgroups: community size, discharge diagnosis, and EMS return code. Differences between the phases for data other than mortality were analyzed with the Wilcoxon signed rank-sum test, the chi-square test, Fisher’s exact test, or Student’s t-test, as appropriate. All reported P values are two-sided and not adjusted for multiple testing.

RESULTS

The study enrolled 8138 patients from 15 communities: 3920 in the 6-month basic-life-support phase (from January 1995 to February 1998) and 4218 in the advanced-life-support phase (from February 1998 to November 2000). In each community, the two phases were separated by a run-in period of 6 to 36 months to allow for training in advanced life support. In general, patients in the two phases had similar characteristics (Table 1).

Table 2 shows the EMS responses during the two phases. The median response intervals were similar in the two phases. Advanced-life-support crews responded to 56.6% of patients in the advanced-life-support phase. Although the use of respiratory support measures increased in this phase, fewer than 3.0% of patients received bag–valve–mask ventilation and fewer than 2.0% of patients underwent intubation. Intravenous medications (most often furosemide) were given to 15.0% of patients in the advanced-life-support phase. The use of medications for symptom relief (primarily nebulized salbutamol) increased markedly (from 15.7% to 59.4%, P<0.001). In addition, the rate of intubation in the emergency department decreased from 5.3% to 3.1% (P<0.001). There was no significant change in the presence of aspiration on chest radiography, although surveillance for this outcome was much more thorough in the advanced-life-support phase. There was only a very small difference between phases for hospital admission rates (67.8% vs. 65.0%), and no significant difference in mean length of hospital stay in days.

We evaluated a number of clinically important subgroups (Table 4). A reduction in mortality during the advanced-life-support phase was suggested for patients with the final diagnosis of congestive heart failure (15.1% vs. 10.9%) but not for those with other discharge diagnoses. However, a test for interaction did not confirm a statistically demonstrable difference between the effect of out-of-hospital advanced-life-support measures in patients with congestive heart failure and in patients with other diagnoses.

In another subgroup analysis, there was evidence that the benefit of advanced-life-support measures on mortality was seen only among patients in the larger cities in the study (those with more than 100,000 people) but not in the smaller communities. Finally, as compared with the basic-life-support phase, a reduction in mortality was seen in the advanced-life-support phase when the EMS return code was recorded as “not urgent” (11.7% vs. 9.8%) but not when the EMS return code was recorded as “urgent.” However, a test for
interaction showed that neither of these trends was statistically significant.

**DISCUSSION**

In this trial, we evaluated the effect of out-of-hospital advanced life support on the outcomes of patients with respiratory distress. Although there was a significant reduction in overall mortality during the advanced-life-support phase of the trial, the magnitude of the observed decrease did not exceed the prespecified, minimal, clinically important difference of 2 percentage points. In addition, there was a significant increase in the proportion

| Table 1. Baseline Characteristics of the 8138 Patients in the OPALS Respiratory Distress Study.* |
|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| **Characteristic** | **Basic-Life-Support Phase (N=3920)** | **Advanced-Life-Support Phase (N=4218)** |
| Age — yr | | |
| Mean | 70.8±16.6 | 70.2±17.2 |
| Range | 16–107 | 16–102 |
| Male sex — no. (%) | 1882 (48.0) | 1934 (45.9) |
| Population of city — no. (%) | | |
| <30,000 | 63 (1.6) | 90 (2.1) |
| 30,000–99,999 | 729 (18.6) | 601 (14.2) |
| 100,000–199,999 | 642 (16.4) | 769 (18.2) |
| 200,000–500,000 | 1438 (36.7) | 1586 (37.6) |
| >500,000 | 1048 (26.7) | 1172 (27.8) |
| EMS return code — no. (%) | | |
| Urgent | 1418 (36.2) | 1413 (33.5) |
| Prompt | 2445 (62.4) | 2579 (61.1) |
| Deferrable | 30 (0.8) | 100 (2.4) |
| Declined transport | 27 (0.7) | 126 (3.0) |
| EMS severity status score of “severe” — no./total no. (%) † | 1383/3651 (37.9) | 1444/4073 (35.5) |
| Initial GCS score of 15 — no./total no. (%) ‡ | 3050/3548 (86.0) | 3588/4160 (86.2) |
| Initial heart rate — beats per minute | 99.6±22.3 | 101.2±22.7 |
| Initial respiratory rate — breaths per minute | 28.4±7.9 | 28.5±8.3 |
| Final diagnosis — no./total no. (%) | | |
| Congestive heart failure | 1009/3605 (28.0) | 861/3649 (23.6) |
| Chronic obstructive pulmonary disease | 670/3605 (18.6) | 702/3649 (19.2) |
| Pneumonia | 500/3605 (13.9) | 468/3649 (12.8) |
| Other respiratory condition | 258/3605 (7.2) | 341/3649 (9.3) |
| Asthma | 269/3605 (7.5) | 279/3649 (7.6) |
| Other cardiovascular condition | 151/3605 (4.2) | 175/3649 (4.8) |
| Myocardial infarction | 86/3605 (2.4) | 106/3649 (2.9) |
| Bronchitis | 130/3605 (3.6) | 159/3649 (4.4) |
| Lung cancer | 140/3605 (3.9) | 106/3649 (2.9) |
| Congestive heart failure or chronic obstructive pulmonary disease | 60/3605 (1.7) | 29/3649 (0.8) |
| Other condition | 332/3605 (9.2) | 423/3649 (11.6) |

* Plus–minus values are means ±SD. GCS denotes Glasgow Coma Scale.
† Scores range from minor to moderate, severe, life-threatening, and “vital signs absent.”
‡ Scores range from 3 to 15, with higher scores indicating a better condition.
Table 2. EMS Response for the 8138 Patients in the OPALS Respiratory Distress Study.*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Basic-Life-Support Phase (N=3920)</th>
<th>Advanced-Life-Support Phase (N=4218)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Paramedics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary-care paramedics on scene — no. (%)</td>
<td>3920 (100.0)</td>
<td>1829 (43.4)</td>
</tr>
<tr>
<td>Advanced-care paramedics — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On scene</td>
<td>0</td>
<td>2389 (56.6)</td>
</tr>
<tr>
<td>On scene in 11 min</td>
<td>0</td>
<td>1988 (47.1)</td>
</tr>
<tr>
<td><strong>EMS return code</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients with EMS return code “urgent” — no./total no. (%)</td>
<td>0/1418</td>
<td>866/1413 (61.3)</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bag–valve–mask ventilation — no. (%)</td>
<td>92 (2.3)</td>
<td>123 (2.9)</td>
</tr>
<tr>
<td>Endotracheal intubation — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attempted</td>
<td>NA</td>
<td>70 (1.7)</td>
</tr>
<tr>
<td>Successful</td>
<td>NA</td>
<td>61 (1.4)</td>
</tr>
<tr>
<td>Administration of intravenous medications — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furosemide</td>
<td>NA</td>
<td>637 (15.1)</td>
</tr>
<tr>
<td>Morphine</td>
<td>NA</td>
<td>609 (14.4)</td>
</tr>
<tr>
<td>Fluid bolus</td>
<td>NA</td>
<td>62 (1.5)</td>
</tr>
<tr>
<td>Administration of medications for symptom relief — no. (%)</td>
<td>614 (15.7)</td>
<td>2507 (59.4)</td>
</tr>
<tr>
<td>Nebulized salbutamol — no. (%)</td>
<td>585 (14.9)</td>
<td>2268 (53.8)</td>
</tr>
<tr>
<td>Sublingual nitroglycerin — no. (%)</td>
<td>29 (0.7)</td>
<td>397 (9.4)</td>
</tr>
<tr>
<td><strong>Response intervals — min</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call receipt to crew notification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>0.5–1.2</td>
<td>0.5–1.1</td>
</tr>
<tr>
<td>Crew notification to vehicle arrival at scene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5.9</td>
<td>6.3</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>4.3–8.1</td>
<td>4.6–8.4</td>
</tr>
<tr>
<td>Crew notification to ambulance with basic-life-support team at scene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5.9</td>
<td>6.1</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>4.3–8.1</td>
<td>4.5–8.3</td>
</tr>
<tr>
<td>Crew notification to ambulance with advanced-life-support team at scene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>NA</td>
<td>6.4</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>NA</td>
<td>4.7–8.5</td>
</tr>
<tr>
<td>Vehicle arrival at scene to arrival at patient’s side</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>1.0–2.0</td>
<td>2.0–2.0</td>
</tr>
<tr>
<td>Arrival at patient’s side to departure from scene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>11.7</td>
<td>14.8</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>9.0–14.9</td>
<td>11.4–18.7</td>
</tr>
<tr>
<td>Departure from scene to arrival at hospital</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>6.1</td>
<td>6.9</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>3.9–9.6</td>
<td>4.5–10.6</td>
</tr>
</tbody>
</table>

* NA denotes not applicable.
of patients with a cerebral-performance category score of level 1. These improvements in outcome were achieved despite the fact that providers of advanced life support attended fewer than 60% of patients in the second phase of the study and the two advanced-life-support interventions (endotracheal intubation and the administration of intravenous medication) were performed in only 1.4% and 15.0% of patients, respectively.

We performed subgroup analyses to determine whether the survival benefit varied from group to group. The subgroup of patients with a discharge diagnosis of congestive heart failure, as compared with those with other diagnoses, was more likely to have a reduction in mortality during the advanced-life-support phase. However, an interaction test did not confirm a significant difference in effect among patients with the most common discharge diagnoses. Patients in cities with a population of more than 100,000 were also more likely to benefit during the second phase of the trial, as were patients with an EMS return code of “not urgent.”

Previous data regarding the benefit of advanced life support for patients with shortness of breath are limited. To our knowledge, there have been no previous controlled trials and no previous studies that clearly show improved survival with advanced airway measures or the administration of medication for patients with congestive heart failure.\textsuperscript{6,7,12,17} Three small studies evaluated the feasibility but not the effectiveness of techniques to maintain positive airway pressure in patients being transported in ambulances.\textsuperscript{18–20} For patients with asthma, several small studies evaluated the administration of beta-agonists in out-of-hospital settings and showed an improvement in the peak expiratory flow rate but no improvement in the rate of deaths among patients.\textsuperscript{5,8,9}

An important potential limitation of our study is that it was designed as a before-and-after controlled trial rather than as a randomized trial and, as such, it had a historical rather than a contemporaneous control group. It was not possible for individual patients to undergo randomization because the paramedics considered the random with-

### Table 3. Mortality, Functional Status, and Other Outcomes of Patients from the Two Study Phases.\textsuperscript{*}

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Basic-Life-Support Phase (N = 3920)</th>
<th>Advanced-Life-Support Phase (N = 4218)</th>
<th>Absolute Change (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall mortality — no. (%)†</td>
<td>560 (14.3)</td>
<td>522 (12.4)</td>
<td>1.9 (0.4 to 3.4)</td>
<td>0.01</td>
</tr>
<tr>
<td>Cerebral-performance category score, level 1 — no./total no. (%)</td>
<td>1559/2983 (52.3)</td>
<td>1723/2756 (62.5)</td>
<td>10.3 (7.7 to 12.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Outcomes in emergency department — no./total no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCS score of 15 on arrival</td>
<td>1055/1215 (86.8)</td>
<td>1274/1455 (87.6)</td>
<td>0.7 (−1.9 to 3.3)</td>
<td>0.57</td>
</tr>
<tr>
<td>Status of patient on arrival</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved</td>
<td>927/3784 (24.5)</td>
<td>1876/4096 (45.8)</td>
<td>21.3 (19.2 to 23.4)</td>
<td></td>
</tr>
<tr>
<td>Unchanged</td>
<td>2673/3784 (70.6)</td>
<td>2033/4096 (49.6)</td>
<td>21.0 (18.9 to 23.1)</td>
<td></td>
</tr>
<tr>
<td>Worsened</td>
<td>182/3784 (4.8)</td>
<td>177/4096 (4.3)</td>
<td>0.5 (−0.4 to 1.4)</td>
<td></td>
</tr>
<tr>
<td>Lost vital signs en route</td>
<td>2/3784 (0.1)</td>
<td>10/4096 (0.2)</td>
<td>0.2 (0.0 to 0.4)</td>
<td></td>
</tr>
<tr>
<td>Underwent intubation</td>
<td>190/3583 (5.3)</td>
<td>110/3580 (3.1)</td>
<td>2.2 (1.3 to 3.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aspiration</td>
<td>45/2155 (2.1)</td>
<td>67/3471 (1.9)</td>
<td>0.2 (0.6 to 0.9)</td>
<td>0.68</td>
</tr>
<tr>
<td>Death</td>
<td>46/3657 (1.3)</td>
<td>46/3702 (1.2)</td>
<td>0.0 (−0.5 to 0.5)</td>
<td>1.0</td>
</tr>
<tr>
<td>Outcomes in hospital</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admission — no./total no. (%)</td>
<td>2478/3657 (67.8)</td>
<td>2405/3702 (65.0)</td>
<td>2.8 (0.6 to 5.0)</td>
<td>0.01</td>
</tr>
<tr>
<td>Length of stay — days</td>
<td>9.8±13.2</td>
<td>9.4±12.2</td>
<td>0.4</td>
<td>0.20</td>
</tr>
<tr>
<td>Disposition to home — no./total no. (%)</td>
<td>2415/3665 (65.9)</td>
<td>2457/3668 (67.0)</td>
<td>1.1 (−1.1 to 3.3)</td>
<td>0.32</td>
</tr>
<tr>
<td>Death — no./total no. (%)</td>
<td>514 (13.1)</td>
<td>476 (13.1)</td>
<td>1.7 (0.3 to 3.2)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

\textsuperscript{*} For patients with asthma, several small studies evaluated the administration of beta-agonists in out-of-hospital settings and showed an improvement in the peak expiratory flow rate but no improvement in the rate of deaths among patients.\textsuperscript{5,8,9}

\textsuperscript{†} For patients with asthma, several small studies evaluated the administration of beta-agonists in out-of-hospital settings and showed an improvement in the peak expiratory flow rate but no improvement in the rate of deaths among patients.\textsuperscript{5,8,9}

\textsuperscript{6,7,12,17} Three small studies evaluated the feasibility but not the effectiveness of techniques to maintain positive airway pressure in patients being transported in ambulances.\textsuperscript{18–20} For patients with asthma, several small studies evaluated the administration of beta-agonists in out-of-hospital settings and showed an improvement in the peak expiratory flow rate but no improvement in the rate of deaths among patients.\textsuperscript{5,8,9}
holding of potentially lifesaving procedures to be unethical. Such a study would have been logistically difficult to carry out in any case. In addition, the primary outcome measure, death, was not subject to ascertainment bias. Selection bias was minimized by the population-based approach of including all patients from the study communities. A program to administer medications for symptom relief (nebulized salbutamol and sublingual nitroglycerin) was introduced toward the end of the first phase of this study. Although this program was not specifically related to advanced life support, it may have been a factor that influenced the benefit in the second phase of the study. Positive-airway-pressure therapy was also introduced in some emergency departments during the study period; this could have influenced the outcome for some of the patients in the study.

The implications of this study require careful consideration. The patients in the second phase of the study had a significantly lower mortality than those in the first phase. We estimate that 53 is the number needed to treat for the entire cohort with shortness of breath, and in the study regions with 2.5 million people, approximately 161 lives would be saved each year.

However, it is less clear which interventions should be considered essential and how they should be implemented. In this study, very few patients underwent intubation, and of the intravenous medications, only furosemide was given to a large number of patients (14.4%). The most substantial change in therapeutic intervention was the marked increase in the use of medications for symptom relief; this intervention is not a component of advanced life support, and it was implemented as part of a separate program. Thus, the benefit of the intervention in this trial may have been primarily due not to the availability of advanced-life-support techniques but to the use of nebulized salbutamol and sublingual nitroglycerin. However, it is difficult to analyze the effect of individual measures in this study, since the patients treated with any given intervention likely differed from those who did not receive that intervention and it would be difficult to define a comparable subgroup within the control sample. The reduction in mortality among patients in this study was entirely due to a reduction in the in-hospital mortality, with no change in the mortality in the emergency department. Although many of the patients who died presumably did so soon after admission,

### Table 4. Mortality among Clinically Important Subgroups.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Patients</th>
<th>Patients Who Died before Discharge</th>
<th>Difference (95% CI)</th>
<th>P Value for Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population of city — no./total no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30,000</td>
<td>63 (1.6)</td>
<td>90 (2.1)</td>
<td>7/63 (11.1)</td>
<td>12/90 (13.3)</td>
</tr>
<tr>
<td>30,000–99,999</td>
<td>729 (18.6)</td>
<td>601 (14.2)</td>
<td>92/729 (12.6)</td>
<td>102/601 (17.0)</td>
</tr>
<tr>
<td>100,000–199,999</td>
<td>644 (16.4)</td>
<td>769 (18.2)</td>
<td>102/644 (15.8)</td>
<td>85/769 (11.1)</td>
</tr>
<tr>
<td>200,000–500,000</td>
<td>1438 (36.7)</td>
<td>1600 (37.9)</td>
<td>211/1438 (14.7)</td>
<td>199/1600 (12.4)</td>
</tr>
<tr>
<td>&gt;500,000</td>
<td>1048 (26.7)</td>
<td>1175 (27.9)</td>
<td>148/1048 (14.1)</td>
<td>124/1175 (10.6)</td>
</tr>
<tr>
<td>Discharge diagnosis — no./total no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>1009/3605 (28.0)</td>
<td>861/3649 (23.6)</td>
<td>152 (15.1)</td>
<td>94 (10.9)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>670/3605 (18.6)</td>
<td>702/3649 (19.2)</td>
<td>51 (7.6)</td>
<td>52 (7.4)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>500/3605 (13.9)</td>
<td>468/3649 (12.8)</td>
<td>113 (22.6)</td>
<td>94 (20.0)</td>
</tr>
<tr>
<td>Asthma</td>
<td>269/3605 (7.5)</td>
<td>279/3649 (7.6)</td>
<td>0 (0.0)</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td>EMS return code — no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not urgent</td>
<td>2475 (63.1)</td>
<td>2679 (63.5)</td>
<td>290 (11.7)</td>
<td>263 (9.8)</td>
</tr>
<tr>
<td>Urgent</td>
<td>1418 (36.2)</td>
<td>1413 (33.5)</td>
<td>270 (19.0)</td>
<td>258 (18.3)</td>
</tr>
</tbody>
</table>
these data do raise the question of whether other interventions occurring after the patients arrived at the hospital played a role in the improvement in outcome.

Analyses of the benefit for patients with specific discharge diagnoses are of some interest, but the decision to dispatch an advanced-life-support team cannot be made on the basis of a subsequently determined discharge diagnosis. There was more evidence of a survival benefit among patients with an EMS return code of “not urgent” than among those with a code of “urgent,” so it is unclear whether patients were more likely to benefit if they were less ill. Finally, the benefit of an advanced-life-support program must be balanced against the relatively high cost of its implementation.

The OPALS Respiratory Distress Study showed that the introduction of an EMS advanced-life-support program and interventions for symptom relief significantly reduced mortality for patients with shortness of breath. It is unclear whether these data are sufficient to justify implementation of the entire program of interventions described here. Further research should target populations and evaluate the optimal treatment regimens for patients with out-of-hospital respiratory distress.

Supported by peer-reviewed grants from the Emergency Health Services Branch of the Ontario Ministry of Health and Long-Term Care and the Canadian Health Services Research Foundation and by a Distinguished Investigator Award from the Canadian Institutes of Health Research (to Dr. Stiell).

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

We thank the OPALS Study Group investigators and other members of the OPALS Study Coordinating Center: Tammy Beaudoin (research coordinator), David Brissom (research coordinator), Irene Harris (administrator secretary), and My-Linh Tran (database coordinator). We thank Cathy Francis of the Ministry of Health and Long-Term Care for her support.

APPENDIX


REFERENCES


Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants

Ciaran S. Phibbs, Ph.D., Laurence C. Baker, Ph.D., Aaron B. Caughey, M.D., Ph.D., Beate Danielsen, Ph.D., Susan K. Schmitt, Ph.D., and Roderic H. Phibbs, M.D.

ABSTRACT

BACKGROUND
There has been a large increase in both the number of neonatal intensive care units (NICUs) in community hospitals and the complexity of the cases treated in these units. We examined differences in neonatal mortality among infants with very low birth weight (below 1500 g) among NICUs with various levels of care and different volumes of very-low-birth-weight infants.

METHODS
We linked birth certificates, hospital discharge abstracts (including interhospital transfers), and fetal and infant death certificates to assess neonatal mortality rates among 48,237 very-low-birth-weight infants who were born in California hospitals between 1991 and 2000.

RESULTS
Mortality rates among very-low-birth-weight infants varied according to both the volume of patients and the level of care at the delivery hospital. The effect of volume also varied according to the level of care. As compared with a high level of care and a high volume of very-low-birth-weight infants (more than 100 per year), lower levels of care and lower volumes (except for those of two small groups of hospitals) were associated with significantly higher odds ratios for death, ranging from 1.19 (95% confidence interval [CI], 1.04 to 1.37) to 2.72 (95% CI, 2.37 to 3.12). Less than one quarter of very-low-birth-weight deliveries occurred in facilities with NICUs that offered a high level of care and had a high volume, but 92% of very-low-birth-weight deliveries occurred in urban areas with more than 100 such deliveries.

CONCLUSIONS
Mortality among very-low-birth-weight infants was lowest for deliveries that occurred in hospitals with NICUs that had both a high level of care and a high volume of such patients. Our results suggest that increased use of such facilities might reduce mortality among very-low-birth-weight infants.
PARALLELING THE LITERATURE ON ADULT CARE,4-3 many studies of neonatal care have shown a lower mortality rate in hospitals with higher volumes of patients than in those with lower volumes.4-7 Other studies have examined the association between the level of neonatal care and outcomes. Neonatal care is formally regionalized, with assigned levels of care and specific guidelines that define the characteristics of infants who should be delivered, and cared for, at each level of care. Each neonatal intensive care unit (NICU) that offers a lower level of care must have a formal contractual relationship with a NICU that provides tertiary care.8 Higher levels of care are associated with lower neonatal mortality, particularly among infants with very low birth weight (below 1500 g).4-6,9-17

Growth in the number of NICUs in community hospitals throughout the United States over the past two decades (i.e., deregionalization) has resulted in increasing numbers of high-risk newborns receiving care in low-volume units offering midlevel care.6,11,18-21 It is uncertain whether lower-volume, lower-level NICUs are associated with worse outcomes. One of the complexities in addressing this question is that the units with the highest level of care are also typically those with the highest volume, making it difficult to ascertain whether both volume and level are independent predictors of neonatal outcome.

We,4,6 as well as other investigators,5,9-17,22-25 have previously demonstrated a relationship between the level of NICU care and neonatal outcome. However, most previous studies, including our own, involved relatively small samples or narrowly defined networks and thus could not adequately assess interaction between volume and level of care. In addition, most studies were based on data collected before the routine use of surfactant-replacement therapy, which has substantially improved mortality rates among very-low-birth-weight infants.

These infants are a vulnerable group and thus particularly likely to be affected by hospital services; in 2000, they accounted for only 1.4% of births but 51% of infant deaths.26 In the current study, we used data collected from all hospitals in California from 1991 to 2000 to examine the effects of NICU level of care and patient volume on mortality among very-low-birth-weight infants. These data reflect outcomes reported after the reduction in mortality associated with the introduction of surfactant-replacement therapy in 1990 and, for the most part, after the increased use of antenatal corticosteroid therapy that occurred after the publication of the results of the National Institute of Child Health and Human Development consensus conference in 1994.27-30

**METHODS**

**STUDY DESIGN**

We obtained data on very-low-birth-weight infants born in California hospitals and on in-hospital infant and fetal deaths for the period from January 1, 1991, to December 31, 2000 (66,838 infants). California birth and death certificates were linked to hospital-discharge abstracts for both mothers and infants. The death certificates included both infant and fetal death certificates. More than 99% of the maternal and infant discharge abstracts were successfully linked with infant birth certificates.31,32 The birth-certificate data were also successfully linked to infant discharge abstracts from the receiving hospital for 99% of the infants who were transferred to another hospital. The study was approved by the human subjects committees at Stanford University and the California Office of Statewide Health Planning and Development, and the requirement for obtaining informed consent was waived.

**LEVELS OF CARE**

We defined levels of care as follows: level 1, no NICU; level 2, a NICU that provides care for mildly ill infants but does not provide mechanical ventilation; level 3A, a NICU that provides mechanical ventilation with restrictions (e.g., only for infants with a birth weight above 1000 g); level 3B, a NICU that provides mechanical ventilation without restrictions but does not provide major surgery; level 3C, a NICU that provides major neonatal surgery but neither open-heart surgery nor extracorporeal membrane oxygenation (ECMO); and level 3D, a NICU that provides cardiac surgery requiring cardiopulmonary bypass or ECMO. These definitions are based on the draft version of the American Academy of Pediatrics report on NICU levels of care4,33 to differentiate NICUs in community hospitals from true tertiary or regional perinatal centers (level 3C or 3D). We used the draft rather than the final version because the draft was a more accurate reflection of how hospitals in California were actually operating. The final version does not allow for NICUs that provide mechanical ventilation without re-
strictions but that do not provide major surgery (level 3B in the draft version); many California NICUs were actually providing this level of care during the study period.

We assigned levels of NICU care to each hospital, for each year, empirically from our data (see Section A-1 of the Supplementary Appendix, available with the full text of this article at www.nejm.org). For each year, we also counted the number of very-low-birth-weight infants who received care at each hospital (both those born in the hospital and those born elsewhere).

**NEONATAL AND FETAL DEATHS**

Because of improvements in neonatal care, the standard definition of a “neonatal death” — death within 28 days after birth — may be biased by the exclusion of continuously hospitalized infants who die after 28 days. Thus, we use the term “neonatal-related deaths” to refer to all neonatal deaths plus any deaths that occurred between 29 days and 1 year after delivery if the infant was continuously hospitalized. In 2000, deaths after 28 days accounted for 7.5% of all neonatal deaths.

Differences among hospitals in the level of obstetrical care, especially the ability to perform very rapid cesarean deliveries, can result in the live birth of infants who would otherwise die in utero. Thus, the exclusion of in-hospital fetal deaths would introduce a systematic bias against hospitals with large or high-risk obstetrical services. To arrive at a conservative estimate of the number of fetal deaths that occurred after the mother was admitted to the hospital, we identified in-hospital fetal deaths using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes (see Table A-2 of the Supplementary Appendix) from the mother’s discharge abstract for procedures that are performed continuously hospitalized infants who would otherwise die in utero. When added to the neonatal deaths, in-hospital fetal deaths account for 22.8% of total deaths.

Data on infants with a birth weight below 500 g (5157 infants) were excluded to be consistent with previous studies and because of the variability in decisions about whether to treat such infants. Because some congenital anomalies can increase the risk of death among infants with very low birth weight, we used ICD-9-CM codes to identify and exclude infants with such anomalies (7667 infants) (see Table A-3 of the Supplementary Appendix). We also excluded fetal deaths that we could not confirm had occurred after hospital admission (5777 fetal deaths), resulting in a final sample of 48,237 infants. A total of 6892 infants were transferred between hospitals; these infants remained in the sample, and deaths among them were attributed to the birth hospital.

**STATISTICAL ANALYSIS**

We used logistic regression to estimate odds ratios for mortality associated with the NICU level of care and annual volume of very-low-birth-weight infants. The dependent variable was in-hospital neonatal-related or fetal death. The standard errors for the hospital-level independent variables were corrected for within-hospital clustering with the use of the “cluster” option in Stata software, version 9.34 We controlled for the year to offset the decline in neonatal mortality over the course of the study period.27,28

Regressions run separately for each level of care showed that the effects of the volume of very-low-birth-weight infants varied according to the NICU level. For ease of presentation, we created categorical variables for the volume for each level of care.

We tested several different nonlinear functional forms using birth weight and gestational age but used categorical variables for the final model because they produced a better fit. We used separate birth-weight functions for male singletons, female singletons, and multiple births with 100-g intervals up to 1000 g and 250-g intervals from 1000 to 1500 g. For gestational age, we used 2-week intervals through 33 weeks.

We tested a wide range of clinical and demographic variables from the birth certificate and discharge data to control for risk factors, and we considered only those variables that were present at birth (see Table A-4 of the Supplementary Appendix). We developed the model using a random 50% sample and then validated it by applying the estimated coefficients to the remaining data. A Hosmer–Lemeshow test revealed an acceptable fit (P=0.13).35 When applied to the entire data set, the model again fit well (P=0.34 by the Hosmer–Lemeshow test; area under the receiver-operating-characteristic curve=0.86).

**RESULTS**

The number of NICUs increased slightly between 1991 and 2000, and there was a noticeable shift upward in the levels of care provided (Fig. 1). Most
of the new NICUs in California in the 1990s were low- or moderate-volume units, as were most of the NICUs that upgraded their level of service (Table A-1 of the Supplementary Appendix).

The percentage of very-low-birth-weight deliveries in hospitals with level 3B, 3C, and 3D NICUs that treated more than 100 such infants decreased from 35.6% in 1991 to 21.5% in 2000 (Table 1), with most of this decline offset by the increase in deliveries at hospitals with level 3B or 3C NICUs that treated 26 to 50 very-low-birth-weight infants annually. The percentage of very-low-birth-weight deliveries in NICUs that treated 51 to 100 of these infants was constant over this time, and there were only minor changes at NICUs with other levels of care.

There was a wide range in the unadjusted mortality rates among NICU level-of-care and volume groups (Table 2). Mortality decreased as patient volume increased within each level of care and with higher levels of care within each volume group. Table 2 also shows the distribution of several risk factors for death according to the level of care.

As compared with deliveries at hospitals with a level 3B, 3C, or 3D NICU that treated at least 100 very-low-birth-weight infants per year, deliveries at hospitals with lower-level and lower-volume NICUs were associated with an increased risk of death (Table 3), adjusted for the risk factors shown in Table 2. (Table A-5 of the Supplementary Appendix, which shows odds ratios for death associated with the other covariates in the model.) The odds ratios decreased as volume increased within each level of care and as the NICU level of care increased within each volume group. The risk of death was significantly higher in level 3B and 3C NICUs that treated 50 or fewer very-low-birth-weight infants per year than in units with larger volumes. The risk of death for NICUs with various combinations of lower levels of care and patient volumes were significantly increased, with the exception of two very small groups of hospitals: those with level 2 NICUs that treated more than 25 very-low-birth-weight infants (four hospitals in 2000) and level 3A NICUs that treated more than 50 such infants (three hospitals in 2000). Although the number of NICUs in these two groups was smaller than normal for categorical variables, model specification tests showed that they should not be combined with smaller NICUs with the same level of care. A NICU that treats 50 very-

**Figure 1. Number of NICUs, According to Level of Care, in California, 1991–2000.**

Levels of care were empirically determined by the authors on the basis of a modified version of American Academy of Pediatrics definitions. Level 2 denotes an intermediate-care NICU, with no mechanical ventilation; level 3A denotes mechanical ventilation with restrictions (e.g., only for infants whose birth weight is greater than 1000 g); level 3B denotes no restrictions on mechanical ventilation but no major surgery; level 3C denotes major neonatal surgery but no cardiac surgery and no extracorporeal membrane oxygenation (ECMO); and level 3D denotes cardiac surgery, ECMO, or both.
low-birth-weight infants per year corresponds to an average NICU census of about 15 patients. Thus, most of the increase in the risk of death was accounted for by hospitals with small to moderate-size NICUs.

Results were materially unchanged when we included infants with congenital anomalies and when we excluded in-hospital fetal deaths. We also performed an analysis stratified according to birth weight. Although the associations between mortality and NICU level and volume were greater for the smallest infants (below 1000 g), they were still significant for the larger infants (see Table A-6 of the Supplementary Appendix). The results of a model limited to infants born after 1995 were consistent with the overall results (data not shown). The very strong correlation between NICU volume and number of deliveries, as well as the lack of other data about obstetrical services, limited our ability to investigate the role of obstetrical factors. However, to address the possibility that obstetrical volume accounted for our results, we created a model based on estimates that included obstetrical volume; the effects on our results were minimal (see Table A-6 of the Supplementary Appendix).

We estimated the number of potentially preventable deaths on the basis of the odds ratios from Table 3 and the distribution of very-low-birth-weight deliveries across NICUs in 2000 from Table 1. Because the distance from the mother’s home to a hospital determines the feasibility of delivery at that hospital, this analysis was restricted to geographic areas with at least 100 very-low-birth-weight deliveries in 2000. Since most births occurred in the large urban areas of California, this restriction excluded only 8% of such deliveries. Assuming that only 90% of the deliveries of very-low-birth-weight infants in the large urban areas could be shifted to hospitals with tertiary-level NICUs that care for at least 100 such infants annually, we estimated that 21% of the deaths of very-low-birth-weight infants in the year 2000 were potentially preventable (see Table A-7 of the Supplementary Appendix).

**DISCUSSION**

Our study shows strong associations between both NICU level and volume at the delivery hospital and mortality. Our analysis of data from a 10-year period strengthens the evidence from previous studies that used data from a period of 1 or 2 years. Mortality was lowest when very-low-birth-weight deliveries occurred in hospitals with tertiary-level NICUs that treat more than 100 of these infants annually.

High-volume, high-level NICUs represent the minority of NICUs in California. Fewer than 25% of very-low-birth-weight infants were delivered in such hospitals in 2000, and the proportion of such deliveries occurring in these hospitals has declined over time.

Some limitations of our study should be noted. Because of the observational design, it was not possible to assess whether there was a causal relationship between NICU features and neonatal mortality. Factors other than NICU level and volume may explain the observed associations. For example, hospitals with large, high-level NICUs may also have better obstetrical care. The ability to provide rapid emergency cesarean sections not only prevents some fetal deaths but also results in the delivery of many infants in healthier condition, with associated improvement in survival. Our results were robust when obstetrical volume

<table>
<thead>
<tr>
<th>Level of Care and No. of Infants</th>
<th>1991</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>6.1</td>
<td>4.9</td>
</tr>
<tr>
<td>&gt;10</td>
<td>2.6</td>
<td>4.0</td>
</tr>
<tr>
<td>Level 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td>11–25</td>
<td>3.8</td>
<td>4.6</td>
</tr>
<tr>
<td>&gt;25</td>
<td>8.0</td>
<td>6.4</td>
</tr>
<tr>
<td>Level 3A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤25</td>
<td>1.2</td>
<td>0.9</td>
</tr>
<tr>
<td>26–50</td>
<td>2.3</td>
<td>4.5</td>
</tr>
<tr>
<td>&gt;50</td>
<td>3.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Level 3B or 3C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤25</td>
<td>2.4</td>
<td>3.5</td>
</tr>
<tr>
<td>26–50</td>
<td>7.8</td>
<td>19.1</td>
</tr>
<tr>
<td>Level 3B, 3C, or 3D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51–100</td>
<td>25.9</td>
<td>25.8</td>
</tr>
<tr>
<td>&gt;100</td>
<td>35.6</td>
<td>21.5</td>
</tr>
</tbody>
</table>

* The numbers of very-low-birth-weight infants are the total numbers treated at each hospital, including infants transferred to or received from other facilities. Levels of care were empirically determined on the basis of a modified version of American Academy of Pediatrics definitions. Level 2 denotes an intermediate-care NICU, with no mechanical ventilation; level 3A mechanical ventilation with restrictions (e.g., only for infants whose birth weight is greater than 1000 g); level 3B no restrictions on mechanical ventilation but no major surgery; level 3C major neonatal surgery but no cardiac surgery and no extracorporeal membrane oxygenation (ECMO); and level 3D cardiac surgery, ECMO, or both.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3A</th>
<th>Levels 3B and 3C</th>
<th>Levels 3B, 3C, and 3D†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–10</td>
<td>&gt;10</td>
<td>1–10</td>
<td>&gt;10</td>
<td>1–10</td>
</tr>
<tr>
<td>No. of infants</td>
<td>2636</td>
<td>1379</td>
<td>872</td>
<td>2270</td>
<td>4071</td>
</tr>
<tr>
<td>No. of deaths (%)</td>
<td>909 (34.5)</td>
<td>405 (29.4)</td>
<td>275 (31.5)</td>
<td>597 (26.3)</td>
<td>863 (21.2)</td>
</tr>
<tr>
<td>Birth weight, sex, and plurality (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singleton female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500–599 g</td>
<td>4.3</td>
<td>4.3</td>
<td>3.9</td>
<td>4.2</td>
<td>3.6</td>
</tr>
<tr>
<td>600–699 g</td>
<td>3.8</td>
<td>3.2</td>
<td>4.1</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>700–799 g</td>
<td>2.8</td>
<td>2.5</td>
<td>2.3</td>
<td>3.0</td>
<td>3.1</td>
</tr>
<tr>
<td>800–899 g</td>
<td>3.0</td>
<td>2.9</td>
<td>2.3</td>
<td>2.4</td>
<td>3.2</td>
</tr>
<tr>
<td>900–999 g</td>
<td>3.3</td>
<td>3.1</td>
<td>3.0</td>
<td>3.4</td>
<td>3.7</td>
</tr>
<tr>
<td>1000–1249 g</td>
<td>7.7</td>
<td>8.4</td>
<td>7.2</td>
<td>9.5</td>
<td>8.7</td>
</tr>
<tr>
<td>1250–1499 g</td>
<td>12.2</td>
<td>13.5</td>
<td>14.1</td>
<td>11.7</td>
<td>11.4</td>
</tr>
<tr>
<td>Singleton male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500–599 g</td>
<td>5.8</td>
<td>4.0</td>
<td>5.4</td>
<td>4.1</td>
<td>3.5</td>
</tr>
<tr>
<td>600–699 g</td>
<td>4.6</td>
<td>4.1</td>
<td>4.4</td>
<td>3.5</td>
<td>3.9</td>
</tr>
<tr>
<td>700–799 g</td>
<td>4.0</td>
<td>3.4</td>
<td>4.1</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>800–899 g</td>
<td>3.1</td>
<td>3.6</td>
<td>3.6</td>
<td>2.2</td>
<td>3.3</td>
</tr>
<tr>
<td>900–999 g</td>
<td>4.8</td>
<td>4.1</td>
<td>2.6</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>1000–1249 g</td>
<td>10.5</td>
<td>9.5</td>
<td>11.0</td>
<td>10.8</td>
<td>11.0</td>
</tr>
<tr>
<td>1250–1499 g</td>
<td>14.7</td>
<td>13.5</td>
<td>16.5</td>
<td>14.4</td>
<td>12.3</td>
</tr>
<tr>
<td>Multiple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500–599 g</td>
<td>1.4</td>
<td>1.0</td>
<td>0.9</td>
<td>1.9</td>
<td>1.7</td>
</tr>
<tr>
<td>600–699 g</td>
<td>1.0</td>
<td>1.7</td>
<td>1.8</td>
<td>1.6</td>
<td>2.0</td>
</tr>
<tr>
<td>700–799 g</td>
<td>1.1</td>
<td>1.7</td>
<td>0.9</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>800–899 g</td>
<td>1.0</td>
<td>1.4</td>
<td>1.1</td>
<td>0.7</td>
<td>1.9</td>
</tr>
<tr>
<td>900–999 g</td>
<td>1.0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.7</td>
<td>2.0</td>
</tr>
<tr>
<td>1000–1249 g</td>
<td>3.3</td>
<td>5.3</td>
<td>3.7</td>
<td>5.2</td>
<td>5.5</td>
</tr>
<tr>
<td>1250–1499 g</td>
<td>6.6</td>
<td>7.2</td>
<td>5.5</td>
<td>8.1</td>
<td>7.1</td>
</tr>
<tr>
<td>Gestational age (%)</td>
<td>&lt;24 wk</td>
<td>24–25 wk</td>
<td>26–27 wk</td>
<td>28–29 wk</td>
<td>30–31 wk</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>13.8</td>
<td>14.8</td>
<td>15.9</td>
<td>17.5</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>12.4</td>
<td>12.8</td>
<td>17.3</td>
<td>19.5</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td>14.0</td>
<td>12.6</td>
<td>15.6</td>
<td>16.9</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td>11.5</td>
<td>12.5</td>
<td>16.3</td>
<td>19.5</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>10.2</td>
<td>14.1</td>
<td>19.0</td>
<td>21.1</td>
<td>13.9</td>
</tr>
<tr>
<td></td>
<td>10.9</td>
<td>8.9</td>
<td>12.4</td>
<td>19.5</td>
<td>18.1</td>
</tr>
<tr>
<td></td>
<td>10.6</td>
<td>12.7</td>
<td>15.1</td>
<td>21.9</td>
<td>15.8</td>
</tr>
<tr>
<td></td>
<td>10.3</td>
<td>12.9</td>
<td>17.0</td>
<td>19.9</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>11.1</td>
<td>12.4</td>
<td>15.3</td>
<td>21.7</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>10.2</td>
<td>12.8</td>
<td>17.1</td>
<td>20.5</td>
<td>17.6</td>
</tr>
<tr>
<td></td>
<td>9.5</td>
<td>13.6</td>
<td>17.6</td>
<td>21.6</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td>8.9</td>
<td>13.1</td>
<td>17.5</td>
<td>22.4</td>
<td>17.3</td>
</tr>
</tbody>
</table>

| Black race (%)‡ | 8.9   | 12.2  | 15.4  | 13.3  | 21.0  | 14.6  | 18.5  | 12.7  | 10.6  | 16.2  | 20.7  | 18.4  |

| Maternal educational level (%)§ | 8.9   | 12.2  | 15.4  | 13.3  | 21.0  | 14.6  | 18.5  | 12.7  | 10.6  | 16.2  | 20.7  | 18.4  |

| Insurance (%) | Medicaid | 54.1 | 60.0 | 41.7 | 36.7 | 53.1 | 27.3 | 31.4 | 80.0 | 33.8 | 38.1 | 42.1 | 46.6 |
|              | Uninsured | 9.7  | 5.9  | 5.0  | 5.5  | 5.2  | 2.9  | 3.2  | 3.9  | 4.1  | 3.9  | 3.8  | 2.4  |
|              | Health maintenance organization | 15.2 | 17.5 | 28.3 | 35.2 | 23.1 | 46.0 | 48.9 | 7.7  | 33.2 | 34.9 | 39.3 | 25.3 |

| Fetal or infant condition (%) | Small for gestational age¶ | 8.5  | 8.2  | 13.2 | 10.2 | 9.7  | 8.9  | 7.7  | 3.6  | 10.0 | 11.1 | 9.8  | 8.8  |
|                               | Large for gestational age¶ | 11.8 | 9.4  | 12.6 | 11.0 | 18.2 | 18.1 | 14.4 | 15.0 | 13.3 | 15.5 | 19.7 | 17.9 |
|                               | Hydrops due to isoimmunization‖ | 0.34 | 0.07 | 0.00 | 0.00 | 0.05 | 0.00 | 0.14 | 0.07 | 0.07 | 0.15 | 0.11 | 0.17 |
|                               | Hemolytic disorder | 1.9  | 2.2  | 1.8  | 1.5  | 2.2  | 2.0  | 2.0  | 3.1  | 2.8  | 2.7  | 3.0  | 2.1  |
|                               | Fetal distress | 6.8  | 6.6  | 7.8  | 5.5  | 5.5  | 5.6  | 5.3  | 6.2  | 5.7  | 5.7  | 7.1  | 6.2  |
Table 2. (Continued.)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3A</th>
<th>Levels 3B and 3C</th>
<th>Levels 3B, 3C, and 3D†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–10 Infants</td>
<td>&gt;10 Infants</td>
<td>11–25 Infants</td>
<td>&gt;25 Infants</td>
<td>≤25 Infants</td>
</tr>
<tr>
<td>High-risk maternal condition (%)</td>
<td>6.6</td>
<td>6.8</td>
<td>9.6</td>
<td>8.4</td>
<td>8.1</td>
</tr>
<tr>
<td>Infant affected by maternal condition**</td>
<td>0.8</td>
<td>0.6</td>
<td>1.4</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Oligohydramnios</td>
<td>6.6</td>
<td>4.4</td>
<td>7.3</td>
<td>6.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Placental hemorrhage</td>
<td>3.8</td>
<td>6.0</td>
<td>5.5</td>
<td>5.6</td>
<td>6.1</td>
</tr>
<tr>
<td>Premature rupture of membranes</td>
<td>2.3</td>
<td>1.4</td>
<td>1.9</td>
<td>1.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Prolapsed cord</td>
<td>0.72</td>
<td>0.22</td>
<td>0.57</td>
<td>0.88</td>
<td>0.64</td>
</tr>
<tr>
<td>Other††</td>
<td>0.56</td>
<td>0.48</td>
<td>0.41</td>
<td>0.66</td>
<td>0.65</td>
</tr>
</tbody>
</table>

* Data include all in-hospital, very-low-birth-weight deliveries and fetal deaths (a total of 48,237). Infants with a birth weight below 500 g and infants with major congenital anomalies were excluded.
† No NICU with a 3D level of care had fewer than 51 patients.
‡ The Asian, Native American, and Hispanic categories had no significant effect on mortality and were excluded from the final model. Data on race or ethnic group were obtained from information recorded on the birth certificate, as reported by the mother.
§ The other educational categories (12 years and ≤4 years of college) had no significant effect on mortality and were excluded from the final model. Data on maternal education were obtained from information recorded on the birth certificate, as reported by the mother.
¶ Information on these conditions was based on ICD-9-CM codes for small for gestational age and large for gestational age (764 and 766, respectively).
‖ Nonimmune hydrops was classified as a congenital anomaly, and infants with this condition were excluded, along with infants who had other congenital anomalies. Hemolytic disease without a diagnosis of hydrops was included in the hemolytic-disorders variable.
** Data on maternal hypertensive disorders and noxious substances were based on ICD-9-CM codes 760.0, 760.7, 760.72, and 760.73.
††† Data on chronic maternal circulatory and respiratory diseases and incompetent cervix were based on ICD-9-CM codes 760.3 and 761.0, respectively.
was added to the model, but the role of obstetrical volume merits further investigation.

Although our model controlled for many potential confounders, we could address only those variables available from birth certificates and discharge abstracts. These data are of high quality, but they do not include information on all the potential differences in mortality. The risk factors we assessed did not differ significantly among the level-of-care and volume groups, but it is possible that unmeasured differences among the groups affected the results. Given that some high-risk cases are selectively referred to large tertiary-care centers, we would expect such factors to introduce a bias against the highest-level NICUs; consequently, they should not explain our findings. Our exclusion of infants with life-threatening congenital anomalies eliminated any bias due to referrals that were restricted to infants with treatable anomalies.

The only outcome we assessed was mortality; other outcomes, such as intraventricular hemorrhage and chronic lung disease, are also important. A recent study showed that a higher NICU volume was associated with a lower risk of intraventricular hemorrhage. However, the relationship between volume and outcomes other than mortality requires additional study.

Studies using data from the Vermont Oxford Network showed weaker relationships between NICU volume and mortality compared to what we observed. Further, in these analyses, volume explained less of the variance in mortality than it did in our study. One potential explanation for these differences is that our data included a broader sampling of hospitals, particularly community hospitals. Data from the Vermont Oxford Network also demonstrated considerable variation in outcomes across hospitals after taking the effects of NICU level and volume into account.

On the basis of our model, we estimated that increased regionalization of NICU care may have the potential to prevent 21% of deaths among very-low-birth-weight infants. This estimate relies on several assumptions, including a causal relationship between large, high-level NICUs and reduced mortality, and a very high level of regionalization. Our observation that 92% of the very-low-birth-rate deliveries in our study occurred in urban areas with more than 100 such deliveries suggests that it would be geographically feasible to regionalize the vast majority of these deliveries in California. To do so would probably require the addition of some large perinatal centers, which, ideally, would be strategically located to maximize geographic access and could be created through the mergers of existing smaller NICUs. There could be some disadvantages to closing facilities that are not included in our estimates, and efforts to increase regionalization are likely to draw some opposition. Although it would be more difficult to regionalize very-low-birth-weight deliveries in more sparsely populated areas of the United States, our results suggest that reductions in mortality could be achieved by moving from low to moderate volumes, which may be a more feasible goal in these areas.

In conclusion, our study showed that the NICU volume and level in the hospitals where very-low-

<table>
<thead>
<tr>
<th>Table 3. Odds Ratios for Mortality among Very-Low-Birth-Weight Infants, According to NICU Level of Care and Annual Patient Volume.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of Care and No. of Infants</strong></td>
</tr>
<tr>
<td><strong>Level 1</strong></td>
</tr>
<tr>
<td>( \leq 10 )</td>
</tr>
<tr>
<td>&gt;10</td>
</tr>
<tr>
<td><strong>Level 2</strong></td>
</tr>
<tr>
<td>( \leq 10 )</td>
</tr>
<tr>
<td>11–25</td>
</tr>
<tr>
<td>&gt;25</td>
</tr>
<tr>
<td><strong>Level 3A</strong></td>
</tr>
<tr>
<td>( \leq 25 )</td>
</tr>
<tr>
<td>26–50</td>
</tr>
<tr>
<td>&gt;50</td>
</tr>
<tr>
<td><strong>Level 3B or 3C</strong></td>
</tr>
<tr>
<td>( \leq 25 )</td>
</tr>
<tr>
<td>26–50</td>
</tr>
<tr>
<td><strong>Level 3B, 3C, or 3D</strong></td>
</tr>
<tr>
<td>51–100</td>
</tr>
<tr>
<td>&gt;100</td>
</tr>
</tbody>
</table>

* The area under the receiver-operating-characteristic curve was 0.86. The reference group was hospitals with a level 3B, 3C, or 3D NICU that treat at least 100 very-low-birth-weight infants per year. Standard errors were corrected for clustering of patients within hospitals. The model included birth weight, gestational age, sex, multiple gestation, black race, maternal educational level, type of insurance, year (2000 was the reference variable), several obstetrical conditions (premature rupture of the membranes, fetal distress, placental complications, polyhydramnios, and oligohydramnios), and fetal and neonatal conditions (small for gestational age, exceptionally large for gestational age, hydrops, and hemolytic disorders). Infants with major congenital anomalies or a birth weight below 500 g were excluded.
birth-weight infants are born is strongly associated with mortality; the mortality was lowest for deliveries that occurred in hospitals with high-level and high-volume NICUs. Less than a quarter of very-low-birth-weight infants are born in hospitals with such NICUs, and this percentage has been declining over time. Our results suggest that increased regionalization of perinatal care might reduce mortality among very-low-birth-weight infants.

Supported by a grant (HD-36914) from the National Institute of Child Health and Human Development and the Agency for Healthcare Research and Quality.

Dr. Baker reports receiving consulting fees from a consortium of hospitals in the state of Georgia, from Blue Cross of California (Wellpoint), and from Catholic Healthcare West and serving as an expert witness in legal matters related to certificate-of-need legislation. Dr. Phibbs reports serving as a scientific adviser for the Leapfrog Group’s evidence-based hospital referral program and as chairman of the neonatal subcommittee. No other potential conflict of interest relevant to this article was reported.

REFERENCES


34. Stata user’s guide. 8th ed. College Station, TX: Stata Press, 2003.

Copyright © 2007 Massachusetts Medical Society.
Bariatric Surgery for Morbid Obesity

Eric J. DeMaria, M.D.

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

A 44-year-old obese woman (height, 1.7 m [65 in.]) has seen her primary care physician for the past 10 years for management of conditions related to her obesity, including diabetes, hypertension, and gastroesophageal reflux disease. Despite efforts to lose weight, her body weight has increased from 109 to 127 kg (240 to 280 lb), and her body-mass index (BMI) — the weight in kilograms divided by the square of the height in meters — from 40.0 to 46.6. During a routine office visit, the patient asks her physician whether bariatric surgery might be a treatment option for her. The physician does not recommend referral for surgical evaluation, citing concerns about the variable effectiveness of the procedure and the associated risks, as well as the lack of long-term outcome data. The patient then seeks a specialist in bariatric surgery for evaluation, without the assistance of her physician.

The Clinical Problem

Obesity has become an epidemic condition in the United States and around the world. In the United States, the percentage of adults who are obese (defined as having a BMI of 30 or more) increased from 15.3% in 1995 to 23.9% in 2005.\(^1\) Approximately 4.8% are considered to be extremely or morbidly obese (having a BMI of 40 or more).\(^2\) Worldwide, it is estimated that more than 300 million people are obese.\(^3\)

Obesity, particularly abdominal obesity, is associated with increased risks of hypertension, diabetes, hyperlipidemia, sleep apnea, coronary heart disease, and stroke.\(^3,4\) In the United States, health care expenditures related to obesity and associated medical conditions amount to $100 billion annually,\(^5\) and in 2000, obesity was estimated to contribute to approximately 400,000 deaths.\(^6\) It has been suggested that in the 21st century, increasing rates of obesity may lead to a decline in overall life expectancy in the United States.\(^7\)

Pathophysiology and Effect of Therapy

The pathophysiology of obesity is complex and poorly understood, but it includes genetic, behavioral, psychological, and other factors.\(^8\) Family studies suggest that heredity may explain 67% of the population variance in BMI.\(^9\) However, genetic factors are unlikely to account fully for the rapid increase in the prevalence of obesity. Declining rates of physical activity\(^10\) and increases in the consumption of energy-dense foods\(^11\) may play a role.

Bariatric surgical procedures reduce caloric intake by modifying the anatomy of the gastrointestinal tract. These operations are classified as either restrictive or malabsorptive. Restrictive procedures limit intake by creating a small gastric reservoir with a narrow outlet to delay emptying. Malabsorptive procedures bypass vary-
Review identified only two small, randomized, controlled trials and three cohort studies, all of which were considered to have a high risk of bias in their design. Nonetheless, their summary assessments, as well as those of two meta-analyses, suggest a typical weight loss of 20 to 50 kg (44 to 110 lb) with various bariatric procedures as compared with a modest weight gain in medically treated patients.

The only large, well-controlled prospective study of bariatric surgery is the Swedish Obese Subjects (SOS) trial. A total of 2010 surgically treated obese patients (BMI, 34 or more for men and 38 or more for women) were compared with 2037 control subjects who were matched for 18 variables, including age, sex, weight, and several cardiac risk factors. Weight changes were significantly greater in the surgical group than in the control group among 3505 patients followed for 2 years (23.4% of body weight lost vs. 0.1% gained) and among 1268 patients followed for 10 years (16.1% of body weight lost vs. 1.6% gained). In a study of 1035 patients who underwent bariatric surgery, the mean BMI decreased from 50.0 to 32.6 at a median of 2.0 years of follow-up. In general, weight loss with malabsorptive procedures tends to be greater than weight loss with solely restrictive procedures.

Improvement in the conditions that are often associated with obesity has been consistently reported after bariatric surgery. In a meta-analysis by Buchwald et al., 77% of patients with preoperative diabetes no longer required medication after surgery. Similar improvements were seen for patients with hyperlipidemia (83%), hypertension (66%), and sleep apnea (88%). The SOS data suggest that some of these benefits are less marked at 10 years than at 2 years, although they are still significant.

It has not been clearly established whether bariatric surgery results in reduced mortality as compared with medical management of obesity, although such a benefit is suggested in the results of several matched cohort studies. Sjostrom has presented data from the SOS trial showing that unadjusted overall mortality was reduced by 31.6% in the surgery group as compared with the control group, which was a significant reduction.
The recommendations resulting from the conference included the following criteria for bariatric surgery: a BMI of 40 or higher, or a BMI of 35 or higher in a patient with a high-risk condition such as severe sleep apnea, obesity-related cardiomyopathy, or severe diabetes mellitus. Additional criteria included failure of medical weight control and an absence of medical or psychological contraindications, as well as the patient’s understanding of the procedure and its risks and strong motivation to comply with the postsurgical regimen. These criteria make clear the need for a multidisciplinary approach that includes medical, surgical, nutritional, and psychological assessment.

Evaluation of the surgical candidate should include a comprehensive nutritional and weight history, covering weight trends, previous weight-loss efforts, and perceived obstacles to successful weight management. Current weight, height, and BMI should be determined. Although measurement of waist circumference provides additional information regarding health risks, this information is not particularly useful in persons with a BMI above 40.

Secondary causes of obesity should be considered, although they do not usually account for severe obesity. Routine screening for Cushing’s syndrome and hypothyroidism is not necessary unless clinical suspicion is high. A medication history should be obtained; antidepressants, oral contraceptives, oral hypoglycemic agents, and other drugs can be associated with weight gain.

The medical evaluation should include assessment for the conditions that commonly accompany severe obesity, including diabetes, hypertension, hyperlipidemia, coronary artery disease, sleep apnea, pulmonary hypertension, and musculoskeletal disease. Careful selective investigation of these conditions serves several purposes. It facilitates optimal medical management before surgery, identifies problems that may influence the perioperative and postoperative course, and provides a baseline set of clinical data for evaluating the benefit of surgery.

The psychological evaluation of the candidate for bariatric surgery is one of the most important and difficult elements of the clinical assessment. A majority of patients presenting for bariatric procedures have one or more psychiatric disorders; some studies suggest that patients with a diagnosis of an Axis I or Axis II disorder (according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition) are likely to lose less weight after surgery than those without such diagnoses. Other psychosocial factors that have been associated with a suboptimal surgical outcome include disturbed eating habits (e.g., binge eating), substance abuse, low socioeconomic status, limited social support, and unrealistic expectations of surgery.

Patients undergoing bariatric surgery often believe they will lose more weight than is consistent with clinical experience and may think that minimal personal effort or risk is involved. Unfortunately, such ideas are increasingly fostered by inaccurate information obtained from unreliable sources on the Internet and elsewhere. Preoperative education is important in improving the patient’s understanding of the anticipated consequences of the procedure.

Specific contraindications to bariatric surgery are few. They include mental or cognitive impairment that limits the patient’s ability to understand the procedure and thus precludes informed consent. Very severe coexisting medical conditions, such as unstable coronary artery disease or advanced liver disease with portal hypertension, may in some instances render the risks of surgery unacceptably high.

Perioperative care of patients undergoing bariatric surgery requires specialized expertise and
C Roux-en-Y gastric bypass

D Biliopancreatic diversion with duodenal switch

- Common bile duct
- Roux limb
- Biliopancreatic limb
- Alimentary limb
- Common channel

- Stomach
- Common bile duct
- Biliopancreatic limb
- Alimentary limb
- Common channel
facilities. Studies have demonstrated that the likelihood of postoperative complications is significantly associated with annual surgical experience. The risks are greatest when surgeons perform fewer than 25 operations and hospitals host fewer than 50 operations per year, and the risks are lowest when surgeons perform more than 100 operations and hospitals host more than 150 operations per year.32-34

The University HealthSystem Consortium evaluated 1143 bariatric surgical procedures performed between October 2003 and March 2004 at 29 institutions in the United States.35 Roux-en-Y gastric-bypass operations accounted for 1049 (92%) of the procedures. Among the patients who underwent these procedures, the mean time in the operating room was 3.8 hours, 7.7% required intensive care, and the mean hospital stay was 3.5 days. Ninety-four (8%) of the procedures were restrictive. In the group of patients who underwent restrictive operations, the mean time in the operating room was 2.3 hours, only 1.1% required intensive care, and the mean hospital stay was 1.6 days.

A comprehensive plan for long-term patient care is necessary for bariatric surgery to have a reasonable chance of being both safe and successful. The operation should not be performed unless systematic follow-up is available and unless the patient has made a commitment to participate in such care. As in the preoperative evaluation, postoperative management requires a coordinated approach involving expertise in medicine, surgery, psychology, and nutrition.36 Unfortunately, many patients do not receive systematic postoperative care, and they may have suboptimal outcomes as a result.37,38

Estimates of the median hospital costs for bariatric surgery range from approximately $10,000 to $14,000.39 The Medicare physician fees for 2007 are $800 to $1,000 for laparoscopic adjustable gastric banding and $1,300 to $2,000 for Roux-en-Y gastric bypass and for biliopancreatic diversion with duodenal switch.40 Longer-term costs are more difficult to determine, but one analysis suggested that lifetime medical costs could be $59,000 to $75,000.41

### Adverse Effects

In several large series, the mortality rate associated with bariatric surgery was 0.1% to 2.0%.20,42-45 In the meta-analysis by Buchwald et al., operative mortality rates were 0.5% for gastric bypass, 0.1% for gastric banding, and 1.1% for malabsorptive procedures.15 Common causes of death among patients undergoing bariatric surgery include pulmonary embolism and anastomotic leaks. Factors that have been found to contribute to increased mortality include lack of experience on the part of the surgeon or the program, advanced patient age, male sex, severe obesity (BMI ≥50), and coexisting conditions.20,32,42,44-50

Nonfatal perioperative complications include venous thromboembolism, anastomotic leaks, wound infections, bleeding, incidental splenectomy, incisional and internal hernias, and early small-bowel obstruction. In the SOS trial, postoperative complications occurred in 13% of patients, including bleeding in 0.5%, embolism or thrombosis in 0.8%, wound complications in 1.8%, and pulmonary complications in 6.1%.37

Postoperative gastrointestinal complications of bariatric surgery are common. Nausea and vomiting occur in more than 50% of patients undergoing restrictive procedures, partly as a result of eating too much or too rapidly but sometimes because of anastomotic stricture or other mechanical consequences of the operation.36 The dumping syndrome, a complex of neurohormonally mediated symptoms that include facial flushing, lightheadedness, palpitations, fatigue, and diarrhea, occurs in as many as 70% of patients after Roux-en-Y gastric bypass.24 Typically triggered by the ingestion of concentrated sugar, this syndrome may discourage patients from eating foods with a high sugar content, thus contributing to the beneficial effects of the operation.51 Deficiencies of iron, calcium, folate, vitamin B₁₂, and other nutrients occur after procedures with a component of malabsorption, such as gastric bypass. With the more extensive procedures, such as biliopancreatic diversion, protein malnutrition and deficiencies of the fat-soluble vitamins (A, D, E, and K) may occur. All of these deficiencies require regular monitoring and replacement.24,36 Other gastrointestinal complications include dehydration, bowel obstruction, anastomotic leaks, strictures, erosions, ulcers, adhesions, internal and incisional hernias, and cholelithiasis.16,36,52

Patients who have undergone bariatric surgery may require subsequent readmission or reoperation. In the study by the University HealthSystem Consortium, the rate of repeat operation was...
The changing popularity of specific bariatric surgical procedures over time suggests that the ideal procedure has not been definitively established. In the United States, the Roux-en-Y gastric bypass (open or laparoscopic) is the most common operation, but in Europe, laparoscopic adjustable gastric banding is performed more frequently. Comparative studies indicate that laparoscopic gastric bypass is similar to open surgery in terms of weight loss, with fewer complications and less postoperative pain, but specific training in laparoscopic techniques is required.

**Areas of Uncertainty**

The potential benefit of bariatric surgery for patients with mild obesity (BMI of 30 to 35) remains unclear. In a recent randomized, controlled trial, patients in this range who underwent laparoscopic adjustable gastric banding had more weight loss, greater resolution of the metabolic syndrome, and greater improvement in the quality of life than did patients who received medical therapy. However, this remains an area of controversy. It is likewise uncertain whether patients with extremely severe obesity are appropriate candidates for bariatric surgery. There are reasons for concern that the operative risk may be increased for such patients, partly because of technical difficulties in performing the procedure and practical management issues. At least one report has suggested that mortality rates may be increased among patients with a BMI of 70 or more. However, there is also a lack of comparison data with which to determine risk and longevity in patients who have not had surgery.

The role of bariatric procedures in patients outside the commonly defined age range (18 to 60 years) is not well established. It has been demonstrated that the levels of risk and benefit among obese adolescents who undergo surgery are similar to those among older patients. However, many younger patients may not have sufficient insight to appreciate the consequences of the decision to undergo surgery or to cooperate fully with follow-up care. In addition, the long-term consequences of surgery are less clear in this population. For the elderly, at least one report has indicated that the benefits in terms of weight loss and improvement in coexisting conditions, although significant, are not as great as in younger patients.

**Guidelines**

Most clinical guidelines regarding the role of bariatric surgery have followed the lead of the 1991 NIH Health Consensus Development Conference in concluding that such procedures should be considered for patients who have a BMI of 40 or more or for those who have a BMI of 35 or more with coexisting medical conditions. These criteria are endorsed by the National Heart, Lung, and Blood Institute in guidelines for the treatment of obesity published in 1998; they are also endorsed in more recent guidelines published by the Institute for Clinical Systems Improvement, the American Society for Bariatric Surgery, the European Association for Endoscopic Surgery, and other organizations. The American College of Physicians has adopted a somewhat more conservative approach, recommending that surgery be considered only in patients with a BMI of 40 or more who also have coexisting conditions. All these guidelines generally concur that patients should have made previous attempts to lose weight, should be free of medical and psychological contraindications, and should be cared for by a multispecialty team with experience in bariatric surgery and perioperative care.

**Recommendations**

The patient described in the vignette is a candidate for bariatric surgery on the basis of her BMI and coexisting medical conditions. She should be evaluated by an experienced surgeon at a center with established expertise in bariatric procedures and should undergo a comprehensive medical, surgical, nutritional, and psychological assessment. It is important that her expectations for surgery be discussed in advance and that she re-
ceive full information about the anticipated risks and results of the operation. She should be required to make a commitment to an appropriate postoperative regimen of diet, exercise, and medical and surgical follow-up care. On the basis of my own experience, I would recommend that she undergo a laparoscopic Roux-en-Y gastric bypass, with some discussion and consideration of other surgical options. The detailed plan for her care should be discussed with her primary care physician, who should be recruited in the effort to provide the patient with appropriate preoperative and postoperative medical and psychological support.

Dr. DeMaria reports receiving consulting fees from Power Medical Interventions and grant support for the Duke Endosurgery Center from Tyco Healthcare and Stryker. No other potential conflict of interest relevant to this article was reported.

REFERENCES


39. Livingston EH. Hospital costs associated with bariatric procedures in the...
A healthy 65-year-old man presented with a 4-month history of left flank pain and nocturia. Dermatologic examination revealed multiple eruptive seborrheic keratoses, which he reported had developed over the previous 1 to 2 years (Panels A and B). Abdominal ultrasound examination and computed tomography (CT) revealed a localized mass (7 cm by 5 cm in diameter) of the left lower renal pole with central necrosis. A left total nephrectomy was performed, and histopathological examination confirmed a renal-cell carcinoma. The tumor infiltrated to, but did not penetrate, Gerota’s fascia. The cutaneous findings were consistent with the diagnosis of the Leser–Trélat sign, which is usually associated with gastrointestinal adenocarcinoma. This sign is a controversial physical finding, however, since seborrheic keratoses are common with aging. Four months after the initial diagnosis, thoracic CT showed multiple metastatic lesions in the lung, for which the patient received immunotherapy and chemotherapy. He continues to receive treatment with sunitinib.

Copyright © 2007 Massachusetts Medical Society.
Presentation of Case

A 61-year-old man was referred to the thoracic oncology service of this hospital for management of a thymoma. He had been well until 6 weeks earlier, when he experienced the sudden onset of sharp left anterior chest pain, which was worse adjacent to the sternum and when he took a deep breath. He went to the emergency room of another hospital. Computed tomographic (CT) scanning of the chest performed with a pulmonary-embolism protocol revealed no evidence of pulmonary embolism. However, a lobulated, soft-tissue mass, 4 cm in diameter, was found in the left anterior mediastinum, adjacent to the main pulmonary artery and left main pulmonary artery and contiguous with the pericardium. There was no mediastinal or hilar lymphadenopathy, and the lungs were clear. He was referred to a thoracic surgeon at that hospital.

A whole-body positron-emission tomographic (PET) scan obtained 2 weeks after the onset of symptoms (Fig. 1) showed increased uptake of tracer in the area of the lesion, with no other areas of increased uptake. A bone scan was negative. The results of pulmonary-function testing were consistent with moderate mixed obstructive and restrictive disease, with a forced expiratory volume in 1 second of 1.96 (56% of the predicted value) and a forced vital capacity of 3.12 (59% of the predicted value). A bronchoscopic examination performed 1 month after the onset of symptoms showed no endobronchial lesions. Mediastinoscopy was performed, and examination of biopsy specimens of four mediastinal lymph nodes with frozen sections revealed no malignant tumor. A left anterior, mediastinal exploration with biopsy (Chamberlain procedure) was then performed; there was a poorly circumscribed mass in the aortopulmonary window that appeared to invade the pericardium. Biopsy specimens were obtained, and a diagnosis was made of a predominantly cortical, type B1 thymoma, according to the World Health Organization (WHO) criteria. Cytologic examination of specimens of left-lung washings were negative for tumor. The patient was referred to this hospital for further treatment.

His chest pain had resolved spontaneously, and he did not have cough, dyspnea, sputum production, fever, or muscle weakness. There was no history of myasthenia gravis, pure red-cell aplasia, or hypogammaglobulinemia. The patient was obese and had hypertension and hypercholesterolemia. He had undergone parathyroidectomy for hyperparathyroidism and had been treated for Lyme disease; the Guillain–Barré
syndrome had developed after a tetanus toxoid injection, and he had undergone neck irradiation for pertussis in childhood. He had smoked cigarettes for 25 years and had quit smoking 20 years earlier. He was a retired university professor, with no unusual occupational exposures, and he did not have an unusual travel history. His medications were lisinopril, simvastatin, and gabapentin.

On physical examination, the patient’s vital signs were normal. He had a body-mass index (the weight in kilograms divided by the square of the height in meters) of 37. There was no jugular venous distention. There was no cervical or supraclavicular lymphadenopathy. The chest was clear on auscultation. There was no tenderness on palpation of the ribs or sternum. The remainder of the examination was normal.

A decision regarding management was made.

**DIFFERENTIAL DIAGNOSIS**

Dr. Jo-Anne O. Shepard: The initial examination, a CT study performed with intravenous contrast material (Fig. 1A and 1B), did not show a pulmonary embolism. A soft-tissue mass was identified in the left anterior mediastinum. It was noncalcified, with soft-tissue attenuation, and was somewhat lobulated in configuration. The mass was contiguous with the main pulmonary artery and the pericardium. There was no specific, direct evidence of vascular invasion, but the question of pericardial invasion was raised.

Whole-body PET (18F-fluorodeoxyglucose [FDG]) scanning (Fig. 1C), performed shortly after the initial examination, showed a focus of FDG uptake in the anterior mediastinum that corresponded with the CT finding. The increased uptake on the PET scan suggested a malignant lesion.

The radiologic differential diagnosis at this point would include a thymic neoplasm, a lymphoma or a germ-cell neoplasm, and — potentially but less likely — a solitary focus of metastatic disease to the mediastinum. We did not see other sites of abnormal lymphadenopathy either on the CT scan or on the PET scan.

Dr. Cameron D. Wright: An anterior mediastinal mass was discovered incidentally while this patient was undergoing evaluation for acute chest pain. The relationship between the pain and the tumor that was eventually identified is not clear. The original investigation focused on possible lung cancer with hilar lymphadenopathy, despite the absence of an obvious lung lesion. Although in rare cases lung cancer presents in this manner, a careful review of the CT scan suggests that the mass is probably an anterior mediastinal lesion.

Figure 1. Chest CT and Whole-Body PET Scans Obtained before Admission.

A contrast-enhanced CT scan (Panel A) reveals a lobulated, soft-tissue mass in the left anterior mediastinum (arrow) lateral to the main and left pulmonary arteries. At a slightly more caudal level (Panel B), the mass (arrow) is contiguous with the pericardium (arrowhead). An 18F-fluorodeoxyglucose PET scan (Panel C) shows intense focal uptake within the mass (arrow), a finding that is consistent with a metabolically active lesion and suggests a malignant condition.
rather than lymphadenopathy. The mass is smooth in contour and homogeneous, and it does not look like a cluster of lymph nodes.

The most likely causes of an anterior mediastinal mass in a patient of this age are lymphoma and thymoma. Given the appearance of the mass on CT, the age of the patient, and the absence of symptoms, the most likely diagnosis is thymoma. My first choice for a diagnostic procedure in this situation would have been a CT-guided core needle biopsy; if this procedure had been nondiagnostic, I would have performed an anterior mediastinotomy with an incisional biopsy, which was the diagnostic procedure performed in this case.

Examination of the biopsy specimen showed that this tumor was a thymoma. Although thymomas are the most common mediastinal tumors, accounting for about 50% of all anterior mediastinal tumors, their overall incidence is very low, with only about 1.5 cases per million persons per year.\(^1\) The usual age at the time of diagnosis is the sixth decade. The sex distribution is equal. About one third to one half of the patients are symptomatic.\(^2\) Most symptoms are due to pressure from the tumor on surrounding structures or invasion of these structures, and include chest pain, dyspnea, cough, and the superior vena cava syndrome. Paraneoplastic syndromes, which are common, include myasthenia gravis, occurring in about one third of patients; hypogammaglobulinemia; and red-cell aplasia. Patients with thymomas are at increased risk for the development of another malignant tumor.\(^3\) This patient has no evidence of a paraneoplastic syndrome or another malignant condition and does not have symptoms related to the mass.

### Pathological Discussion

**Dr. Robert P. Hasserjian:** Multiple biopsy specimens were obtained from the mediastinal tumor and tissue from the ascending aorta, pulmonary artery, and pericardium. Examination of all the specimens showed a tumor composed predominantly of small lymphoid cells with the immunophenotype of cortical thymocytes (Fig. 2). Thymic epithelial cells were scattered throughout the tumor and were present in increased numbers within vaguely nodular areas that were histologically reminiscent of thymic medulla. These features are diagnostic of a thymoma.

Thymomas are neoplasms of thymic epithelial cells\(^4\) that vary greatly in both their morphologic features and their tendency to invade or recur. This histologic and biologic variability has resulted in controversy regarding their classification.\(^5-8\) The WHO recently proposed a classification system for thymoma based on the morphologic characteristics of the neoplastic epithelial cells and their relationship to the thymic lymphocytes (Table 1).\(^9\) In this case, the presence of numerous thymic T cells and pale areas recapitulating medullary differentiation were diagnostic of a type B1 thymoma.

The diagnosis, classification, and staging of thymoma may be difficult with small biopsy samples, such as those that were obtained in this case. The presence of thymocytes can mimic either normal thymus or a lymphoblastic lymphoma. Classification can be inaccurate because thymomas are often heterogeneous, and diverse components may not have been sampled. A B1 thymoma such as this one may occur as a pure histologic type or it may be associated with areas of type A thymoma, in which case a diagnosis of type AB thymoma is made; AB thymomas are associated with a lower risk of invasion or recurrence than type B1 thymomas (Table 1).\(^10-12\) Less often, a B1 thymoma is seen with a B2 thymoma, which is a more aggressive tumor. In this case, because of the small biopsy samples, the presence of areas of type A or type B2 thymoma elsewhere in the tumor cannot be ruled out. Finally, the biopsy samples were too small to assess capsular invasion or invasion of adjacent structures, and the findings were thus inconclusive with regard to tumor stage. Four specimens of mediastinal lymph nodes were negative.

### Discussion of Management

**Staging and Resectability of Thymoma**

**Dr. Wright:** My most important task as a thoracic surgeon assessing a patient with a thymoma was to determine whether the tumor was amenable to a complete resection and its likelihood of recurrence after surgery alone. Stage, WHO histologic type, tumor size, and status with regard to complete resection are independent predictors of the prognosis in thymoma.

The most commonly used staging system for thymoma is the postsurgical system devised by Masaoka in 1981\(^13\) (Table 2). Unfortunately, it is difficult to assign a stage preoperatively because of the difficulty of assessing tumor invasion on
CT scans. When the tumor is small (<5 cm), is mostly surrounded by fat, and does not abut the lung, hilum, or great vessels, the surgeon can predict with confidence that a complete resection can be done. Extensive abutment of the pericardium or lung without a fat plane means that it is difficult to be sure that there is no microscopical invasion of these structures. The presence of symptoms is a good predictor of invasion and should be weighed in deciding on the probable stage of the tumor.

In this case, the tumor is smaller than 5 cm in diameter, but the CT scan indicates that there may be invasion of the pericardium.

The WHO histologic type of thymoma can help predict resectability (Table 1). Type A thymomas are almost always noninvasive, so if this tumor were type A, I would be confident that it could be resected, regardless of its size. Both type A and type AB thymomas rarely recur and seldom require additional therapy after resection, even if focal invasion is present. Type B thymomas (B1, B2, and B3) are often invasive, and a B3 tumor, in particular, is likely to be at an advanced stage and unresectable. This patient’s thymoma was a B1 tumor, which has an intermediate likelihood of invasion or recurrence.

Finally, the size of the tumor can help predict the likelihood that the tumor will be resectable and the chance of recurrence. A diameter larger than 8 cm is an independent predictor of recurrence. This patient’s tumor is 4 cm in diameter, and tumors of this size are often resectable; however, the surgeon who performed the biopsies was convinced that it was invading the pericardium.

Stage I and II thymomas are almost always

---

**Figure 2. Biopsy Specimens of Mediastinal Tumor.**

A low-magnification view (Panel A, hematoxylin and eosin) shows sheets of small lymphocytes with interspersed paler areas (arrow) resembling thymic medulla; the inset shows normal thymus and thymic medulla (arrows). On high magnification (Panel B, hematoxylin and eosin), large epithelial cells with delicate chromatin and small nucleoli (arrows) are scattered among the predominant population of small lymphocytes. Immunoperoxidase staining (Panels C and D) shows that the majority of the cells within the tumor are cortical thymocytes expressing terminal deoxynucleotidyl transferase (TdT), an immature lymphoid marker (Panel C); the neoplastic thymic epithelial cells are negative for TdT (Panel C, arrows) but express cytokeratin protein (Panel D, arrows).
completely resectable and cured with surgery alone, whereas stage III tumors may recur, most commonly in the ipsilateral pleural space from droplet metastases (Table 2). Thus, stage I and II lesions require no adjuvant therapy, whereas many stage III thymomas have positive margins and require postoperative therapy. I discussed the intraoperative findings in this patient with the referring surgeon; he thought that the thymoma invaded at least the pericardium and was thus a stage III tumor. Therefore, I was concerned that the tumor would not be completely resectable.

A decade ago, the standard treatment for thymoma was resection in all patients whose tumors appeared to be resectable on the basis of the preoperative evaluation; if the tumor proved to be invasive or could not be completely resected, radiotherapy was offered as an adjuvant postoperative treatment. This approach led to incomplete resection in many patients, and about one third of patients with stage III tumors, such as this one, had a recurrence. This observation, together with the favorable experience with induction chemotherapy for lung cancer, led many centers to recommend induction therapy before attempted resection of invasive thymomas, and my colleagues and I follow this approach at this hospital.

Although there are no randomized, controlled trials showing that induction therapy improves survival among patients with locally advanced thymomas, multiple case series have suggested an improvement in survival with induction therapy before surgery. This patient had a type B thymoma that was clinically estimated to be stage

### Table 1. Histologic and Clinical Features of Thymomas According to the WHO Classification.

<table>
<thead>
<tr>
<th>WHO Type</th>
<th>Epithelial Cells</th>
<th>Lymphocytes</th>
<th>Invasive Resection</th>
<th>Complete Stage</th>
<th>Recurrence</th>
<th>20-Yr Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Spindle-cell morphologic features, resembling medullary epithelial cells</td>
<td>Sparse; mature medullary thymocyte type</td>
<td>11</td>
<td>100</td>
<td>1.2</td>
<td>0</td>
</tr>
<tr>
<td>AB</td>
<td>Mixed type A and type B features</td>
<td>Mixed type A and type B features</td>
<td>42</td>
<td>99</td>
<td>1.5</td>
<td>5</td>
</tr>
<tr>
<td>B1</td>
<td>Sparse; both cortical and medullary type, recapitulating thymic architecture</td>
<td>Predominant; immature cortical thymocyte type, with areas of mature medullary thymocyte type</td>
<td>47</td>
<td>95</td>
<td>1.7</td>
<td>9</td>
</tr>
<tr>
<td>B2</td>
<td>More numerous than in type B1; oval nuclei with prominent nucleoli and indistinct cytoplasm, resembling cortical epithelial cells</td>
<td>Predominant; immature cortical thymocyte type</td>
<td>69</td>
<td>91</td>
<td>2.3</td>
<td>18</td>
</tr>
<tr>
<td>B3</td>
<td>Predominant; oval, often grooved nuclei and clear cytoplasm with distinct cell borders; cytologically atypical</td>
<td>Sparse; immature cortical thymocyte type</td>
<td>85</td>
<td>92</td>
<td>2.5</td>
<td>29</td>
</tr>
</tbody>
</table>

* The information is from Marx et al. and Okumura et al.

### Table 2. The Masaoka Staging System and Results of Thymoma Treatment in 1320 Patients.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Definition</th>
<th>Complete Resection</th>
<th>Recurrence</th>
<th>5-Yr Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Completely encapsulated tumor</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>II</td>
<td>Tumor that invades adjacent thymus, mediastinal fat, or mediastinal pleura</td>
<td>100</td>
<td>4</td>
<td>98</td>
</tr>
<tr>
<td>III</td>
<td>Tumor that invades surrounding structures such as lung, pericardium, or great vessels</td>
<td>85</td>
<td>28</td>
<td>89</td>
</tr>
<tr>
<td>IVA</td>
<td>Tumor with pleural or pericardial metastases</td>
<td>42</td>
<td>34</td>
<td>71</td>
</tr>
<tr>
<td>IVB</td>
<td>Tumor with lymphogenous or hematogenous metastases</td>
<td>NA</td>
<td>34</td>
<td>52</td>
</tr>
</tbody>
</table>

* The information is from Kondo and Monden. NA denotes not applicable.
III because of pericardial invasion, so I recommended induction therapy and referred him to the medical oncology department.

**INDUCTION CHEMOTHERAPY FOR STAGE III THYMOMA**

**Dr. Panos Fidias:** In this patient with stage III thymoma, consideration of induction chemotherapy was based on experience with unresectable thymomas. The administration of corticosteroids alone may lead to a response, and corticosteroids in combination with octreotide are associated with response rates in the range of 30% among patients with unresectable thymomas. Combination chemotherapy in a patient with advanced, unresectable thymoma usually involves cisplatin-based regimens; however, responses have been reported with other regimens, such as cyclophosphamide, doxorubicin, and vincristine (CAV) or CAV with the addition of prednisone, with or without bleomycin. Typically, responses are seen in 30 to 90% of patients, and the median survival ranges from 2 to 4 years.

More recently, studies evaluating systemic treatment for locally advanced tumors that may not be completely resectable have shown that induction therapy with a combination of cyclophosphamide, doxorubicin administered by means of continuous infusion, cisplatin, and prednisone or cisplatin and etoposide with either preoperative or postoperative radiotherapy resulted in high overall response rates, with improved tumor resectability and progression-free survival.

On the basis of these studies, we recommended preoperative chemotherapy with two cycles of cisplatin (33 mg per square meter of body-surface area on days 1, 2, and 3) and etoposide (100 mg per square meter on days 1, 2, and 3), and we asked Dr. Choi from the radiation oncology department to consider concomitant radiotherapy. We planned to administer two additional cycles of the same drugs after the surgical removal of the tumor.

**PREOPERATIVE RADIOTHERAPY IN STAGE III THYMOMA**

**Dr. Noah C.H. Choi:** The conventional approach of surgery and postoperative radiotherapy for patients with stage III thymoma has resulted in a 5-year survival rate of about 60%, without noticeable improvement over the past 20 years. Even after postoperative radiotherapy, patients with incomplete resection have a 5-year survival rate of 20 to 40%; 50% of the recurrences are in the pleural cavity. However, radiotherapy with or without chemotherapy results in 5-year survival rates of 30 to 50% for patients with inoperable thymomas, indicating that some tumors can be controlled by these treatment approaches.

Furthermore, a study has shown that patients with marginally resectable stage III thymomas who received radiotherapy preoperatively had improved recurrence-free and overall survival as compared with those who received radiotherapy after incomplete resection. These findings suggest that preoperative radiotherapy or chemoradiotherapy may increase survival by improving the rate of complete resection and reducing local and pleural recurrences.

Since this patient’s tumor was close to the heart, the esophagus, and the spinal cord, we used intensity-modulated radiotherapy (see Fig. 1 of the Supplementary Appendix, available with the full text of this article at www.nejm.org). With this technique, we can plan the radiation dose to fit the target volume and spare the adjacent normal organs and tissues better than we can with standard three-dimensional radiotherapy. In this patient, intensity-modulated radiotherapy consisted of a total dose of 45 Gy administered in 25 fractions over a period of 5 weeks. The radiotherapy was given concomitantly with chemotherapy.

**Dr. Shepard:** After the patient had received chemotherapy and radiotherapy, repeated CT scanning with contrast material (Fig. 3) and whole-body PET–CT scanning were performed. The CT scan showed a marked decrease in the size of the mass, and there was more fat interspersed between the mass and pericardium as compared with previous examinations. The PET–CT scan showed no evidence of abnormal FDG uptake.

**SURGICAL MANAGEMENT OF THYMOMA**

**Dr. Wright:** The patient had an excellent response to induction treatment, with the PET scan showing resolution of the hypermetabolic focus and the CT scan showing a marked reduction in the size of the tumor. Although complete resolution of FDG uptake on a PET scan suggests an absence of gross viable tumor, it does not necessarily indicate a complete absence of morphologically viable tumor cells. Extrapolation from other thoracic cancers (esophageal and lung cancers) suggests that normal uptake on PET scanning after induction therapy indicates a favorable prognosis. However,
since the radiographic response is a rather poor discriminator of the response to therapy, most thoracic cancers that are treated with induction therapy are subsequently resected, regardless of the degree of radiographic response. I recommended that this patient undergo a mediastinotomy and resection of the residual mass and thymus gland.

At surgery, I found a heavily scarred, residual mass adherent to the pericardium, projecting into the left pleural space and densely adherent to the left phrenic nerve. I removed all mediastinal fatty tissue, encompassing the entire thymus gland, and I resected the pericardium underneath the thymus gland and around the mass. There was no evidence of tumor invasion through the pericardium. I did not want to resect the left phrenic nerve, since the patient had preexisting lung disease and I was concerned about impaired lung function if the diaphragm was paralyzed. I dissected the phrenic nerve from the edge of the residual mass, completed the resection, marked the phrenic-nerve margins with metallic clips in case we thought more radiotherapy would be beneficial, and oriented the specimen for the pathologist. A frozen section of the area close to the phrenic nerve was negative for tumor. The patient had an uneventful recovery and was discharged to his home after 5 days in the hospital.

**PATHOLOGICAL FINDINGS**

*Dr. Hasserjian:* The resection specimen contained a small area of tumor (Panel A, hematoxylin and eosin) resembling the thymoma in the previous biopsy specimens. However, most of the tumor (Panel B, hematoxylin and eosin) is characterized by dense fibrosis with scattered epithelial cells almost entirely devoid of lymphocytes (inset), unlike the original tumor. These changes probably reflect the effects of adjuvant therapy.
background (Fig. 4B). Histologic changes in thymoma after therapy are not well documented in the literature, although tumor necrosis and reduced proliferation of neoplastic cells were reported in one study. In the current case, most of the residual tumor did not resemble any particular thymoma type and probably represented involutional changes in the tumor that were induced by the chemotherapy and radiotherapy. The tumor extended focally to the lateral resection margin on permanent sections; it did not invade the attached pericardium.

Dr. Wright: The patient received two additional cycles of chemotherapy and is well 4 years after the diagnosis, with no evidence of recurrent disease on serial CT scans. Long-term follow-up is required for patients with thymomas, since recurrences are routinely seen more than 10 years after treatment.

Dr. Harris: Was additional radiotherapy considered because of the positive margin?

Dr. Choi: We considered a postoperative booster dose of radiation to the tumor-bed region where the resection margin was positive for microscopic residual disease. However, the patient remained short of breath for several months after surgery. Thus, no additional radiotherapy was given.

ANATOMICAL DIAGNOSIS

Thymoma, WHO type B1, with individual changes after preoperative chemoradiation.

Dr. Wright reports receiving consulting fees fromClosure; and Dr. Fidias, lecture fees from Genentech. No other potential conflict of interest relevant to this article was reported.

REFERENCES


LANTERN SLIDES UPDATED: COMPLETE POWERPOINT SLIDE SETS FROM THE CLINICOPATHOLOGICAL CONFERENCES

Any reader of the Journal who uses the Case Records of the Massachusetts General Hospital as a teaching exercise or reference material is now eligible to receive a complete set of PowerPoint slides, including digital images, with identifying legends, shown at the live Clinicopathological Conference (CPC) that is the basis of the Case Record. This slide set contains all of the images from the CPC, not only those published in the Journal. Radiographic, neurologic, and cardiac studies, gross specimens, and photomicrographs, as well as unpublished text slides, tables, and diagrams, are included. Every year 40 sets are produced, averaging 50-60 slides per set. Each set is supplied on a compact disc and is mailed to coincide with the publication of the Case Record.

The cost of an annual subscription is $600, or individual sets may be purchased for $50 each. Application forms for the current subscription year, which began in January, may be obtained from the Lantern Slides Service, Department of Pathology, Massachusetts General Hospital, Boston, MA 02114 (telephone 617-726-2974) or e-mail Pathphotoslides@partners.org.
Government in Medicine
Jeffrey M. Drazen, M.D.

Although I am not a provider of reproductive medical services, I was alarmed to read of the recent decision of the Supreme Court in Gonzales v. Carhart. Why should I feel so concerned? The practical consequences of the “partial-birth abortion” bill are so far from my medical practice in pulmonary and intensive care medicine that the ruling should have no impact on me. Indeed, since most health care practitioners will not be directly affected by this decision, why should we care at all? It is because, as Charo1 and Greene2 point out in this issue of the Journal, with this decision the Supreme Court has sanctioned the intrusion of legislation into the day-to-day practice of medicine.

In 2005, we all saw the disastrous consequences of congressional interference in the case of Terri Schiavo. In that case, the courts wisely decided that Congress should not be practicing medicine. They correctly ruled that wrenching medical decisions should be made by those closest to the details and subtleties of the case at hand. Such decisions must be made on an individual basis, with the best interests of the patient foremost in the practitioner’s mind.

It is not that physicians do not want oversight and open discussion of delicate matters but, rather, that we want these discussions to occur among informed and knowledgeable people who are acting in the best interests of a specific patient. Government regulation has no place in this process. In 1997, another editor of the Journal, Jerome Kassirer, took Congress to task for practicing medicine without a license.3 He cited a number of instances, including the passage of a forerunner of the bill that the Supreme Court upheld last week. With Gonzales v. Carhart, the judicial branch has regretfully joined the legislative branch in practicing medicine without a license.

Aspirin and Colon Cancer — Targeting Prevention?
Sanford D. Markowitz, M.D., Ph.D.

The compelling evidence that chronic use of aspirin or certain nonsteroidal antiinflammatory drugs (NSAIDs) can substantially lower the risk of colon cancer has important implications, especially because colon cancer is the second leading cause of cancer death. Aspirin and nonselective NSAIDs each inhibit the generation of prostaglandins by inhibiting the two cyclooxygenase (COX) enzymes that initiate prostaglandin synthesis, COX-1 and COX-2. NSAIDs that are selective for COX-2 also inhibit the generation of prostaglandins. COX-1 is constitutively expressed in the colon, but COX-2 is inducible and markedly up-regulated in many colon cancers. Interventional trials have shown a decreased risk of the development of colon adenomas in high-risk subjects.


Copyright © 2007 Massachusetts Medical Society.
who were given aspirin or COX-2 selective NSAID antagonists, and observational trials have associated a decreased risk of colon cancer with aspirin use.\textsuperscript{1,6}

In the normal colon, 15-prostaglandin dehydrogenase (15-PGDH), a prostaglandin-degrading enzyme, acts to down-regulate the prostaglandin pathway. A potent suppressor of the growth of human colon tumor-cell lines in immunodeficient mice, 15-PGDH also inhibits the development of murine intestinal neoplasias.\textsuperscript{7-9} These findings and the fact that the expression of 15-PGDH is abolished in human colonic neoplasms highlight the oncogenic potential of the prostaglandin pathway.

In this issue of the Journal, Chan et al. report findings that strongly support the primary inhibition of COX-2–mediated synthesis of prostaglandins in the prevention of colon cancer by aspirin.\textsuperscript{10} They demonstrate that aspirin specifically prevents the subgroup of colon cancers in which COX-2 is most highly induced. Hence, colon cancers, or their adenomatous precursors, that express a high level of COX-2 are presumably “addicted” to COX-2 enzymatic activity and are, therefore, particularly susceptible to COX-2 inhibitors. This work raises a number of questions. How is the cancer-promoting activity of COX-2 mediated? Is COX-2 the only or the best target in the pathway to colon cancer? Can we identify the people who are most likely to benefit from COX-2 inhibition? Are some established colon cancers responsive to inhibitors of this pathway?

Both COX-1 and COX-2 enzymatically convert arachidonic acid, a 20-carbon fatty acid, to the prostaglandin precursor PGH\textsubscript{2}, from which different prostaglandin synthases generate individual members of series-2 prostanooids. These prostanooids include prostaglandins PGE\textsubscript{2}, PGD\textsubscript{2}, PGI\textsubscript{2} (prostacyclin), PGF\textsubscript{2\alpha}, and the eicosanoid TXA\textsubscript{2} (thromboxane) (Fig. 1).\textsuperscript{9}

COX inhibitors thus affect a wide swath of biologic pathways, but despite the complex biology, current thinking ascribes the cancer-preventive activity of aspirin and NSAIDs principally to their ability to block the generation of PGE\textsubscript{2}, the most abundant colonic prostaglandin, by COX-2. This idea is supported not only by the preventive activity of COX-2–selective NSAIDs against colon tumors but also by the resistance to intestinal tumorigenesis of the COX-2 knockout mouse and of knockout mice lacking receptors for PGE\textsubscript{2}, as well as by the augmented development of intestinal tumors in mice given PGE\textsubscript{2} orally.\textsuperscript{9,11,12}

In selected PGE\textsubscript{2}-responsive colon cancer–cell lines, the addition of PGE\textsubscript{2} increases the growth, migration, and invasiveness of the cells and also increases their resistance to apoptosis and secretion of pro-angiogenic molecules.\textsuperscript{9} These biologic activities are mediated by the up-regulation by PGE\textsubscript{2} of the expression of a number of genes, including genes encoding pro-angiogenic molecules.\textsuperscript{9} PGE\textsubscript{2} also increases expression of the anti-apoptotic Bcl-2 oncogene and the proliferation-promoting cyclin D1 oncogene.\textsuperscript{9} The consistent resistance of established human colon cancers and most colon cancer–cell lines to COX inhibitors clearly suggests that neoplastic progression is coupled to genes that enable escape from COX-2 dependence.

PGE\textsubscript{2} signaling is initiated by binding to EP1, EP2, EP3, or EP4 G protein–coupled prostaglandin receptors.\textsuperscript{9} EP2 has a role in promoting colonic neoplasia: stimulation of EP2 by PGE\textsubscript{2} activates the kinase activity of the phosphatidylinositol-3-kinase (PI3K) oncogene and the transcription-factor activity of β-catenin. Stimulation of EP4 by PGE\textsubscript{2} triggers phosphorylation of the epidermal growth factor receptor. Downstream targets that these pathways activate are shown in Figure 1. Intriguingly, many of the pathways and genes activated by PGE\textsubscript{2} also participate in feedback loops that elevate COX-2 levels. Such feedback loops may amplify the activity of the COX-2 pathway and may magnify the potency of COX-2 inhibitors.

The RAS, PI3K, and β-catenin–signaling pathways that are activated by COX-2 (Fig. 1) are also activated by mutations of those genes in colon cancer. Perhaps such mutations, in addition to contributing to tumor progression, free advanced colon neoplasms from dependence on COX-2 and thereby contribute to the resistance of established cancers to COX-2 inhibitors. The findings of Chan et al. raise the question whether mutations in the downstream Ras, PI3K, and APC or β-catenin genes occur preferentially in colon cancers that arise among aspirin users or that express low levels of COX-2. Equally, these findings suggest that such mutations may not be present in some high COX-2–expressing cancers that occur in non-users of aspirin. A related question is whether colon cancers with the highest COX-2 levels retain responsiveness to aspirin or to COX-2 in-
Hibitors added to palliative or adjuvant chemotherapies.

Chan et al. found that only one third of colon cancers that are positive for COX-2 are prevented by regular aspirin use (relative risk, 0.64); moreover, chronic use of aspirin or COX-2 inhibitors carries attendant toxic effects. For these reasons, we need to ask whether there are alternative strategies for targeting the COX pathway that have better efficacy or lower rates of adverse effects.
Shades of Dry — Curing Urinary Stress Incontinence
Kris Strohbehn, M.D.

Urinary incontinence is a common condition affecting 20 to 40% of older women.¹ The two most common types of urinary incontinence are stress incontinence, the involuntary loss of urine resulting from increased abdominal pressures (such as with a cough or Valsalva’s maneuver), and urge incontinence, the involuntary loss of urine after an unwanted contraction of the detrusor muscle.² Some patients have mixed incontinence with both types of conditions.

The frequency of both types of urinary incontinence increases with age, with peaks in preva-
ence around menopause and after the age of 65 years.\textsuperscript{3,4} Obesity and multiparity are recognized risk factors for stress incontinence. Even so, many nulliparous teenage women, especially athletes, have episodes of stress incontinence at times of increased physical activity.\textsuperscript{5} The degree to which incontinence bothers a woman may vary, depending on her activity level. Many women adjust their activity to reduce the frequency of leakage, which may adversely affect other health and quality-of-life measures.

The age-adjusted rate of inpatient surgical procedures for urinary incontinence in women in the United States increased from 0.32 per 1000 women in 1979 to 0.60 per 1000 women in 1997.\textsuperscript{6} Much of the rise can be attributed to an increase by a factor of 3 in the number of procedures among women 50 years of age or older (from 0.5 to 1.5 per 1000 women). Given the aging of the U.S. population, the total number of procedures for treatment of urinary incontinence is expected to continue to rise.\textsuperscript{6,7} Yet there are few data from randomized trials to inform surgical decision making.

In this issue of the\textit{Journal}, Albo et al.\textsuperscript{8} report the results of a multicenter, randomized, controlled trial sponsored by the National Institutes of Health (NIH), comparing two surgical treatments for stress incontinence: the pubovaginal sling procedure and the Burch procedure. (Descriptions and illustrations are included in the article.) Both procedures have been considered highly effective, with reported cure rates of 80 to 90%, although in observational studies the Burch procedure has had lower reported cure rates if there is evidence of a very weak urethral sphincter muscle. In the study by Albo et al., at 24 months, the cure rates were significantly higher for the pubovaginal sling than for the Burch procedure for both overall incontinence (47% vs. 38%) and stress incontinence (66% vs. 49%).\textsuperscript{8} Cure rates declined over time in both groups, emphasizing the need for a longer-term study of outcomes. Despite the higher cure rates, the sling group had more complications, including urinary retention, voiding dysfunction, urgency symptoms, urinary tract infections, and reoperation for problems with voiding.

The cure rates in both groups were lower than those commonly reported, an observation that is probably explained by the use of strict criteria to define cure and the variable criteria for cure used in previous studies. Stress incontinence can be considered to be a subjective symptom (described by the patient), an objective finding on examination (the patient coughs with a full bladder and urine is observed coming from the urethra), or a condition that is determined by subjective and objective findings, including urodynamic evaluation.\textsuperscript{2,\textsuperscript{9}} Accordingly, outcome tools can be subjective (including bladder diaries and validated surveys to determine quality-of-life scores, severity of incontinence, and treatment satisfaction) or objective (weighing a patient’s incontinence pads, cough stress testing to observe incontinence, and other complex urodynamic tests). At present, there is no single outcome that adequately measures success after treatment of urinary incontinence. There are as many “shades of dry” as there are shades of gray.

A commentary published in 2002 underscored the dependence of surgical success rates on the outcome measure used for stress incontinence, noting cure rates of 70 to 83% if patients’ satisfaction was the only outcome measure; however, cure rates were reportedly as low as 6 to 9% according to very strict NIH criteria (no symptoms of stress incontinence, negative objective testing, and no new problems due to the intervention).\textsuperscript{9} Cure rates in the study by Albo et al. likewise varied, depending on whether subjective, objective, or combined outcome tools were reported.

In comparing outcomes, then, is a patient’s report of satisfaction more important, or are results of objective testing (e.g., leakage on a cough stress test) more important? A recent survey\textsuperscript{11} found a poor correlation between attitudes of clinicians and patients toward outcomes. Contrary to patients’ perception of “bothersomeness,” a majority of clinicians thought that small or infrequent episodes of leakage were acceptable after treatment.

Patients’ perception of bother may also be influenced by new-onset postoperative voiding problems or urgency. Such problems are well described after procedures for stress incontinence, possibly owing to increases in urethral resistance. If urethral resistance exceeds the strength of a woman’s voiding detrusor muscle, urinary retention and voiding dysfunction may result; subsequent detrusor-muscle overactivity may lead to urgency, frequency, nocturia, or urge incontinence. At present, there is no testing that accu-
rately identifies patients in whom these problems will develop after surgery for stress incontinence, although data from the study by Albo et al. indicate that such complications are significantly more likely with the pubovaginal sling than with the Burch procedure. If postoperative urinary retention requiring intermittent catheterization develops in a woman who had mildly bothersome stress incontinence, she is likely to be dissatisfied even if her stress incontinence is reduced. However, a patient who had severely bothersome stress incontinence and a similar postoperative complication might be highly satisfied if her incontinence has been eliminated. The primary determinants of perceived “success” in these cases are the difference between preoperative and postoperative “bothersomeness” of symptoms and whether the surgery met the patients’ expectations.

Although Albo et al. provide important information for patients and clinicians in deciding between the Burch procedure and the pubovaginal sling, new techniques are rapidly expanding the available options. A new generation of mesh synthetic slings has been introduced in the past decade, with cited advantages of lower rates of urinary retention, smaller incisions, less pain, quicker recovery, and lower cost and complications. Such slings are placed blindly with needles through the retropubic or transobturator space, guiding the sling to the midurethral area through a small vaginal incision. However, risks include possible vaginal, urethral, or bladder erosion of the synthetic materials. Injuries to adjacent structures (bowel, bladder, urethra, and large blood vessels) have also been reported owing to the blind nature of needle passage. Data from a recent randomized trial suggested that midurethral mesh slings have success rates similar to those with the Burch procedure,12 but these slings have not been directly compared with pubovaginal slings.

Randomized trials such as that of Albo et al. greatly advance our ability to counsel patients and effectively compare surgical options for treatment of stress incontinence. Objective and subjective outcome data are often discordant, and there is no consensus on which “shade” of dry is most important. Future surgical trials should continue to use multiple outcome measures, including ones directed by patients, to identify whether surgical procedures have met patients’ expectations and goals.13,14

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

From the Department of Obstetrics and Gynecology, Dartmouth Medical School, and the Division of Urogynecology and Reconstructive Pelvic Surgery, Dartmouth–Hitchcock Medical Center — both in Lebanon, NH.


Copyright © 2007 Massachusetts Medical Society.
The Supreme Court and Abortion Rights
George J. Annas, J.D., M.P.H.

Since the Supreme Court’s landmark 1973 abortion-rights decision in *Roe v. Wade*, the law has taken the lead in defining the contours of the continuing public debate over reproductive liberty. Ever since then, abortion opponents have tried to make abortion more burdensome by limiting *Roe*, and these continuing challenges are the reason there have been so many Supreme Court decisions about abortion, including the Court’s 1992 decision in *Planned Parenthood of Southeastern Pennsylvania v. Casey*, which unexpectedly reaffirmed the core of *Roe*.

In the wake of *Casey*, political efforts to restrict abortion have switched to outlawing one specific medical procedure, which its opponents label “partial-birth abortion,” and more than 30 states and the federal government have made it a crime to perform this procedure. In 2000, in *Stenberg v. Carhart*, the Court ruled 5 to 4 that these laws are unconstitutional. In April 2007, also by a 5 to 4 vote, the Court reached the opposite conclusion in *Gonzales v. Carhart*. This is the first time the Court has ever held that physicians can be prohibited from using a medical procedure deemed necessary by the physician to benefit the patient’s health. The importance of the decision to physicians and their patients cannot be appreciated without an understanding of the constitutional law of reproductive liberty as it has developed during the past 40 years.

**The Right to Privacy**

The first case to embrace the concept of reproductive liberty was *Griswold v. Connecticut*, in which the Court ruled in 1965 that a Connecticut statute criminalizing the use of contraceptives violated the constitutional right to privacy that married couples had in sexual relations. Later, in 1972, the Court found that even outside marriage, a person had a “right to privacy . . . to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision to bear or beget a child.”

The following year, in *Roe*, the Court struck down a Texas law that made it a crime for physicians to perform an abortion unless it was necessary to save the life of the patient; there were no exceptions for the woman’s health. The Court held that women have a constitutional right of privacy that is fundamental and “broad enough to encompass a woman’s decision . . . to terminate her pregnancy.” Because the right is fundamental, states that wish to restrict abortion rights were required to demonstrate a compelling interest to restrict the exercise of this right. The Court ruled that the state’s interest in the life of the fetus became compelling only at the point of viability, when the fetus can survive independently of its mother. Even after the point of viability, the state cannot favor the life of the fetus over the life or health of the pregnant woman. Under the right of privacy, physicians must be free to use their “medical judgment for the preservation of the life or health of the mother.” On the same day that the Court decided *Roe*, it also decided *Doe v. Bolton*, in which the Court defined health very broadly:

The medical judgment may be exercised in the light of all factors — physical, emotional, psychological, familial, and the woman’s age — relevant to the well-being of the patient. All these factors may relate to health. This allows the attending physician the room he needs to make his best medical judgment.

*Roe* and *Doe* together established that both physician and patient were protected by the constitutional right of privacy. In later cases, the Court continued to defer to the medical judgment of the attending physician. For example, in 1976 in *Planned Parenthood of Central Missouri v. Danforth*, the Court concluded that state legislatures could not determine when viability occurred; rather this “essentially medical concept . . . is, and must be, a matter for the judgment of the responsible physician.”

---

**THE RIGHT TO PRIVACY**

The first case to embrace the concept of reproductive liberty was *Griswold v. Connecticut*, in which the Court ruled in 1965 that a Connecticut statute criminalizing the use of contraceptives violated the constitutional right to privacy that married couples had in sexual relations. Later, in 1972, the Court found that even outside marriage, a person had a “right to privacy . . . to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision to bear or beget a child.”

The following year, in *Roe*, the Court struck down a Texas law that made it a crime for physicians to perform an abortion unless it was necessary to save the life of the patient; there were no exceptions for the woman’s health. The Court held that women have a constitutional right of privacy that is fundamental and “broad enough to encompass a woman’s decision . . . to terminate her pregnancy.” Because the right is fundamental, states that wish to restrict abortion rights were required to demonstrate a compelling interest to restrict the exercise of this right. The Court ruled that the state’s interest in the life of the fetus became compelling only at the point of viability, when the fetus can survive independently of its mother. Even after the point of viability, the state cannot favor the life of the fetus over the life or health of the pregnant woman. Under the right of privacy, physicians must be free to use their “medical judgment for the preservation of the life or health of the mother.” On the same day that the Court decided *Roe*, it also decided *Doe v. Bolton*, in which the Court defined health very broadly:

The medical judgment may be exercised in the light of all factors — physical, emotional, psychological, familial, and the woman’s age — relevant to the well-being of the patient. All these factors may relate to health. This allows the attending physician the room he needs to make his best medical judgment.

*Roe* and *Doe* together established that both physician and patient were protected by the constitutional right of privacy. In later cases, the Court continued to defer to the medical judgment of the attending physician. For example, in 1976 in *Planned Parenthood of Central Missouri v. Danforth*, the Court concluded that state legislatures could not determine when viability occurred; rather this “essentially medical concept . . . is, and must be, a matter for the judgment of the responsible physician.”
attending physician.” By the end of the 1980s, a pattern in Court decisions could be discerned in which abortion regulations that significantly burdened a woman’s decision, treated abortion differently from other similar medical or surgical procedures, interfered with the exercise of professional judgment by the attending physician, or were stricter than accepted medical standards were struck down by the Court.3

Privacy as a constitutional right became a one-word description of liberty to make decisions regarding marriage, procreation, contraception, sterilization, abortion, family relationships, child rearing, and sexual relationships free of governmental interference.2,10

T H E R I G H T T O L I B E R T Y

One strategy to change Roe was to change the composition of the Supreme Court by appointing anti–Roe justices. Because of new justices on the Court in 1992, in Casey, the Court had its first real opportunity to overturn Roe v. Wade. Many Court observers thought it would. Instead, in an unusual procedure for the Court, three potentially anti–Roe justices, Justices Sandra Day O’Connor, David Souter, and Anthony Kennedy, joined together to write a joint opinion confirming the “core holding” of Roe. (They were joined in most of their opinion by two justices who would have simply upheld Roe, making this a 5-to-4 decision.) Most centrally, the authors of the joint opinion believed that although the pressure to overrule Roe has grown “more intense,” doing so would severely and unnecessarily damage the Court’s legitimacy by undermining “the Nation’s commitment to the rule of law.”2

Specifically, the three justices wrote that they were reaffirming “Roe’s essential holding” that before the point of viability a woman has a right to choose abortion without undue state interference, that after the point of viability the state can restrict abortion “if the law contains exceptions for pregnancies which endanger the woman’s life or health,” and that “the state has legitimate interests from the outset of the pregnancy in protecting the health of the woman and the life of the fetus that may become a child.” The Court applied these principles to uphold laws mandating much more detailed requirements for abortion, as well as a mandatory 24-hour waiting period, but struck down a spousal-notification requirement as an “undue burden.” Thus, after Casey, Roe stood for the proposition that pregnant women have a “personal liberty” right (“privacy” went unmentioned) to choose to terminate their pregnancies before the point of viability and that the state cannot “unduly burden” such a right by erecting barriers that effectively prevent the exercise of that choice.2,11 Of course, a major problem was definitional: burdensome regulations were acceptable, “unduly burdensome” ones were not — but it was not clear what qualified as which. Put another way, the state could demonstrate its concern for life by requiring that physicians make women seeking abortions jump through new and burdensome hoops (including offers of detailed and accurate information on abortion, the status of the fetus, adoption, sources of help for childbirth, and a 24-hour waiting period), as long as doing so did not “unduly burden” women by actually preventing them from being able to make a decision to have an abortion.

With the loss of all hope that the Court would overrule Roe wholesale, anti–Roe advocates switched strategies dramatically, focusing on criminalizing a specific procedure that they believed would horrify most Americans and that they labeled “partial-birth abortion.” The first such bill passed Congress in 1996 and was vetoed by President Bill Clinton because the prohibition did not contain an exception for the health of the woman, as required by Roe and Casey. In 1997, this time with the support of the American Medical Association, the bill passed Congress again. President Clinton vetoed it, again for failure to contain a health exception.12

“P A R T I A L - B I R T H A B O R T I O N ” A N D T H E S T A T E S

Proponents of the ban took their cause to the individual states, a majority of which enacted substantially identical laws. In 2000, Nebraska’s partial-birth abortion law reached the Supreme Court. The Nebraska law carried a penalty of up to 20 years in prison for physicians who performed the procedure. The law reads in relevant part:

No partial-birth abortion shall be performed in this state, unless such a procedure is necessary to save the life of the mother whose life is endangered by a physical disorder, physical illness, or phys-
ical injury, including a life-endangering physical condition caused by or arising from the pregnancy itself.

[A “partial-birth abortion” is] an abortion procedure in which the person performing the abortion partially delivers vaginally a living unborn child before killing the unborn child and completing the delivery. . . . [The statute further defines the phrase “partially delivers vaginally a living unborn child before killing the unborn child” as] deliberately and intentionally delivering into the vagina a living unborn child, or a substantial portion thereof, for the purpose of performing a procedure that the person performing such procedure knows will kill the unborn child and does kill the unborn child [emphasis added].

This ban applies throughout pregnancy and has no exception to protect the woman’s health, only to save her life. In a 5-to-4 opinion in Stenberg v. Carhart, the Court found this law unconstitutional for two reasons. First, the description of the banned procedure was too close to dilation and evacuation (D&E), another procedure that was permitted and widely used for second-trimester abortions. Therefore, this law would discourage physicians from using the lawful procedure, which would place an undue burden on their patients. Second, the law failed to provide an exception for instances in which the procedure was deemed necessary by the physician to protect the woman’s health, as required by Roe and Casey. Justice John Paul Stevens, in his concurring opinion, noted that the extreme anti-Roe rhetoric as exemplified in the partial-birth abortion debate obscured the fact that during the 27-year period since Roe was decided, the core holding of Roe “has been endorsed by all but 4 of the 17 Justices who have addressed the issue.”

A notable dissenting opinion was written by Justice Kennedy, who had specifically endorsed the core of Roe in Casey. Kennedy argued that the outlawing of “partial-birth abortion” was consistent with Casey because of the interest the state has throughout pregnancy in protecting the life of the fetus that may become a child. In his view, the banned procedure conflates abortion and childbirth in a way that “might cause the medical profession or society as a whole to become insensitive, even disdainful, to life, including life of the human fetus.” He also argued that such a ban was not unduly burdensome to women because state legislatures can determine that specific medical procedures, like this one, are not medically necessary.

Justice Stephen Breyer, the author of the Stenberg majority opinion, stated that a more precise law, with a health exception, could be constitutional. In 2003, Congress passed a slightly revised law. It did not contain a health exception, but its preface did contain a declaration that the outlawed procedure was never medically necessary for the health of the woman. President Bush signed it into law on November 5, 2003. By the time the Court ruled on the constitutionality of this law in April 2007, in Gonzales v. Carhart, there were two important changes in the composition of the Court: a new chief justice, John Roberts, who replaced the consistently anti-Roe Chief William Rehnquist, and Justice Samuel Alito, who replaced Justice Sandra Day O’Connor, who was consistently pro-Roe (as interpreted in Casey). The federal law provides that

(a) Any physician who, in or affecting interstate or foreign commerce, knowingly performs a partial birth abortion and thereby kills a human fetus shall be fined under this title or imprisoned not more than 2 years, or both. This subsection does not apply to a partial birth abortion that is necessary to save the life of a mother whose life is endangered by a physical disorder, physical illness, or physical injury, including a life-endangering physical condition caused by or arising from the pregnancy itself.

(b) (1) The term “partial birth” abortion means an abortion in which the person performing the abortion

(A) Deliberately and intentionally vaginally delivers a living fetus until, in the case of a head-first presentation, the entire fetal head
is outside the body of the mother, or, in the case of breech presentation, any part of the fetal trunk past the navel is outside the body of the mother, for the purpose of performing an overt act that the person knows will kill the partially delivered living fetus; and

(B) Performs the overt act, other than completion of delivery, that kills the partially delivered living fetus [emphasis added].

The Court decided, 5 to 4, that this law was constitutional. Justice Kennedy wrote the majority opinion for himself, Justices Antonin Scalia and Clarence Thomas, and the two new justices. In it he substantially adopts his dissenting opinion in Stenberg as the Court’s new majority opinion. Although he concludes that his decision is consistent with Stenberg, all three U.S. District courts and all three Courts of Appeal that had examined this federal law found it unconstitutional under the principles in Casey and Stenberg, primarily because of the vagueness of the definition and the lack of a health exception.4

As to the vagueness argument, Kennedy writes that the new law is no longer vague because it clarifies the distinction between the prohibited procedure (which he calls “intact D&E”) and standard D&E abortions because the former requires the delivery of an intact fetus, whereas the latter requires “the removal of fetal parts that are ripped from the fetus as they are pulled through the cervix.” In addition, the new federal law specifies fetal landmarks (e.g., the “navel”) instead of the vague description of a “substantial portion” of the “unborn child.”4

Since the law applies to fetuses both before and after the point of viability, Kennedy concedes that under Casey the law would be unconstitutional “if its purpose or effect is to place a substantial obstacle in the path of a woman seeking an abortion before the fetus attains viability.”4 Kennedy finds Congress’s purpose is twofold: first, lawmakers wanted to “express respect for the dignity of human life” by outlawing “a method of abortion in which a fetus is killed just inches before completion of the birth process,” because use of this procedure “will further coarsen society to the humanity of not only newborns, but of all vulnerable and innocent human life. . . .” Second, Congress wanted to protect medical ethics, finding that this procedure “confuses the medical, legal and ethical duties of physicians to preserve and promote life. . . .”4

The key to Kennedy’s legal analysis is his conclusion that these reasons are constitutionally sufficient to justify the ban because under Casey “the State, from the inception of pregnancy, maintains its own regulatory interest in protecting the life of the fetus that may become a child [and this interest] cannot be set at naught by interpreting Casey’s requirement of a health exception so it becomes tantamount to allowing the doctor to choose the abortion method he or she might prefer.”4

Kennedy then goes on to write that “respect for human life finds an ultimate expression in the bond of love the mother has for her child,” and that “while no reliable data” exist on the subject, “it seems unexceptional to conclude some women come to regret their choice to abort the infant life they once created and sustained. . . . Severe depression and loss of esteem can follow.” Such regret, Justice Kennedy believes, can be caused or exacerbated if women later learn what the procedure entails, suggesting that physicians fail to describe it to patients because they “may prefer not to disclose precise details of the means [of abortion] that will be used. . . .”4

The final, important issue is whether the prohibition would “ever impose significant health risks on women” and whether physicians or Congress should make this determination. Kennedy picks Congress: “The law need not give abortion doctors unfettered choice in the course of their medical practice, nor should it elevate their status above other physicians in the medical community. . . . Medical uncertainty does not foreclose the exercise of legislative power in the abortion context any more than it does in other contexts.”4 Furthermore, Kennedy argues, the law does not impose an “undue burden” on women for another reason: alternative ways of killing a fetus have not been prohibited. In his words, “If the intact D&E procedure is truly necessary in some circumstances, it appears likely an injection that kills the fetus is an alternative under the Act that allows the doctor to perform the procedure.”4

Justice Ginsburg’s Dissent

Writing for the four justices in the minority, Justice Ruth Bader Ginsburg observes, “Today’s de-
cision is alarming. It refuses to take Casey and Stenberg seriously. It tolerates, indeed applauds, federal intervention to ban nationwide a procedure found necessary and proper in certain cases by the American College of Obstetricians (ACOG). It blurs the line, firmly drawn in Casey, between previability and postviability abortions. And, for the first time since Roe, the Court blesses a prohibition with no exception safeguarding a woman’s health.”

Ginsburg argues that the majority of the Court has overruled the conclusion in Stenberg that a health exception is required when “substantial medical authority supports the proposition that banning a particular abortion procedure could endanger women’s health.” This conclusion, bolstered by evidence presented by nine professional organizations, including the ACOG, and conclusions by all three U.S. District Courts that heard evidence concerning the Act and its effects, directly contradicted the congressional declaration that “there is no credible medical evidence that partial-birth abortions are safe or are safer than other abortion procedures.” Even Justice Kennedy agreed that Congress’s finding was untenable.

Justice Ginsburg concludes that this leaves only “flimsy and transparent justifications” for upholding the ban. She rejects those justifications, arguing that the state’s interest in “preserving and promoting fetal life” cannot be furthered by a ban that targets only a method of abortion and that cannot save “a single fetus from destruction” by its own terms but may put women’s health at risk. Ultimately, she believes that the decision rests entirely on the proposition, never before enshrined in a majority opinion and explicitly repudiated in Casey, that “ethical and moral concerns” unrelated to the government’s interest in “preserving life” can overcome what had been considered fundamental rights of citizens.

The majority seeks to bolster its conclusion by describing pregnant women as in a fragile emotional state that physicians may take advantage of by withholding information about abortion procedures. Justice Ginsburg concludes that the majority’s solution to this hypothetical problem is to “deprive women of the right to make an autonomous choice, even at the expense of their safety.” She continues, “This way of thinking [that men must protect women by restricting their choices] reflects ancient notions about women’s place in the family and under the Constitution — ideas that have long since been discredited.”

Ginsburg further notes that the majority simply cannot contain its hostility to reproductive rights as articulated in Roe and Casey, calling physicians “abortion doctors,” describing the fetus as an “unborn child” and as a “baby,” labeling second-trimester abortions as “late term,” and dismissing “the reasoned medical judgments of highly trained doctors . . . as ‘preferences’ motivated by ‘mere convenience.’”

Ginsburg makes two final points. First, although the Court invites a lawsuit to challenge the Act “as applied,” it gives “no clue” as to how such a lawsuit should be brought. Surely, she asks, “the Court cannot mean that no suit to challenge the ban [based on how it affects an actual woman or her physician] may be brought until a woman’s health is immediately jeopardized.” Second, she argues that the opinion threatens to undercut the “rule of law” and the “principle of stare decisis,” both of which the Court affirmed in Casey, concluding that, “A decision so at odds with our jurisprudence should not have staying power.” As described in Casey, stare decisis is a doctrine that the principals use to decide whether previous cases, called precedents, should be followed by courts and not abandoned under “political pressure” or as an “unprincipled emotional reaction.”

**DISCUSSION**

The major change in the law this opinion brings with it is the new willingness of Congress and the Court to disregard the health of pregnant women and the medical judgment of their physicians. This departure from precedent was made possible by categorizing physicians as unprincipled “abortion doctors” and infantilizing pregnant women as incapable of making serious decisions about their lives and health. The majority opinion ignores or marginalizes long-standing principles of constitutional law, substituting the personal morality of Justice Kennedy and four of his colleagues.

The majority asserts that giving Congress constitutional authority to regulate medical practice is not new but identifies no case in which Congress had ever outlawed a medical procedure. Its reliance on the more than 100-year-old case of Jacobson v. Massachusetts is especially inapt.
son was about mandatory smallpox vaccination during an epidemic. The statute had an exception for “children who present a certificate, signed by a registered physician, that they are unfit subjects for vaccination,” and the Court implied that a similar medical exception would be constitutionally required for adults. It is not just abortion regulations that have had a health exception for physicians and their patients — all health regulations have.16–18

On the other hand, those who expect Roe to be overturned by this Court may be disappointed. Although Justice Alito has replaced Justice O’Connor and is likely to vote in the opposite direction on Roe-related issues, Justice Kennedy is the new swing vote on the Court, and he insists that he is upholding the principles of Roe v. Wade as reaffirmed in Casey.3 Just as the question of whether a specific abortion regulation was an “undue burden” was once a determination Justice O’Connor could effectively make for the Court, the meaning of Roe v. Wade is, at least for now, up to Justice Kennedy.

CONCLUSIONS

Some physicians will surely be tempted to view the decision as a narrow victory for antiabortion forces that is unlikely to have more than a marginal effect on medical practice. This view is understandable but misses the potential broader impact of the opinion on the regulation of medical practice and the doctor–patient relationship generally. Until this opinion, the Court recognized the importance of not interfering with medical judgments made by physicians to protect a patient’s interest.15 For the first time, the Court permits congressional judgment to replace medical judgment.

For physicians who are disturbed or dismayed by this opinion — for example, the ACOG has termed it “shameful and incomprehensible”10 — there are concrete actions to consider. One is to seek an amendment of the act in Congress to protect women’s health — for instance, by adding a specific exemption for cases in which “in the reasonable medical judgment of the attending physician, an alternative procedure poses a significant risk to the health of the pregnant woman.” Although it would be better simply to repeal the law, this amendment could actually pass because it permits legislators to be against using the despised procedure but at the same time protecting the health of women.

A second, admittedly much more difficult, response is for physicians to become conscientious objectors in particular circumstances. This means doctors will do what is medically necessary to preserve and protect the lives and health of their patients as required by medical ethics, regardless of what politicians attempt to dictate. Unlike antiabortion conscientious objection, this kind does not come with legal immunity. There is danger of prosecution, and this approach will be a viable option only if physicians are assured of the financial and moral support of the medical profession (especially of the ACOG) and, I think, of the legal profession as well. I believe the American Bar Association should agree as an organization to actively support any physician who is prosecuted under this law for doing what he or she believed at the time was in the patient’s best medical interests. This strategy means that a physician who is accused of violating this law would challenge its constitutionality as part of his or her defense in the criminal action, what the Court seemed to mean by an “as applied” case.

Many state legislatures will now enact new laws restricting abortion access to see how far they can go, just as happened after Roe. Other states, especially those like New York that had made abortion legal before Roe, may codify the basic protections of Roe into state law.20 In anti-Roe states, there are likely to be increased requirements for physicians to present their patients with more and more information designed to discourage pregnant women from having abortions, such as viewing ultrasonographic images of their fetuses. Some states will also attempt to outlaw other abortion procedures that the members of their legislatures find personally or religiously objectionable, including standard D&E. In the past, members of state legislatures could vote for all sorts of restrictions and bans, knowing that the courts would almost certainly find them unconstitutional. Thus, they could be publicly in favor of abortion restrictions and at the same time privately assure their pro-choice constituents that such restrictions would have no effect on women. Now that states (and Congress) have been given the green light to regulate medicine on the basis of their own views of morals and ethics, detached from medicine and science, these legislators may have to make real decisions.
For the sake of their patients and the profession of medicine, physicians will have to pay more attention to politics.

George Annas is co-chair of the Health Rights and Bioethics Committee of the American Bar Association’s Individual Rights and Responsibilities Section.

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

This article (10.1056/NEJMrle072595) was published at www.nejm.org on May 2, 2007.

From the Department of Health Law, Bioethics, and Human Rights, Boston University School of Public Health, Boston.


Copyright © 2007 Massachusetts Medical Society.
A Healthy Tan?
Gregory Barsh, M.D., Ph.D., and Laura D. Attardi, Ph.D.

For the past 500 million years or so, the tumor suppressor and transcription factor known as p53 has been the guardian of the genome in multicellular organisms. It mediates the response to DNA damage by inducing cell-cycle arrest and facilitating DNA repair or, if the damage is particularly severe, by triggering apoptosis. Both responses are critical to cancer prevention. More recently — about 1 million years ago — a tanning response evolved in our hominid ancestors, in which the accumulation of melanin granules in keratinocytes provides physical protection against the DNA-damaging effects of sunlight. The tanning response to “sunseeking” behavior is familiar not only to millions of beachgoers today but also to the dermatologists and oncologists who care for patients with melanoma and squamous-cell carcinoma. Although tanning has long been correlated with the cellular response to DNA damage, a new study by Cui et al. provides a direct link between tanning and p53, as well as some clues to the underlying molecular mechanisms.

An apt motto for the skin — “to protect and serve” — captures its roles in barrier, endocrine, and homeostatic functions but only hints at the underlying complexity. In response to ultraviolet (UV) irradiation, epidermal keratinocytes release a host of cytokines and paracrine signaling molecules that target lymphocytes, macrophages, and melanocytes. This network of cellular interactions in the epidermis brings about a pleiotropic response to UV irradiation, including erythema, immunosuppression, and of course, tanning.

The study by Cui et al. rests upon two planks of previous work. First, it was established that UV irradiation activates the p53 protein to stimulate transcription of its target genes in keratinocytes, as in other types of cells exposed to DNA-damaging agents. Second, studies of human pigmented variation showed that mutations of the gene encoding the melanocortin 1 receptor (MC1R), expressed mainly on the cell surface of melanocytes, cause a characteristic phenotype of red hair, fair skin, and inability to tan. This receptor was discovered by virtue of its ability to elevate cyclic AMP levels in response to melanocortins, a family of small peptides — including α-melanocyte-stimulating hormone and adrenocorticotrophic hormone — that are derived from a single precursor protein, pro-opiomelanocortin (POMC).

Cui et al. observed that mice genetically deficient in p53 failed to tan after UV irradiation. (Tanning in mice is measured in their ears, where epidermal melanocytes — which respond in much the same way as do human melanocytes — are abundant.) Cui et al. went on to show that the likely mechanism behind the failure to tan involves the ability of p53 to stimulate transcription of the POMC gene in keratinocytes. They posit that the transcription of POMC in sun-exposed keratinocytes leads to increased release of α-melanocyte-stimulating hormone, activation of the MC1R on melanocytes, and consequently, increased melanogenesis, melanocytic differentiation, and transfer of pigment-containing organelles (melanosomes) to keratinocytes — which together are responsible for the tanning response (Fig. 1).

A direct connection between p53, melanocortin signaling, and the tanning response makes sense from both an evolutionary and a teleologic perspective. At the same time, important questions remain to be answered. The ear of an adult mouse is a reasonable proxy for human skin, but hypopigmentation or inability to tan is not a general feature of the Li–Fraumeni syndrome, in which normal p53 function is often substantially compromised. In addition, melanocortin signaling is likely to be one of several downstream tanning effectors. People of northern European ancestry with MC1R mutations are indeed unable to tan, but their fair skin is the product of evolutionary selection, and it is likely that polymorphisms in genes other than MC1R impair tanning. Finally, the role of keratinocyte-derived POMC as a prerequisite for MC1R activation should
be interpreted cautiously, since the MC1R has high levels of basal activity and the deletion of Pomc in inbred mice has no effect on skin or hair pigmentation.\(^5\)

These considerations notwithstanding, the work by Cui et al. presents some intriguing opportunities. Given the genoprotective effects of tanning, it is tempting to envision pharmacologic stimulation of the tanning response, either by activating p53 directly or by intervening down-
stream of p53 to trigger the melanocortin system. There have been intensive efforts directed toward activating p53 in tumors, and these approaches could be harnessed to induce tanning. Topical application of small molecules designed to activate p53 — such as molecules that repress its negative regulator, MDM2 — could be used as a surrogate for sunlight-induced p53 activation. This strategy would require careful management, however, since robust p53 induction would probably induce keratinocyte senescence or apoptosis, the latter mimicking severe sunburn. Alternatively, topically applied drugs that activate the MC1R, for example, would provide another means to the same end.

Cui et al. also speculate that sunseeking behavior could be motivated by the release of keratinocyte-derived β-endorphin (also derived from POMC) into the circulation. The provocative idea of a “tanner’s high” will likely add fuel to the fire of whether potential benefits of artificial tanning outweigh the risks. Although some beneficial effects of cutaneous exposure to UV light, principally increased absorption of vitamin D, are championed by advocates of the oxymoronic “smart tanning” industry, most dermatologists and epidemiologists counsel that it is smarter to get one’s dose of vitamin D orally. Further studies of cutaneous and pigmentedary processes will be necessary to determine whether, and how, one might achieve a truly healthy tan.

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

From the Stanford University School of Medicine, Stanford, CA.


Copyright © 2007 Massachusetts Medical Society.
Prevention of Death in COPD

TO THE EDITOR: Calverley et al. do not sufficiently emphasize some aspects of their study on the use of salmeterol and fluticasone in patients with chronic obstructive pulmonary disease (COPD) (Feb. 22 issue). Their study, called the Towards a Revolution in COPD Health (TORCH) trial, showed that treatment with fluticasone alone actually increased mortality at the end of 3 years, although the increase was not significant. This finding contrasts markedly with retrospective analyses and meta-analyses showing a substantial reduction in mortality from all causes by about 25% associated with the drug. This discrepancy between the results of a well-conducted, randomized, controlled trial and historical analyses highlights how misleading the latter may be. The net effect of therapy with inhaled corticosteroids for patients who have COPD may be detrimental in view of the increased episodes of pneumonia associated with such agents. Another important result of the TORCH study was the failure of inhaled corticosteroids, even when combined with salmeterol, to reduce the annual decline in lung function. The lack of effect of inhaled corticosteroids on mortality and disease progression may reflect resistance to the antiinflammatory effects of corticosteroids in COPD.

Peter J. Barnes, D.M., D.Sc.
National Heart and Lung Institute
London SW3 6LY, United Kingdom
p.j.barnes@imperial.ac.uk

Dr. Barnes reports being a member of advisory boards at and receiving research funding from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis, and Pfizer. No other potential conflict of interest relevant to this letter was reported.

results of the factorial analysis support the conclusions of the editorial by Rabe2 accompanying the TORCH report.

Considering that the combination of salmeterol and fluticasone is superior to either drug alone in reducing exacerbations of COPD and improving health status and lung function, the reduction of mortality associated with salmeterol alone should be balanced against the more favorable pattern of symptomatic effects of the combination of the two drugs, with allowances made for the increased frequency of pneumonia.

Carlo La Vecchia, M.D.
Istituto di Ricerche Farmacologiche Mario Negri
20157 Milan, Italy
lavecchia@marionegri.it

Leonardo M. Fabbri, M.D.
University of Modena e Reggio Emilia
41100 Modena, Italy

Dr. La Vecchia has been asked by GlaxoSmithKline to review and comment on the TORCH study, and financial aspects are under discussion. Dr. Fabbri reports receiving consulting and lecture fees from Altana Pharma, AstraZeneca, Boehringer Ingelheim, Chiesi Farmaceutici, GlaxoSmithKline, Merck, Novartis, Roche, and Pfizer and grant support from Altana Pharma, AstraZeneca, Boehringer Ingelheim, Menarini, Miat, Schering-Plough, Chiesi Farmaceutici, GlaxoSmithKline, Merck, UCR, and Pfizer. No other potential conflict of interest relevant to this letter was reported.


TO THE EDITOR: By assigning a group of symptomatic patients with COPD to a placebo group without the use of long-acting bronchodilators, Calverley et al. violated paragraph 29 of the Declaration of Helsinki, which states that “the benefits, risks, burdens and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and thera-

---

Table 1. Analysis of Main Effects of Treatment with Salmeterol and Fluticasone on Mortality.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Placebo (A)</th>
<th>Salmeterol (B)</th>
<th>Fluticasone (C)</th>
<th>Salmeterol plus Fluticasone (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>1524</td>
<td>1521</td>
<td>1534</td>
<td>1533</td>
</tr>
<tr>
<td>No. of deaths</td>
<td>231</td>
<td>205</td>
<td>246</td>
<td>193</td>
</tr>
<tr>
<td>Probability of death at 3 yr (%)</td>
<td>15.2</td>
<td>13.5</td>
<td>16.0</td>
<td>12.6</td>
</tr>
<tr>
<td>Hazard ratio (95% confidence interval)</td>
<td>0.81 (0.70–0.94)</td>
<td>1.00 (0.87–1.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square</td>
<td>8.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Data for each treatment were adjusted for the other treatment.
To the Editor:

In the editorial accompanying the report on the TORCH study, Rabe refers to a 25% prevalence of venous thromboembolism in patients hospitalized with a severe exacerbation of COPD, citing a study by Tillie-Leblond et al. An instinctive response of defensive medical practice may now be to order a computed tomographic scan with a pulmonary-embolism protocol for every patient with a severe COPD exacerbation on the assumption that a 25% return must be cost-effective.

Tillie-Leblond et al. actually report that “25% of COPD patients with severe unexplained breathlessness have been shown to have pulmonary embolism.” The subjects were selected because of “the absence of a lower respiratory tract infection” — in other words, an exacerbation of COPD “of unknown origin.” Experience tells us that the prevalence of venous thromboembolism with exacerbations of COPD is low — only 3.3% in a recent study of 123 consecutive patients. The discussion in that report echoes that of Robin and McCauley, who bemoaned the ingrained belief that a low partial pressure of arterial oxygen is of positive predictive value for the diagnosis of venous thromboembolism.

Niall Keaney, M.B., Ph.D.
Sunderland Royal Hospital
Sunderland SR4 7TP, United Kingdom
niall.keaney@chs.northy.nhs.uk


The Authors Reply:

Barnes is concerned that inhaled corticosteroids increase mortality in COPD. Although in our study, numerically more patients died in the group that received fluticasone alone than in the placebo group (hazard ratio, 1.06), the difference was not significant (P=0.53). Space limitations precluded a discussion of differences between our data and those from database analyses, but we agree that a large, randomized, controlled trial such as ours provides a more robust test of the effect of inhaled corticosteroids on survival than do such analyses. The rate of decline in lung function cannot be inferred from the spirometric data in our study.

La Vecchia and Fabbri have undertaken an interesting post hoc analysis suggesting that the salmeterol component has a substantial effect on mortality. Factorial analysis assumes that each treatment has the same additive effect in the absence and presence of the other treatment. This was not the case for the TORCH trial. Our data show the clear clinical superiority of combination treatment with salmeterol and fluticasone, including fewer exacerbations and better health status.

Kupfer and Tessler suggest that our study was unethical. During the trial design, there was concern about the safety and efficacy of the component treatments, questions that our data have resolved. Patients in the placebo group received regular short-acting bronchodilators, but our results show that this treatment will not be an appropriate standard of care in the future. The degree of bronchodilator reversibility in our patients was similar to that in other large COPD trials, which failed to show any relationship between clinical outcomes and reversibility.

Duerden suggests that treatment with inhaled corticosteroids may do more harm than good. More patients with pneumonia were reported in groups that received inhaled corticosteroids, although there was no disproportionate mortality from pneumonia among patients receiving inhaled corticosteroid monotherapy, nor did the overall rate of hospitalization for COPD differ from that in the placebo group. We agree that more data are needed to better understand this finding. Large data sets will be needed, since pneumonia was relatively infrequent, as compared with other serious outcomes (about 1000 cases of pneumonia in 780 patients vs. 13,000 exacerbations of COPD in 4000 patients). We calculated
the number needed to treat as the number of patients required to treat one exacerbation, according to published methods. This number is not the same as the number needed to treat to prevent one patient from having an exacerbation, a number more appropriate for a binary event such as mortality. Although attractive in clinical practice, the number needed to treat should be viewed with caution, since it depends on the background event rate in the population under study.

Peter Calverley, M.D.
University Hospital Aintree
Liverpool L9 7AL, United Kingdom
pmacal@liverpool.ac.uk

**Posaconazole Prophylaxis in Hematologic Cancer**

**TO THE EDITOR:** Ullmann et al. and Cornely et al. (Jan. 25 issue) report on posaconazole prophylaxis in patients with hematologic cancers. Ullmann et al. found that posaconazole was superior to fluconazole for protection against invasive aspergillosis, and Cornely et al. found that posaconazole was superior to fluconazole and also to itraconazole in preventing fungal infections. After the widespread use of fluconazole and voriconazole as prophylaxis and antifungal treatment, an increase in the risk of infections with resistant fungi was observed. Selection pressure due to continuous exposure to azoles appears to play a crucial role in the emergence of resistance to these drugs. Prophylactic use of such a highly active and broad-spectrum antifungal agent as posaconazole, even in high-risk patients, could favor the emergence and amplification of resistant strains. In addition, such use might be associated with a risk of cross-resistance with other azoles, reducing their efficacy in the treatment of life-threatening fungal infections.

Stefan Weiler, M.D.
Romuald Bellmann, M.D.
Innsbruck Medical School
A-6020 Innsbruck, Austria
romuald.bellmann@i-med.ac.at

Dr. Weiler reports receiving salary support from Pfizer and Torres-Chiesi. No other potential conflict of interest relevant to this letter was reported.

**TO THE EDITOR:** Aspergillosis and infection with not-uncommon opportunistic fungi were the most frequent breakthrough infections in the studies of posaconazole prophylaxis reported by Ullmann et al. and Cornely et al. One point mutation in aspergillus cytochrome P-450 (CYP) demethylase could result in posaconazole resistance. It is unclear whether the cases of aspergillosis occurred because aspergillus isolates developed resistance or tolerance to posaconazole or because some patients with mucositis, poor oral intake, or both had suboptimal posaconazole levels. Since another triazole, voriconazole, has emerged as the preferred treatment for aspergillosis, it will be important to determine whether preexposure of

Julie Anderson, M.A.
GlaxoSmithKline
Greenford UB6 0HE, United Kingdom

Bartolome Celli, M.D.
Caritas St. Elizabeth’s Medical Center
Boston, MA 02135-2997


molds colonizing the respiratory tract (especially aspergillus) to posaconazole would result in a lack of fungicidal activity of subsequent voriconazole-based treatment for breakthrough aspergillosis. Risk stratification based on host characteristics, description of the outcomes of treatment for breakthrough aspergillosis in the setting of posaconazole prophylaxis, and analysis of cost-effectiveness are needed to identify the subgroup of high-risk patients with hematologic cancer who may benefit from posaconazole as primary prophylaxis.

Dimitrios P. Kontoyiannis, M.D.
University of Texas M.D. Anderson Cancer Center
Houston, TX 77030
dkontoyi@mdanderson.org

Russell E. Lewis, Pharm.D.
University of Houston
Houston, TX 77030

Dr. Kontoyiannis reports receiving research support and honoraria from and serving on speakers’ bureaus for Merck, Fujisawa, and Enzon and serving on an advisory board for Schering–Plough. Dr. Lewis reports receiving research support from Merck, Fujisawa, and Pfizer and serving on advisory boards for Pfizer and Schering–Plough. No other potential conflict of interest relevant to this letter was reported.


TO THE EDITOR: Cornely et al. state that in patients undergoing chemotherapy for acute myelogenous leukemia or the myelodysplastic syndrome, posaconazole improved overall survival and was more effective in preventing invasive fungal infections than either fluconazole or itraconazole. Patients were randomly assigned to receive posaconazole or either fluconazole or itraconazole, but only one of the two comparison drugs was used at each center. This protocol implies two separate randomized trials: one comparing posaconazole (in 239 patients) with fluconazole (in 240 patients), and the second comparing posaconazole (in 65 patients) with itraconazole (in 58 patients). In the Supplementary Appendix (available with the full text of the article at www.nejm.org), the authors partly present the results in this way, and we agree with them that there is not sufficient statistical power to draw any conclusion from the latter comparison. The statement that posaconazole is better than itraconazole is therefore not upheld by the results of this trial.

Cees van Nieuwkoop, M.D.
Jaap T. van Dissel, M.D., Ph.D.
Leiden University Medical Center
2300 RC Leiden, the Netherlands
c.van_nieuwkoop@lumc.nl

TO THE EDITOR: With regard to the study by Ullmann et al., it would be useful to know whether the numbers of patients with any hepatic adverse events and any hepatic serious adverse events in the posaconazole group exceeded those in the fluconazole group, and whether the patients with increased hepatic enzymes in the posaconazole group had levels that exceeded three times the upper limit of the normal range. Such findings may be harbingers of serious hepatic injury.1

David S. Krause, M.D.
147 Sawgrass Dr.
Blue Bell, PA 19422
dskmd@aol.com

Dr. Krause reports serving as a consultant for Nektar Therapeutics. No other potential conflict of interest relevant to this letter was reported.

we cannot speculate on the attributes of mutations that may arise as a result of posaconazole prophylaxis. Nevertheless, we agree with Weiler and Bellmann and with Kontoyiannis and Lewis that longitudinal surveillance studies are warranted.

As noted by Krause, hepatobiliary disorders are a known side effect of azole therapy, but they are also caused by graft-versus-host disease (GVHD). Patients with GVHD, and especially hepatic GVHD, were allowed to enter our trial with aminotransferase levels that were up to 10 times the upper limit of the normal range, regardless of serum bilirubin levels — values higher than those permitted in other trials. Owing to the randomized, double-blind design of our study, evaluation of changes in liver function in the posaconazole group as compared with the fluconazole group can be considered unbiased. The distribution of treatment-emergent adverse events affecting liver function was similar in the posaconazole and fluconazole groups in our study, as well as in the study of patients with neutropenia by Cornely et al. (Table 1). Post hoc analysis of changes in aminotransferase levels and total bilirubin levels during the exposure period showed similar rates of change in the posaconazole and fluconazole groups (Table 2). The change in both groups re-

Table 1. Treatment-Emergent Hepatic Adverse Events.*

<table>
<thead>
<tr>
<th>Study Population and Hepatic Abnormality</th>
<th>Posaconazole†</th>
<th>Fluconazole or Itraconazole (N = 298)</th>
<th>Fluconazole‡</th>
<th>Itraconazole (N = 58)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Events</td>
<td>Severe or Life-Threatening Events</td>
<td>All Events</td>
<td>Severe or Life-Threatening Events</td>
</tr>
<tr>
<td>Patients with GVHD (study by Ullmann et al.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubinemia</td>
<td>32 (11)</td>
<td>13 (4)</td>
<td>28 (9)</td>
<td>14 (5)</td>
</tr>
<tr>
<td>Increased γ-glutamyltransferase</td>
<td>25 (8)</td>
<td>13 (4)</td>
<td>31 (10)</td>
<td>21 (7)</td>
</tr>
<tr>
<td>Increased hepatic enzymes</td>
<td>21 (7)</td>
<td>14 (5)</td>
<td>21 (7)</td>
<td>8 (3)</td>
</tr>
<tr>
<td>Abnormal hepatic function</td>
<td>14 (5)</td>
<td>6 (2)</td>
<td>18 (6)</td>
<td>10 (3)</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>2 (1)</td>
<td>0</td>
<td>3 (1)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Jaundice</td>
<td>24 (8)</td>
<td>4 (1)</td>
<td>15 (5)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Increased aspartate aminotransferase</td>
<td>11 (4)</td>
<td>2 (1)</td>
<td>11 (4)</td>
<td>6 (2)</td>
</tr>
<tr>
<td>Increased alanine aminotransferase</td>
<td>24 (8)</td>
<td>10 (3)</td>
<td>20 (7)</td>
<td>11 (4)</td>
</tr>
<tr>
<td>Patients with neutropenia (study by Cornely et al.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilirubinemia</td>
<td>27 (9)</td>
<td>11 (4)</td>
<td>34 (11)</td>
<td>19 (6)</td>
</tr>
<tr>
<td>Increased γ-glutamyltransferase</td>
<td>17 (6)</td>
<td>8 (3)</td>
<td>11 (4)</td>
<td>1 (&lt;1)</td>
</tr>
<tr>
<td>Increased hepatic enzymes</td>
<td>17 (6)</td>
<td>7 (2)</td>
<td>16 (5)</td>
<td>4 (1)</td>
</tr>
<tr>
<td>Abnormal hepatic function</td>
<td>2 (1)</td>
<td>1 (&lt;1)</td>
<td>3 (1)</td>
<td>1 (&lt;1)</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>1 (&lt;1)</td>
<td>1 (&lt;1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jaundice</td>
<td>11 (4)</td>
<td>2 (1)</td>
<td>12 (4)</td>
<td>3 (1)</td>
</tr>
<tr>
<td>Increased aspartate aminotransferase</td>
<td>8 (3)</td>
<td>0</td>
<td>10 (3)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Increased alanine aminotransferase</td>
<td>15 (5)</td>
<td>4 (1)</td>
<td>13 (4)</td>
<td>6 (2)</td>
</tr>
</tbody>
</table>

* Treatment-emergent hepatic adverse events were defined according to the National Cancer Institute’s Common Toxicity Criteria (version 2.0, revised March 23, 1998). Adverse events were graded as severe or life-threatening by the investigators.
† The posaconazole group consisted of 301 patients in the study by Ullmann et al. and 304 patients in the study by Cornely et al.
‡ The fluconazole group consisted of 299 patients in the study by Ullmann et al. and 240 patients in the study by Cornely et al.
flects the complicated underlying disease, azole use in general, or both.

Andrew J. Ullmann, M.D.
Johannes Gutenberg University
55101 Mainz, Germany
ullmann@uni-mainz.de

Pranatharthi Chandrasekar, M.D.
Harper Hospital
Detroit, MI 48201

Hernando Patino, M.D.
Schering–Plough Research Institute
Kenilworth, NJ 07033


**Table 2. Selected Results of Liver-Function Tests during the Exposure Period in Patients with Graft-versus-Host Disease.**

<table>
<thead>
<tr>
<th>Result of Liver-Function Test</th>
<th>Posaconazole (N = 275)</th>
<th>Fluconazole (N = 274)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no. of patients (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exceeding the upper limit of the normal range</td>
<td>77 (28)</td>
<td>68 (25)</td>
</tr>
<tr>
<td>At least 2 times the upper limit of the normal range</td>
<td>22 (8)</td>
<td>21 (8)</td>
</tr>
<tr>
<td>At least 3 times the upper limit of the normal range</td>
<td>10 (4)</td>
<td>12 (4)</td>
</tr>
<tr>
<td>At least 5 times the upper limit of the normal range</td>
<td>4 (1)</td>
<td>3 (1)</td>
</tr>
<tr>
<td>At least 10 times the upper limit of the normal range</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Concomitant aspartate and alanine aminotransferase and total bilirubin values at least 3 times the upper limit of the normal range and total bilirubin value more than 2 times the upper limit of the normal range</td>
<td>16 (6)</td>
<td>17 (6)</td>
</tr>
</tbody>
</table>

* Data are shown for treated patients for whom results of liver-function tests were available during the exposure period and for whom at least one set of post-baseline values for aspartate aminotransferase, alanine aminotransferase, and total bilirubin, all obtained on the same day, met the specific criteria.

Kontoyiannis and Lewis hypothesize that mucositis and suboptimal drug exposure are potential reasons for the 2% failure rate in the posaconazole group. Analysis of their well-taken point reveals that none of the 10 patients with grade 3 or 4 mucositis at baseline had an invasive fungal infection. Average and maximum plasma levels of posaconazole in patients with proven or probable breakthrough infection were not different from those in the overall population. Thus, the occurrence of invasive fungal infection does not appear to be associated with a particular plasma level of posaconazole.

In a large, randomized clinical trial comparing fluconazole with itraconazole as prophylaxis in high-risk patients, the efficacy results for the oral solutions of the two drugs did not differ signifi-
cantly. Thus, there was no good evidence to support a recommendation for the use of a specific systemic prophylaxis in patients undergoing remission-induction chemotherapy. In the absence of clinical trials that proved the superior efficacy of any comparison drug in this high-risk population, we did not agree on a single drug to compare with posaconazole and instead left the choice to each participating hospital. This approach allowed us to compare posaconazole with drugs used in current clinical practice.

Oliver A. Cornely, M.D.
University of Cologne
50924 Cologne, Germany
oliver.cornely@uni-koeln.de

Andrew J. Ullmann, M.D.
Johannes Gutenberg University
55101 Mainz, Germany
Cathy Hardalo, M.D.
Schering–Plough Research Institute
Kenilworth, NJ 07033


Treatment of Symptomatic Uterine Fibroids

TO THE EDITOR: In their report on the Randomized Trial of Embolization versus Surgical Treatment for Fibroids (REST), Edwards et al. (Jan. 25 issue) carefully discuss the benefits of uterine-artery embolization as compared with hysterectomy but do not address some important potential concerns associated with embolization. The lack of tissue availability for histopathological examination after uterine-artery embolization has been reported to delay diagnosis of a concomitant cancer or a malignant mass initially misdiagnosed and treated as a uterine fibroma. Furthermore, concern has been expressed about a decrease in ovarian function after embolization. Women with anastomoses of the uterine and ovarian arteries (an uncommon condition) seem to have a predisposition to this adverse outcome. Studies to determine basal follicle-stimulating hormone and estradiol before and after the procedure have been suggested in order to monitor patients for embolization-induced follicle depletion, particularly when preservation of fertility is an important consideration.

Amitabh Parashar, M.D.
Anjali Varma, M.D.
Carilion Clinic
Roanoke, VA 24018
aparashar@carilion.com

Sudha Bedi, M.D.
69 Leroy Ave.
Valhalla, NY 10595


TO THE EDITOR: The global question of fibroid management discussed by Edwards et al. can be looked at in another way. Today, laparoscopic and vaginal approaches should be preferred to the open route for surgical treatment of symptomatic uterine fibroids. Vaginal hysterectomy is now recognized as the method of choice for benign gynecologic disease because of its effectiveness, feasibility, and postoperative advantages over open surgery (significantly shorter hospital stay and better postoperative recovery). In selected cases, laparoscopically assisted vaginal hysterectomy or total laparoscopic hysterectomy can be performed to avoid the open route; the potential risk of complications does not preclude consideration of these procedures. Comparison of uterine-artery embolization with vaginal or laparoscopic approaches in a new randomized trial would better
define the role of embolization in the treatment of women with symptomatic uterine fibroids.

Bruno Borghese, M.D.
Charles Chapron, M.D.
University Hospital Cochin
75014 Paris, France
charles.chapron@cch.aphp.fr


THE AUTHORS REPLY: Parashar et al. discuss the possibility of missing a malignant tumor with uterine-artery embolization. Although the report that they cite acknowledges that the incidence of malignant tumors is unlikely to be greater than that for other nonsurgical treatments or no treatment, this factor is always a concern, and all patients in our study underwent imaging before inclusion. Although such imaging might still miss a leiomyosarcoma, the incidence of such tumors in premenopausal women is extremely low.1 We also acknowledge the concern regarding damage to ovarian function after embolization, and this topic is the subject of ongoing research. Available data suggest that there is no long-term increase in gonadotropin levels, as would be expected with ovarian failure.2 We assessed levels of follicle-stimulating hormone (in many although not all instances on cycle day 3) at baseline and at the 12-month follow-up visit in a subgroup of 73 women in our study. We found a significant but small increase in mean levels of follicle-stimulating hormone after embolization (increase from baseline to 12 months, 6.70 to 7.80 IU per liter; P=0.01). The levels of follicle-stimulating hormone rose significantly more in women who were 45 years of age or older (8.80 to 10.10 IU per liter) than in those under the age of 45 years (5.35 to 6.10 IU per liter, P=0.02), suggesting a greater effect of embolization in older women (unpublished data).

Borghese and Chapron discuss the surgical approach to both hysterectomy and myomectomy. In our trial, all surgeries were performed by the open route, not because of the protocol but because that was the normal practice in the centers involved. A total of 75% of hysterectomies are performed by the abdominal route. Studies suggesting that vaginal hysterectomy should be preferred generally have not included women with an enlarged uterus3 and therefore describe a different group of patients. The majority of hysterectomies are performed abdominally, particularly when there are large fibroids, not because the surgeons are incompetent but because they feel it is in the patient’s interest. An analogy might be the cesarean delivery of a baby when it could be delivered vaginally only with effort, risk, and probable morbidity.

Jon Moss, M.B., Ch.B.
North Glasgow University Hospitals
Glasgow G12 0YN, United Kingdom
jon.moss@northglasgow.scot.nhs.uk

Mary Ann Lumsden, M.D.
University of Glasgow
Glasgow G31 2ER, United Kingdom

Kevin Cooper, M.D., M.Sc.
Aberdeen Royal Infirmary
Aberdeen AB25 2ZN, United Kingdom

the mean decrease was 0.6 percentage point. The smaller decreases in the glycated hemoglobin level with sitagliptin, then, are explained by a lower baseline value than those reported in older studies of metformin, sulfonylureas, and thiazolidinediones.

To illustrate this point, we have added data from published placebo-controlled trials of sitagliptin and vildagliptin to our original analysis (Fig. 1). Furthermore, head-to-head comparisons show that DPP-IV inhibitors have a glucose-lowering effect that is similar to that of the sulfonylurea glipizide, with a considerably lower likelihood of causing hypoglycemia, and to that of the thiazolidinediones, without causing weight gain. We conclude that these new agents, which are effective in combination with insulin sensitizers, will help more persons with diabetes to achieve the difficult goal of improved glycemic control.

Zachary T. Bloomgarden, M.D.
Mt. Sinai School of Medicine
New York, NY 10029
zbloomgard@aol.com

Silvio E. Inzucchi, M.D.
Yale University School of Medicine
New Haven, CT 06520

Dr. Bloomgarden reports serving on speaker panels for Takeda, GlaxoSmithKline, Novo Nordisk, Eli Lilly, Amylin, Merck, and Novartis and serving as a consultant for Merck. Dr. Inzucchi reports serving as an adviser to Merck, Takeda, Pfizer, and Novartis; receiving consulting fees from Merck; receiving honoraria from Merck and Takeda; and receiving research funding from Eli Lilly. No other potential conflict of interest relevant to this letter was reported.


Figure 1. Changes in Glycated Hemoglobin Levels According to Baseline Levels in Six Placebo-Controlled Studies of Sitagliptin and Vildagliptin, as Compared with 73 Studies of Other Agents between 1991 and 2002.

The 73 studies of other agents involved a total of 101 comparisons of treatment among 8987 persons with diabetes. Data are from six studies including Raz et al. and Ristic et al.
TO THE EDITOR: In the article by Nathan, important inaccuracies concerning the review and approval process of the Food and Drug Administration (FDA) should be clarified. The 1962 amendment to the Food, Drug, and Cosmetic Act (FDCA) requires that drug approval be based on “substantial evidence,” a requirement that is generally fulfilled by data from adequate and well-controlled studies. Approval of antidiabetic therapies typically relies on evidence from several well-controlled clinical studies, including a minimum 6-month, placebo-controlled period that continues into an extension period exceeding 1 year to further assess safety and durability of effect. This was the case with the sitagliptin approval, which was based not only on the four phase 3 studies cited by Nathan but also on an extensive nonclinical and phase 1 and 2 clinical development program. Overall, 3276 patients took sitagliptin in 34 studies, with a cumulative exposure of 1339 patient-years — far in excess of the 641 patient-years cited by Nathan. However, since even a large clinical program cannot elucidate all safety issues, the FDA closely monitors the safety of all drugs after approval and takes regulatory action that is consistent with new safety information.

Nathan implies that newer antidiabetic therapies have little clinical value if they are less potent than older agents; this suggests that superior efficacy is now necessary in drug approval. A demand for superior efficacy is not permitted under the FDCA, which calls for the determination of the safety and efficacy of a drug on its own merits. Furthermore, superiority is not the only consideration in selecting a drug for a given patient in the clinic, since relative safety, tolerability, and individual response are important factors in any clinical judgment.

We agree that the failure to use effective therapies and implement lifestyle and dietary modifications contributes to inadequate glycemic control in many persons with diabetes. Comparative clinical trials of available therapies may better inform prescribers and patients and partially address these failures. Well-designed trials of multiple agents probably will not be conducted by manufacturers, and the FDA has limited authority to demand such trials. Any such trials probably would be conducted through an alliance of academic institutions and government agencies. Such efforts to show the comparative safety, effectiveness, and costs of therapies would be more conducive to public health than regulatory measures, which might significantly hinder the development of new therapies that offer choices when other available therapies are contraindicated, are poorly tolerated, or do not meet individual needs.

Ilan Irony, M.D.
Mary H. Parks, M.D.
Robert J. Meyer, M.D.
Food and Drug Administration
Silver Spring, MD 20993-0002
ilan.irony@fda.hhs.gov


TO THE EDITOR: Nathan’s caution against the use of new treatments seems surprising, given not only the increasing prevalence of diabetes but also our declining ability to achieve control of diabetes, as evidenced by the rising average glycated hemoglobin level in the past decade. Our inability to achieve diabetes control is probably due to a lack of effective and durable treatments. Nathan’s suggestion that we continue to use primarily insulin, sulfonylureas, and biguanides (a suggestion that is also made in his recent American Diabetes Association [ADA] guideline and that is similar to the ADA’s recommendations a decade ago) is flawed, since data from the United Kingdom Prospective Diabetes Study (UKPDS) show that these treatments all fail to achieve control over time. A Diabetes Outcome Progression Trial, one of the largest studies involving persons with type 2 diabetes, showed that rosiglitazone resulted in a 32% reduction in the risk of treatment failure as compared with metformin and in a 63% reduction as compared with glyburide — results that Nathan dismissed in his editorial accompanying the report on this trial. The management of type 2 diabetes should focus on achieving as well as sustaining normoglycemia. Might some combination of a glitazone, gliptin, and rimonabant achieve this outcome? We may never find out if others insist on doing the same old thing.

Matthew L. Mintz, M.D.
George Washington University School of Medicine
Washington, DC 20037
mmintz@mfa.gwu.edu
Dr. Mintz reports receiving lecture fees from GlaxoSmithKline, Takeda, and Pfizer. No other potential conflict of interest relevant to this letter was reported.


To the Editor: Nathan criticizes the rapid approval of new drugs for hypoglycemia, stating that the current available treatment is adequate. We think this is inaccurate. Despite existing drugs, the diabetes epidemic is increasing, with relatively few patients achieving glycemic control.\(^1\) Furthermore, as shown in the UKPDS, adherence to glycemic control requires high-dose drug combinations and insulin with resulting side effects.\(^2\) Thus, we believe that new and effective drugs are needed, with future treatment based on metformin combinations.

Regarding the quick FDA approval of sitagliptin, we would like to clarify that although the only published report at the time was ours, the results of other trials have already been presented at the ADA meeting (June 9 to 13, 2006). Furthermore, trials show that metformin–sitagliptin combinations more than double the effect of metformin,\(^3\) with a possible beta-cell protective effect\(^4\) and negligible side effects (in contrast to the hepatotoxicity of troglitazone that was noticed before its FDA approval). Thus, the new gliptins are a welcome addition to our armamentarium. In combination with metformin or other drugs for hypoglycemia, the gliptins offer a safe and effective option for the treatment of diabetes.

Itamar Raz, M.D.
Roy Eldor, M.D.
Hadassah–Hebrew University
91120 Jerusalem, Israel
eldorroy@yahoo.com

THE AUTHOR REPLIES: Bloomgarden and Inzucchi note the potential importance of initial glycated hemoglobin levels in judging the relative effectiveness of new diabetes medications.\(^1\) Although their figure shows a correlation of baseline glycated hemoglobin levels with a glucose-lowering effect, it explains only 14% of the variance. Despite this modest correlation, they suggest that, considered in this light, the DPP-IV inhibitors are not as ineffective as I proposed. However, the scant data available at the time I wrote the Perspective article did not provide any direct proof that sitagliptin is as effective as less expensive agents. If pharmaceutical company-sponsored investigators want to claim similar effects of newer drugs as compared with older ones, they should perform more head-to-head comparison studies. The recently published study cited by Bloomgarden and Inzucchi was an efficacy analysis of glipizide as compared with sitagliptin in which almost one third of the subjects who underwent randomization were excluded from the primary analysis.\(^2\) More patients in the sitagliptin-treatment group discontinued therapy “for lack of efficacy.” In the intention-to-treat analysis, sitagliptin lowered glycated hemoglobin levels by 0.51%.

Irony and colleagues point to “important inaccuracies” in my Perspective article regarding the FDA approval process. They note that the FDA considered data from 30 studies, in addition to the 4 published clinical trials I considered, and that the total cumulative exposure was 1339 patient-years, “far in excess of the 641 patient-years” that I cited. First, I did not suggest that...
the FDA considered only peer-reviewed, published data. I only noted the paucity of peer-reviewed data from clinical trials for clinicians to consider. Second, even if the FDA considers 1339 patient-years of exposure to be a major improvement, the average exposure in the studies was about 5 months — for a drug that may be taken by millions of patients for many years. I join Irion et al. in decrying the limited studies that directly compare therapeutic agents.

Mintz suggests that my curmudgeonly critiques of new glucose-lowering agents, including the DPP-IV inhibitors and thiazolidinediones\(^3\) (as well as inhaled insulin\(^4\)), will inhibit our ability to maintain lower glycemia over time. I would counter that the introduction of more expensive, less effective medications does not contribute to the well-being of the population of patients with diabetes. If combinations of new medications with older ones are proposed to improve long-term diabetes control, at an expense similar to or less than that of older combinations, they should be tested.

Finally, Raz and Eldor note that data published in abstracts suggest that combination therapy with metformin has a synergistic effect. I urge peer-reviewed publication of such studies, assuming they include active comparison treatments so that the additive effects of the newer agents can be compared with older combinations.

David M. Nathan, M.D.
Massachusetts General Hospital
Boston, MA 02114
dnathan@partners.org


---

**Monoclonal Gammopathy of Undetermined Significance**

**TO THE EDITOR:** In the article concerning monoclonal gammopathy of undetermined significance (MGUS) (Dec. 28 issue),\(^1\) Bladé disregards a complication that is especially important in clinical practice. Indeed, the problem of skeletal involvement in this disease should not be viewed only as a way to rule out lytic lesions but should be placed in a more comprehensive context. We\(^2\) and others\(^3\) have reported an increased rate of vertebral fractures among patients with MGUS. This increase is probably secondary to an altered balance between the receptor activator of nuclear factor-κB ligand (RANKL) and osteoprotegerin.\(^2,4\) This is important for two reasons. First, a vertebral fracture is associated with an increased risk of other axial fractures, independently of bone mass. Second, patients with fractures should be offered drugs, such as bisphosphonates, that have demonstrated effectiveness in the secondary prevention of axial fractures. Along these lines, if an excess of RANKL activity is the pathophysiological mechanism underlying bone loss, other molecules, such as denosumab,\(^5\) will prove to be effective in the future.

Salvatore Minisola, M.D.
Jessica Pepe, M.D.
Elisabetta Romagnoli, M.D.
University of Rome La Sapienza
00161 Rome, Italy
salvatore.minisola@fastwebnet.it


**THE AUTHOR REPLIES:** Minisola et al. emphasize that the Mayo Clinic group\(^1\) and their group\(^2\) have reported an increased incidence of vertebral fractures among patients with MGUS. In contrast,
Politou et al. reported that the increased bone resorption observed in MGUS is compensated for by new bone formation; in their study, none of the patients with MGUS had apparent radiographic evidence of osteoporosis or osteopenia (although it should be recognized that this assessment was based on a skeletal survey only).

Pending further data on the risk of fracture among patients with MGUS, I would recommend that the decision to measure bone mineral density in patients with MGUS be made in accordance with the usual guidelines for the population and be predicated on the presence of other risk factors for bone loss. Likewise, treatment should be administered according to the current indications for the treatment of bone loss in the general population. The cost and possible long-term adverse effects in an asymptomatic, long-lasting condition must always be considered. In any event, the type of bisphosphonate, administration schedule, or both in patients with MGUS and bone loss would be different from the type and schedule in patients with multiple myeloma.

Joan Bladé, M.D.
University of Barcelona
08036 Barcelona, Spain
jblade@clinic.ub.es

Since publication of his article, Dr. Bladé reports receiving lecture fees and consulting fees from Novartis. No other potential conflict of interest relevant to this letter was reported.


Pulmonary-Valve Endocarditis

TO THE EDITOR: Isolated pulmonary-valve endocarditis is quite rare, but it shares demographic, clinical, and microbiologic features with the more common tricuspid endocarditis. We describe a 35-year-old intravenous drug abuser who presented with Staphylococcus aureus septicemia and fulminating circulatory failure, complicated by multiple lung abscesses and associated bronchopleural fistula. High-dose vasopressor and inotropic support and positive-pressure ventilation were required, in addition to intravenous antibiotics.

Echocardiography showed a large mobile mass in the right ventricular outflow tract, arising from the pulmonary valve and extending into the right pulmonary artery, causing severe pulmonary regurgitation. The right ventricle was dilated and hypokinetic. The tricuspid valve was structurally normal; however, moderate functional tricuspid regurgitation was present.

Surgery was performed on day 5 of the patient’s hospital stay. The pulmonary trunk was opened during cardiopulmonary bypass, and a vegetation, 9 cm in length, was removed (Fig. 1). The entire pulmonary root, up to the bifurcation of the pulmonary trunk, was excised and replaced with a pulmonary homograft. The patient had an excellent recovery. The bronchopleural fistula was managed conservatively and resolved. Six months
later, the patient remained well, in New York Heart Association functional class I.

It is sometimes argued that right-sided endocarditis is better tolerated and more likely to respond to medical therapy than infection of the mitral and aortic valves. Consequently, an extended trial of antibiotic treatments in these patients may be appealing, especially when there is a risk of prosthetic-valve infection due to recidivism of intravenous drug abuse. The results of retrospective studies suggest that vegetations less than 1 to 2 cm long in right-sided endocarditis usually respond to medical treatment.\(^2,3\)

Whereas for left-sided endocarditis there are established evidence-based guidelines,\(^4\) the indications for surgery and its timing in patients with right-sided endocarditis are much less clear. However, on the basis of current evidence, we suggest that infection with staphylococcus, the presence of large vegetations (\(>2\) cm long), or cardiovascular instability should prompt consideration of early surgical intervention in patients with this condition.

Nicholas Kang, F.R.A.C.S.
Warren Smith, F.R.A.C.P.
Sally Greaves, F.R.A.C.P.
David Haydock, F.R.A.C.S.
Green Lane Cardiothoracic and Cardiovascular Service Auckland 1148, New Zealand

nkang@adhb.govt.nz


Correspondence Copyright © 2007 Massachusetts Medical Society.
Natriuretic Peptides: The Hormones of the Heart

The 1981 discovery by Adolfo J. de Bold of atrial natriuretic peptide, a peptide hormone secreted by the heart in response to volume expansion, revealed a homeostatic mechanism that counterbalanced the salt- and water-retaining actions of the renin–angiotensin–aldosterone system that predominates in terrestrial animals. This finding was followed by the recognition of additional members of this peptide family in many other tissues, including brain tissue, and the discovery of receptors for these peptides throughout the body. This news beguiled physiologists, physicians, and pharmaceutical companies and led to the expectation that the peptides would be the basis of treatments for diseases in which salt and fluid retention is an aggravating factor — these peptides, it was predicted, would be useful in the treatment of a variety of diseases, including hypertension, acute renal failure, and congestive heart failure. The failure of these peptides to reverse or at least reduce some of these conditions has been disappointing, but not all of the findings have been frustrating. Over the past several years, the measurement of circulating natriuretic peptides has emerged as an important tool in the recognition of, and perhaps also the management of, heart failure.

This book reviews these peptides — their discovery, their involvement in cardiac and cardiovascular physiology, the available clinical tests, their role in cardiovascular and other diseases, and current clinical trials of the peptides in patients with heart disease. The editors and authors have been actively involved in studies of the physiological and clinical role of the peptides in cardiovascular disease, as well as in the development and characterization of the laboratory tests used to measure them in clinical and laboratory settings. This is a useful book for scientists and clinicians alike, summarizing the vast clinical and laboratory experience of its contributors and reviewing most of the available literature on natriuretic peptides.

The book contains seven chapters that can be read independently; each chapter ends with its own summary and conclusions. Several of the topics covered are worth noting. Natriuretic peptides are involved in a multitude of actions throughout the body, serving in endocrine, paracrine, and autocrine functions and possibly as neurotransmitters and immunomodulators. The methods used to measure circulating natriuretic peptides are varied. Some of these methods are restricted to research laboratories, whereas others have found their way into clinical practice. The authors and editors of this book call for standardization of these methods. The value of the measurement of natriuretic peptides as a screening and prognostic tool for heart disease, and as a means to differentiate various types of dyspnea, is appropriately emphasized. The failure of natriuretic peptides, or agents that increase their circulating levels, to reverse salt and water retention in congestive heart failure is very likely due to the tonic effect of counterregulatory systems. The authors postulate that combination therapies that reduce or eliminate these opposing mechanisms could be useful and should be the subject of future research.

Patricio Silva, M.D.
Temple University
Philadelphia, PA 19140

Angiogenesis: From Basic Science to Clinical Applications

In 1989, two consecutive reports in Science described the cloning of vascular endothelial growth factor (VEGF) (the protein originally identified by Harold Dvorak and colleagues as vascular
permeability factor), a landmark in the field of angiogenesis. Subsequent studies and clinical trials have since confirmed that VEGF is the dominant angiogenic factor in numerous diseases. Furthermore, numerous recent clinical trials have shown that the inhibition of VEGF can benefit patients with a variety of diseases in which angiogenesis has a role. Napoleone Ferrara, the editor of *Angiogenesis*, was an author of one of the 1989 reports in *Science*; his laboratory not only cloned the VEGF gene but also developed the anti–VEGF monoclonal antibodies bevacizumab and ranibizumab, now approved by the Food and Drug Administration (FDA) for the treatment of malignant and ocular diseases, respectively.

*Angiogenesis* is a thorough overview of the basic science of the field, with an appropriate focus on VEGF biology. In addition to containing comprehensive discussions of VEGF biology and its role in angiogenesis and lymphangiogenesis, the book includes several chapters on the role of VEGF in processes other than angiogenesis, including the effects of VEGF on the nervous system. The book is predominantly devoted to the VEGF signaling pathway — an appropriate choice, as proangiogenic and antiangiogenic approaches in clinical medicine are based almost entirely on aberrations of VEGF signaling.

The strength of the book lies in the expertise and insights of the contributors, who have all made important contributions to the field, including work on the cloning and characterization of numerous angiogenic mediators. The chapters are brief and include clear, easily comprehensible figures that would make excellent slides for lectures on the basic biology of angiogenesis. Several color plates, grouped in the middle of the book, replicate important black-and-white figures from various chapters.

Ferrara clearly states in the preface that *Angiogenesis* is not a comprehensive overview of the field. However, because the fields of basic science and clinical medicine are merging, an expanded discussion on anti-VEGF therapy for malignant diseases would have been beneficial. The single chapter addressing anti-VEGF therapy in patients with cancer is limited to the use of bevacizumab, and most of the discussion focuses on studies in colorectal cancer. Although most clinical studies are in this area, it is important to point out that tyrosine kinase inhibitors that target VEGF receptors have been shown to be effective and are approved by the FDA for the treatment of renal-cell carcinoma.

Overall, *Angiogenesis* is an outstanding overview of the topic in general and of VEGF biology in particular. Clinicians will want to complement the book’s content by reading recent print and online reviews for more up-to-date results from clinical trials in cancer, ocular diseases, and cardiovascular diseases. Because the outcomes of ongoing clinical trials may have an immediate effect on the care of patients, clinicians must stay apprised of real-time changes in the field by seeking various sources for updates of clinical trial results and FDA approvals.

Lee M. Ellis, M.D.
University of Texas M.D. Anderson Cancer Center
Houston, TX 77230

---

**CONGESTIVE HEART FAILURE**


**T**hirteen years have elapsed since the first edition of this book was published. At that time, heart failure was just being recognized
as a cardiovascular condition that had major effects on public health, and that recognition was probably the impetus for writing such a book in the first place. Since then, the remarkable pace of change in our understanding of heart failure has been driven by advances in the biologic sciences, engineering, information technology, and the funding of large trials by industry. The evidence base for choices in the treatment of heart failure is probably greater than in any other area of medicine, partly because mortality from heart failure is so high. Clinical trials of new therapies with mortality as a component of the end point are not prohibitively costly or impractical in size. But there is the rub. As many as 40% of patients die in the 6 months after initial presentation, and the attrition rate thereafter is about 10% per year, so that 50% are dead 3 years after initial presentation, despite new treatments. Although advances have been made to improve survival, the outlook for the patient with heart failure remains poor. The problem of cell loss exceeding the slow rate of cell regeneration in the human heart has not been resolved.

This book is an encyclopedia of knowledge related to heart failure, and for that reason it will be valued by specialists. The chapters, written by leading authorities, are quite detailed and well documented — one chapter has 451 references. Nearly all the contributing authors are American, however, and many ignore data from other parts of the world. A great problem for the editors is that heart failure is now so pervasive in cardiology that a book such as this can rapidly expand into one that seems to cover all of cardiovascular medicine. Most chapters include recent references, and the opening paragraphs identify what is new. There is a well-balanced chapter on the use of digoxin. As in many recent guidelines, simple algorithms for treatment and management are not prominent. The recently revised definition of cardiomyopathy is missing, and topics such as the increasing use of alternative medicines and methods for the delivery of care are largely absent. The book would have been improved by the use of color in some otherwise excellent illustrations.

Several controversial topics are avoided in the book. The dichotomization of heart failure into the categories of systolic and diastolic is an error, because the ejection fraction in patients with heart failure follows an almost normal distribution. Patients should not be placed in imaginary categories; instead, an answer should be sought to the question of why the heart enlarges in some patients with heart failure and not in others. Many readers will wish for the early death of the ugly term HFnlEF, used in the book as an abbreviation for heart failure with a normal ejection fraction. The sections on cardiac resynchronization therapy and the implantable cardioverter–defibrillator persist, as do guidelines, in using the entry criteria of major trials, rather than the characteristics of the actual population in the trials, as a means of identifying who should receive these devices. For cases in which the two populations differ greatly, the actual study population is preferred.

More emphasis should have been given to the importance of adherence to treatment and methods for the delivery of health care to the whole population rather than just within special clinics. Several recent trials have shown that when high-quality care is given to patients with heart failure both within and outside special clinics, there is a greater effect on morbidity and mortality than when many pharmacologic and invasive interventions are used. Over the next few years, the greatest benefit to patients will be achieved by the application of what we know rather than by the introduction of new treatments. If a physician wants to know what we know, this book is the place to start.

Philip A. Poole-Wilson, M.D.
Imperial College London
London SW3 6LY, United Kingdom
p.poole-wilson@imperial.ac.uk

CORRECTION

Orotracheal Intubation (April 26, 2007;356:e15). In the PDF summary of the video, the third sentence under “Confirmation” should have read, “For children, you can use the following formula to estimate the proper depth of tube insertion1: tube depth (in centimeters) = ((child’s age in years) × 2 + 12),” rather than “tube depth = ((child’s age in years) × 2) + 12.” The text has been corrected on the Journal’s Web site at www.nejm.org.
Central Venous Catheterization

Alan S. Graham, M.D., Caroline Ozment, M.D., Ken Tegtmeyer, M.D., Susanna Lai, M.P.H., and Dana A.V. Braner, M.D.

INDICATIONS
Central venous catheterization provides a route for delivery of caustic or critical medications and allows measurement of central venous pressure.

CONTRAINDICATIONS
General contraindications for the placement of a central venous catheter include infection of the area overlying the target vein and thrombosis of the target vein; site-specific and relative contraindications include coagulopathy, although this is not an absolute contraindication. Extreme care must be exercised in patients with coagulopathy and in other patients for whom complications would be life-threatening.

EQUIPMENT
Many institutions stock prepackaged catheter-insertion kits containing the necessary equipment. The catheter should have the appropriate lumen size to deliver the required medications, and its length should be appropriate to reach the junction of the vena cava and the right atrium. Approximate length can be measured against the patient’s external anatomical landmarks. Seven-French 20-cm catheters are the most commonly used. Dialysis or rapid fluid resuscitation requires larger-bore catheters. Each additional lumen decreases the size of the individual lumens, which will decrease the maximal rate at which fluids can be administered. The catheter should be flushed, and compatibility between the guide wire and the needle should be confirmed.

PREPARATION
Explain the procedure to the patient, and obtain written informed consent. Select the insertion site on the basis of the comparisons noted in Table 1. Subclavian and inter-

<table>
<thead>
<tr>
<th>Table 1. Risk of Complications Associated with Internal Jugular, Subclavian, and Femoral Central Venous Catheterization.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complication</strong></td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Pneumothorax (%)</td>
</tr>
<tr>
<td>Hemothorax (%)</td>
</tr>
<tr>
<td>Infection (rate per 1000 catheter-days)</td>
</tr>
<tr>
<td>Thrombosis (rate per 1000 catheter-days)</td>
</tr>
<tr>
<td>Arterial puncture (%)</td>
</tr>
<tr>
<td>Malposition</td>
</tr>
</tbody>
</table>

* NA denotes not applicable.
nal jugular sites are generally preferred because they present a lower risk of infection and fewer mechanical complications. If the patient has challenging anatomy, a scar at the insertion site, or any other indication that could result in a difficult insertion, an expert operator should be in attendance.

Anatomical landmarks for the central approach to internal jugular venous catheterization begin at the apex of the triangle formed by the heads of the sternocleidomastoid muscle and the clavicle. A confluence between the internal jugular vein and the brachiocephalic vein facilitates cannulation at this location. After identifying the landmarks, sterilize the area with chlorhexidine, using a circular motion from the center outward, and then apply a sterile drape.

Administer local anesthesia, using 1 to 2 ml of 1% lidocaine or equivalent, with a 25-gauge needle at the cannulation site. To avoid air embolism, place the patient with head down, in the Trendelenburg position. The head should be rotated 45 degrees away from the site of cannulation; avoid excessive rotation of the head, which can cause collapse of the vein. During the procedure, place the index finger of your nondominant hand on the patient’s carotid artery to diminish the risk of inadvertent puncture of the artery.

Ultrasound Guidance

In numerous studies, ultrasound guidance has been shown to increase the success of first-time catheter placement and to decrease the risk of complications. When using ultrasound guidance, enlist an assistant either to handle the probe or to remove it when it is no longer needed.

The vein and artery appear circular and black on the ultrasound image; the vein is much more compressible when gentle pressure is applied to the skin via the probe. The needle appears echogenic and can be followed into the image of the vein on ultrasound. Newer commercial kits include needles that are more echogenic.

The Procedure

Starting just lateral to the carotid pulse, insert an 18-gauge needle slightly superior to the apex of the triangle. The needle is maintained at an angle of 20 degrees above the coronal plane as it is advanced past the apex of the triangle, with the longitudinal axis in the direction of the ipsilateral nipple. The vein is generally encountered approximately 0.5 in. (1.3 cm) under the skin, though this can vary, depending on regional adiposity.

After venous access is obtained, hold the needle carefully as you disconnect the syringe. The J-shaped end of the guide wire is introduced into the needle and advanced. The wire should thread easily, without resistance, well beyond the end of the needle. If cardiac rhythm changes are noted, pull the wire back until the rhythm normalizes. Then remove the needle, leaving the wire in place. Carefully maintain control of the wire, and make a 1-to-2-mm incision at the site of skin puncture. Advance the dilator over the guide wire. Once the tract is dilated, remove the dilator and thread the catheter over the wire and into the vessel. Then remove the guide wire, confirm blood return, and apply a sterile dressing.

Complications

Risks associated with central venous catheterization include infectious, mechanical, and thrombotic complications. A chest radiograph should be obtained to confirm placement and to assess for complications.

Catheter infections occur by means of one of three mechanisms: local insertion-site infection, which travels down the catheter externally; or hub colonization followed by infection via the intralumenal route or via hematogenous seeding of the
catheter. The Institute for Healthcare Improvement recommends five steps to reduce central-line infections: hand hygiene, adherence to maximal barrier precautions, chlorhexidine skin antisepsis, selection of an optimal catheter site, and daily review of the necessity of the catheter, with prompt removal when the catheter is no longer needed. Implementation of these steps has been conclusively shown to decrease the rate of catheter-related bloodstream infection. Scheduled changing of a catheter over a guide wire or moving a catheter to a new site can increase mechanical and infectious complications, and neither is recommended. Antiseptic-containing hubs and antimicrobial-impregnated catheters have been shown to decrease the rate of catheter-related bloodstream infections. Topical antibiotic ointments are ineffective, promote antibiotic-resistant bacteria, and increase fungal colonization.

Mechanical Complications
Mechanical complications include arterial puncture, hematoma, pneumothorax, hemothorax, arrhythmia, and improper location of the catheter, whether in an accessory vein or in the other vessels of the upper vascular system. Insertion of a catheter into the femoral vein, not shown in this video, has the highest risk of mechanical complications, but the rates of serious mechanical complications for femoral and subclavian insertion are similar. If an artery is punctured, further attempts at that site should be abandoned, and access to an alternative site should be attempted. Internal jugular and subclavian cannulation sites are preferred because of their lower overall rate of mechanical complications. However, these sites carry a small risk of hemothorax and pneumothorax. Ultrasound guidance for internal jugular cannulation significantly reduces the number of attempts required and the risk of complications.

Thrombotic Complications
Central venous cannulation increases the risk of central venous thrombosis, with the concomitant potential risk of venous thromboembolism. Thrombosis may occur as early as the first day after cannulation. The site with the lowest risk for thrombotic complications is the subclavian vein. Prompt removal of the catheter when it is no longer needed decreases the risk of catheter-related thrombosis.

No potential conflict of interest relevant to this article was reported.
A 65-YEAR-OLD MAN WITH A HISTORY OF HYPERTENSION, DIABETES MELLITUS, coronary artery disease with angioplasty and bypass grafting, and deep-vein thrombosis (with placement of an inferior vena cava filter several months earlier) was admitted with heart failure. A right internal jugular catheter was placed for management of congestive heart failure. During placement of the catheter, the guidewire was advanced approximately 50 cm; subsequently, there was difficulty in removing the guidewire, requiring some force to pull it out. A radiograph obtained earlier, confirming the proper placement of a feeding tube, showed that the filter was in the proper position (Panel A). A radiograph obtained after catheter placement showed that the filter was dislodged and in the superior vena cava (Panel B). The dislodged filter was removed through a filter sheath without complication. An inferior venacavogram revealed focal irregularity of the mid-infrarenal inferior vena cava, representing sites where the filter had been attached, but no contrast extravasation was identified. The patient died 2 weeks later from progressive respiratory failure. To avoid this type of complication, advancement of the guidewire should be limited to approximately 10 cm before threading of the catheter.

Copyright © 2007 Massachusetts Medical Society.