The NEW ENGLAND JOURNAL of MEDICINE

Perspective

AUGUST 4, 2005

Public Solicitation of Organ Donors

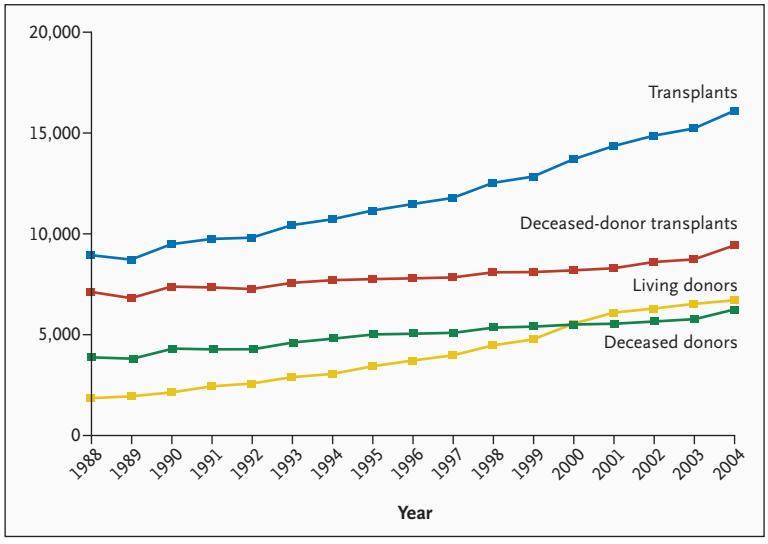
Robert Steinbrook, M.D.

Related articles, pages 444 and 447

During the past half-century, organ donation has been fertile territory for both extraordinary compassion and complex ethical questions. As transplantation has become safer and outcomes have improved, the rules for donation and the fair allocation of organs have struggled to keep pace.

In 2004, there were about 27,000 solid-organ transplantations in the United States — nearly 1600 more than there had been in 2003. Yet the demand for organs remains far greater than the supply. As of July 2005, there were about 89,000 people on waiting lists. For the largest group — the 62,500 patients awaiting kidneys — the expectation is that only about a quarter of them will receive a transplant within the next year. Greater public awareness of the organ shortage could lead to more donations, but doubts about the safety of donation by living persons and the integrity of the allocation process can jeopardize public trust. In January 2002, a donor died after giving a portion of his liver to his brother at Mount Sinai Medical Center in New York. Subsequently, the number of liver transplantations involving living donors in the United States — which had increased from 395 in 2000 to 518 in 2001 — decreased to 362 in 2002 and about 320 each in 2003 and 2004.

During the past year, a passionate controversy has developed about the public solicitation of organs, which takes place over the Internet, on billboards, and through other advertising. This is a complex matter, as discussed by Truog in this issue of the *Journal* (pages 444–446). Until recently, it was exceedingly rare for a kidney to be transplanted from a donor who was unknown to the recipient or whose blood or tissue was not compatible with the recipient's. Solicitation arouses concern about the potential for financial exploitation, the inequitable allocation of organs, and the subversion of the standards for donation. Federal regulations authorize direct-



Kidney Transplantations in the United States, 1988–2004.
Data are from the Organ Procurement and Transplantation Network.

ed donation, which is defined as "the allocation of an organ to a recipient named by those authorized to make the donation." Thus, soliciting organs from a living or dead donor is not unlawful, although the National Organ Transplant Act of 1984 prohibits the transfer "of any human organ for valuable consideration for use in a human transplantation if the transfer affects interstate commerce." Reasonable payments are permitted for the costs of organ procurement and storage and for the expenses incurred by living donors — in travel, housing, and lost wages.

In August 2004, a 32-year-old Houston man with liver cancer received a directed liver donation from a deceased donor after he advertised in newspapers, on a Web site (www.toddneedsaliver.com), and on two highway billboards, and his story had been widely publicized. He died in April 2005, eight months after the transplantation. His controversial campaign might have encouraged a donation that would not have occurred otherwise — or diverted a liver from someone with a greater chance for long-term survival.

Further controversy surrounds MatchingDonors.com. On its Web site, the organization, which is based in the Boston area, describes itself as "a venue where patients and potential donors can meet and communicate, and hopefully expedite a donor agreeing to give a patient a much needed organ." It states that the organization "is a nonprofit corporation and 100% of the money paid for patient memberships is applied to running this site." Potential recipients pay a one-time fee of $595 or a monthly fee of $295 to be listed, but about 70 percent of the 65 patients with active profiles are being listed without charge, according to Dr. Jeremiah Lowney, the organization's medical director and cofounder. There is no inherent reason why it should matter whether a living donor meets his or her intended recipient at work or a place of worship, online, or in some other way. But the involvement of a commercial entity has been criticized.

Since MatchingDonors.com was launched in February 2004, it has paired about 30 living organ donors and recipients. As of July 1, 2005, 12 transplantations had taken place, all of them of kidneys. There was wide media coverage of the first of these transplantations, which occurred at a Denver hospital in October 2004. The recipient, who had been seeking a kidney for five years, said that he had not paid for the kidney but had reimbursed the donor about $5,000 for expenses. The week after the transplantation, the donor was jailed in Tennessee for failure to pay child support. In March 2005, he took — and was said to have failed — a televised polygraph test, during which he was asked if he had profited from the transplantation.[1]

Although the public solicitation of organs has been a factor in very few transplantations, the transplantation of kidneys from living donors is increasingly common. The United Network for Organ Sharing, a private, nonprofit organization based in Richmond, Virginia, operates the Organ Procurement and Transplantation Network (OPTN) under contract with the federal government. The networks share the same board of directors and are collectively known as OPTN/UNOS. At a public hearing held in Chicago in June, Dr. Francis L. Delmonico, a Boston transplantation surgeon who is president of OPTN/UNOS, noted that "anyone can be a live kidney donor who is medically

and psychosocially suitable. It is no longer the case that one has to be HLA-matched. The outcome for a friend, spouse, or anonymous donor is just as good as that from a parent or child." Transplantation of a kidney from a living donor, however, remains a major surgical procedure with attendant risks and complications, as discussed by Ingelfinger in this issue of the *Journal* (pages 447–449). In addition, until recently, the presence of preformed antibodies against HLA antigens or incompatibility with respect to ABO blood type has generally ruled out intended donors.[2]

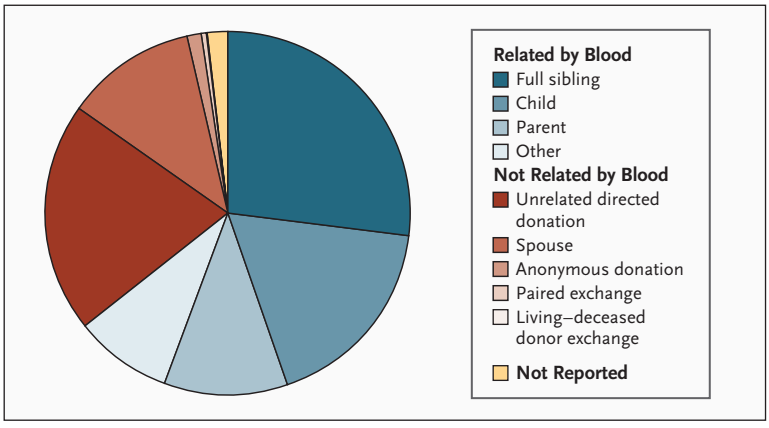
Since 2000, there have been more living kidney donors than cadaveric donors in the United States, although kidneys from cadavers — because both kidneys are usually available — still account for about 60 percent of kidney transplants (see line graph). Nearly two thirds of living donors are related to their recipients by blood (see pie chart). Those without a blood relationship are most commonly spouses or others who have a previous personal relationship with the recipient that is unrelated to the person's need for a transplant. In 2004, 85 people received a kidney through anonymous donation, in which a volunteer approached a transplantation center with no knowledge of a specific recipient. In the same year, 30 people received a kidney through paired exchange, in which two donors provided a kidney to each other's intended recipients because both had been found to be incompatible with the recipient to whom they had wished to donate.

The OPTN receives the majority of its funding from a one-time computer registration fee of $459, which is assessed when a transplantation center accepts a candidate for a transplant from a deceased donor or when a person who had not been on the waiting list receives a transplant from a living donor. Historically, consistent with its federal charge, the OPTN has focused on transplants from deceased donors. Donation by living persons was less common and usually involved people who knew each other. Such donation has primarily been the responsibility of individual transplantation centers.

Now, the OPTN is taking an active role in transplantation from living donors. In October 2004, the Department of Health and Human Services directed it to develop voluntary-allocation guidelines for organ donations from living donors that are made to an anonymous pool, not to specific patients, and other voluntary policies and guidelines as "it believes necessary and appropriate to promote the safety and efficacy of living donor transplantation for the donor and the recipient."[3] The responsibilities include the development of guidelines about the public solicitation of organ donors. In November 2004, OPTN/UNOS announced its opposition to the solicitation of organs from deceased donors who had no personal or family bond with the patient.[4] According to Delmonico, such solicitations "can undermine the allocation system and prevent the best medical use of the organ."

In June 2005, OPTN/UNOS announced that it "will not participate in efforts to solicit living donors for specific transplant candidates."[5] Specifically, it will not create a Web site similar to MatchingDonors.com. OPTN/UNOS will, however, "provide comprehensive resource information to support prospective live donors, including medical criteria for who can donate and individual transplant institutions' protocols for live unrelated donation."[5]

OPTN is establishing quality criteria and guidelines for living-donor kidney and liver programs. Programs must have staff



Relationships between Living Kidney Donors and Transplant Recipients in the United States, 2004.

Data are from the Organ Procurement and Transplantation Network.

with the requisite training and expertise to evaluate donors fully, to ensure that donations are voluntary, and to perform living-donor surgery. In addition, better data about the outcomes of living donors should be forthcoming. The OPTN is also working toward requiring that transplantation centers report within 72 hours the death of a living donor or the donor's need for organ transplantation. Finally, OPTN/UNOS will develop a nationwide mechanism for allocating organs from living donors who have not directed their donation to specific persons.[5] The principles will be similar to those for allocating organs from cadaveric donors. Eventually, there might be a registry for altruistic kidney and liver donors, such as the National Marrow Donor Program, which lists volunteer marrow donors.

The controversy over the public solicitation of organs has forced the transplantation community to address a difficult issue and to reexamine its approach to living donors. As long as there is a profound organ shortage, the challenge of trying to help patients in need of transplants while upholding the integrity of the overall allocation system will not go away.

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

1. Hartman T. Donor fails lie detector test on organ transplant. Rocky Mountain News. March 11, 2005:6A.
2. Delmonico FL. Exchanging kidneys — advances in living-donor transplantation. N Engl J Med 2004;350:1812-4.
3. Letter from Dr. James F. Burdick, director, Division of Transplantation, Healthcare Systems Bureau, Health Resources and Services Administration, Department of Health and Human Services, to Walter K. Graham, Executive Director, United Network for Organ Sharing, October 29, 2004.
4. OPTN/UNOS board opposes solicitation for deceased organ donation. Press release of the United Network for Organ Sharing, November 19, 2004. (Accessed July 14, 2005, at http://www.unos.org/news/newsDetail.asp?id=374.)
5. OPTN/UNOS board addresses information needs of potential living donors. Press release of the United Network for Organ Sharing, June 24, 2005. (Accessed July 14, 2005, at http://www.unos.org/news/newsDetail.asp?id=456.)

The Ethics of Organ Donation by Living Donors

Robert D. Truog, M.D.

Related articles, pages 441 and 447

Most organs for transplantation come from cadavers, but as these have failed to meet the growing need for organs, attention has turned to organs from living donors. Organ donation by living donors presents a unique ethical dilemma, in that physicians must risk the life of a healthy person to save or improve the life of a patient. Transplantation surgeons have therefore been cautious in tapping this source. As surgical techniques and outcomes have improved, however, this practice has slowly expanded. Today, according to the United Network for Organ Sharing (UNOS), almost half of all kidney donors in the United States are living. In 2004, living organ donors also provided a lobe of the liver in approximately 320 cases and a lobe of a lung in approximately 15 cases.

Three categories of donation by living persons can be distinguished: directed donation to a loved one or friend; nondirected donation, in which the donor gives an organ to the general pool to be transplanted into the recipient at the top of the waiting list; and directed donation to a stranger, whereby donors choose to give to a specific person with whom they have no prior emotional connection.

Each type of donation prompts distinct ethical concerns. With directed donation to loved ones or friends, worries arise about the intense pressure that can be put on people to donate, leading those who are reluctant to do so to feel coerced. In these cases, transplantation programs are typically willing to identify a plausible medical excuse, so that the person can bow out gracefully.[1] Equally important, however, are situations in which people feel compelled to donate regardless of the consequences to themselves. In one instance, both parents of a child who was dying of respiratory failure insisted on donating lobes of their lungs in a desperate but unsuccessful attempt to save her life.[2] Such a sense of compulsion

is not unusual. In cases like these, simply obtaining the informed consent of the relative is insufficient — physicians are obligated to prevent people from making potentially life-threatening sacrifices unless the chance of success is proportionately large.

Nondirected donation raises different ethical concerns. The radical altruism that motivates a person to make a potentially life-threatening sacrifice for a stranger calls for careful scrutiny. One recent case involved a man who seemed pathologically obsessed with giving away everything, from his money to his organs, saying that doing so was "as much a necessity as food, water, and air."[3] After donating one kidney to a stranger, he wondered how he might give away all his other organs in a dramatic suicide. Other psychologically suspect motivations need to be ruled out as well. Is the person trying to compensate for depression or low self-esteem, seeking media attention, or harboring hopes of becoming involved in the life of the recipient? Transplantation teams have an obligation to assess potential donors in all these dimensions and prohibit donations that arouse serious concern.[1]

Directed donation to a stranger raises similar ethical questions with a few additional wrinkles. This type of donation usually occurs when a patient advertises for an organ publicly, on television or billboards or over the Internet. Such advertising is not illegal, but it has been strongly discouraged by the transplantation community. Two central objections are that the practice is unfair and that it threatens the view that an organ is a "gift of life," not a commodity to be bought and sold.

Some argue that just as we have a right to donate to the political parties and charities of our choice, so should we be able to choose to whom to give our organs. In practice, however, this means that those who have the most compelling stories and the means to advertise their plight tend to be the ones who get the organs — rather than those most in need. This strikes some ethicists as unfair. Unlike monetary gifts, they argue, organ transplantation requires the involvement of social structures and institutions, such as transplantation teams and hospitals. Hence, the argument goes, these donations are legitimately subject to societal requirements of fairness, and transplantation centers should refuse to permit the allocation of organs on the basis of anything but morally relevant criteria.[4]

The most ethically problematic cases are those in which the recipient is chosen on the basis of race, religion, or ethnic group. In one case, for example, the family of a brain-dead Florida man agreed to donate his organs — but insisted that because of the man's racist beliefs, the recipients must be white. Although the organs were allocated accordingly, Florida subsequently passed a law prohibiting patients or families from placing such restrictions on donation.[5]

Even when the motives for choosing a recipient may be unethical, however, there might be reasons for allowing the donation to proceed. Consider a case that was discussed at a recent public forum hosted by Harvard Medical School's Division of Medical Ethics: a Jewish man in New York learned of a Jewish child in Los Angeles who needed a kidney transplant. The man wanted to help someone of his own faith and decided he was willing to donate a kidney to help this particular child. Despite his discriminatory preference, one might view the donation as permissible, since at least some patients would benefit (the child would receive a kidney, and those below her on the waiting list would move up one notch) and no one would be harmed (those above the girl on the waiting list would not receive the kidney under any circumstances, because the man would not give it to them). Whether directed donation to strangers violates standards of fairness is thus controversial. But if it is permitted, it will be very difficult to prohibit discriminatory preferences, since donors can simply specify that the organ must go to a particular person, without saying why.

The other substantial cause for concern about this type of donation is its potential for making possible the buying and selling of organs. These practices are strictly prohibited by law, yet they seem to be an inherent risk in directed donations to strangers. Wealthy patients in need and healthy donors looking for a quick fix to their financial problems will always be able to find ways around even the most earnest attempts to prevent money from changing hands.

Despite these concerns, efforts to direct organ donations to strangers are not new, dating back at least to the celebrated 1982 case of Jamie Fiske, whose

father successfully mounted a nationwide appeal for someone to donate a liver to her. Today, many such solicitations are transmitted over the Internet, where, when the practice was relatively limited, organ solicitation was managed quietly, on a case-by-case basis, by individual transplantation centers. All this changed, however, with the emergence in 2004 of MatchingDonors.com (as discussed by Steinbrook in this issue of the *Journal*, pages 441–444). This Web site currently claims to have more than 2100 registered potential donors and to have brokered 12 transplantations, with about 20 more recipient–donor pairs matched and awaiting surgery.

Although the business conducted on this organization's Web site does not raise any fundamental ethical issues not already posed by other methods of solicitation, it does introduce a new degree of visibility that increases the magnitude of the issue. Will competing commercial Web sites begin to emerge? How will these sites be held accountable? Dr. Jeremiah Lowney, the medical director of MatchingDonors.com, recently argued that just as a dating service could not be held responsible for a bad date, his Web site has no responsibility for the outcomes of its matches. Furthermore, the Web site has no mechanism for ensuring the quality of the information it provides about transplantation and donation by living persons or for checking the accuracy of information submitted by potential donors and recipients.

Given the life-or-death consequences of the procedure, organ donation should not be governed by the ethics of caveat emptor. Nevertheless, MatchingDonors.com has clearly identified a need, and if this need is not met by a service that can address the ethical challenges, the vacuum will be filled by other enterprises. Entrepreneurs commonly open up useful new markets and services that must eventually become subject to rigorous standards and regulations.

The solicitation of organs over the Internet is probably here to stay, but it will require higher standards of responsibility and accountability than are currently in place. UNOS has more than 20 years of experience in managing the cadaveric-donor pool and is in a good position to extend its jurisdiction to include donation by living donors. The organization recently considered the topic of solicitation and decided not to pursue building a Web site similar to that of MatchingDonors.com but, instead, to provide educational information for anyone who is willing to be a living donor of a kidney and to develop a nationwide mechanism for allocating organs for nondirected donation by living donors.

This effort, however, does not go far enough. The proposal does not address directed donation and leaves many critical aspects of donation by living donors to the transplantation centers. Organ transplantation is big business, and each center is highly motivated to expand its share of the pie. They therefore have intolerable conflicts of interest when it comes to regulating themselves. Instead, UNOS should be charged with standardizing the process for evaluating potential donors, ensuring that independent advocates are assigned to help donors make an informed choice, developing mechanisms to deal with potential injury or death to the donor, setting standards for both directed and nondirected donation, and prohibiting transplantation when the chance of success is insufficient to justify the risks. Comprehensive oversight is necessary if the ethical pitfalls are to be adequately addressed.

Dr. Truog is a professor of medical ethics and anesthesia (pediatrics) in the Department of Social Medicine at the Harvard Medical School and the Division of Critical Care Medicine at Children's Hospital Boston.

An interview with Dr. Truog can be heard at www.nejm.org.

1. Abecassis M, Adams M, Adams P, et al. Consensus statement on the live organ donor. JAMA 2000;284:2919-26.
2. Kolata G. Lungs from parents fail to save girl, 9, and doctors assess ethics. New York Times May 20, 1991:A11.
3. Parker I. The gift: Zell Kravinsky gave away millions: but somehow it wasn't enough. The New Yorker. August 2, 2004:54-63.
4. Kluge E-HW. Designated organ donation: private choice in social context. Hastings Cent Rep 1989;19(5):10-6.
5. Veatch RM. Transplantation ethics. Washington, D.C.: Georgetown University Press, 2000.

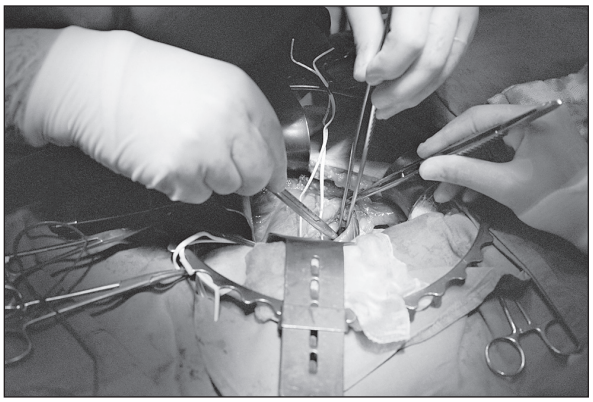
Risks and Benefits to the Living Donor

Julie R. Ingelfinger, M.D.

Related articles, pages 441 and 444

In the five decades since the first renal transplantation from a living donor took place, in 1954, donating a kidney has become common; according to the Organ Procurement and Transplantation Network, 6647 people became living kidney donors in the United States in 2004. Indeed, donating a kidney is sufficiently safe that the emotional benefits to the donor generally far outweigh the risks. During my career as a pediatric nephrologist, I have discussed kidney donation with scores of potential donors, helping the transplantation team to explain to the potential recipient and donor the procedure and its associated risks and benefits. In fact, I feel certain that I would be willing to donate a kidney, should a family member or intimate friend need one.

Advances in immunosuppression have changed the criteria for donation of a kidney by a living person, and someone who is neither a relative of the recipient nor a close HLA match can now donate. Increasing numbers of transplantations from unrelated donors are now being performed (raising many ethical issues that are addressed by Steinbrook and Truog in this issue of the *Journal*, pages 441–444 and 444–446, respectively), and recipients are increasingly likely to receive a transplant that functions well for many years. But what are the risks to the donor? Nowadays, the risks — both physical and emotional — associated with kidney donation are more completely understood than they were in the early years of transplantation; some of the risks were unanticipated.[1]



Antonia Reeve/Getty Images

Although donating may be an emotional necessity when the recipient is the donor's child, spouse, sibling, parent, or other loved one, this is far less likely to be the case for friends or anonymous donors. The current criteria for donor suitability now permit selected persons who have medical conditions such as treated hypertension to donate. Therefore, predonation evaluation assumes greater importance, since donors will undergo an operative procedure that is otherwise unnecessary yet can present medical hazards. The team evaluating a potential donor should not be the same as the team evaluating the recipient, given the obvious conflicts of interest.

Before kidney donation, a number of physiological studies must be performed to establish that the donor is in excellent health and has normal renal function and that the anatomy of the kidney is suitable for transplantation. In 1954, the kidney transplantation between the Herrick twins was performed without vascular imaging of the donated kidney, because at that time arteriography involved greater risk than surgery itself. Imaging of the renal vasculature has since evolved and is now standard procedure before kidney donation. It may identify unsuspected vascular disease and thereby benefit the potential donor, or it may uncover anatomy that renders transplantation technically difficult. Adverse events, however, are always possible whenever contrast material is used, although newer imaging methods, such as three-dimensional computed tomography and magnetic resonance angiography, mean that exposure to contrast medium is minimal for most donors.

In addition to assessing the appropriateness of the donor's kidney for the recipient, it is important to determine that the loss of a kidney would not pose a short- or long-term threat to the donor's general health. Thus, potential donors with conditions such as hypertension, vascular disease,

and the like are generally declined. And donors who are overweight or have certain lifestyle-related problems (e.g., excessive alcohol use) must address these issues successfully before they can donate a kidney. Persons with cancer or certain infectious diseases are excluded from donating, and other conditions, such as renal stones or previous urinary tract infections, may be relative contraindications.

Some conditions once thought to pose little risk to donors are now considered relative contraindications to donation. In X-linked hereditary nephritis, for example, renal failure generally develops in affected male family members, but females, who are carriers of the disease, usually have only microscopic hematuria and may wish to donate a kidney to a brother with renal failure. Although such a carrier will generally remain well when she has two kidneys, her risk of progression of the disease will increase if she loses or donates a kidney and is later subject to common life stresses such as pregnancy.

In addition to the team members who undertake the physical evaluation, transplantation teams generally include a psychiatrist who interviews potential donors to assess their motivation, who examines with them the psychosocial issues involved, and who supports them as they decide whether in fact to donate. The psychological implications of donation must be explored in depth, and there may be psychiatric reasons for declining someone's offer to donate.

What are the risks associated with donating? The short-term physical risks are generally small.[2,3] Although on rare occasions renal donors have died, the death rate is 0.03 percent — similar to or lower than that for any operation involving the use of general anesthesia. Other short-term risks are obvious: the risk of bleeding during or after the procedure and the risks of infection or other immediate problems related to the operation. In addition, donors will lose time from work. It can take a few weeks to recover fully from surgery, although the increasing use of laparoscopic donation means that a donor may now be discharged from the hospital in just a few days. Most return to work within four or five weeks.

Long-term health risks are less apparent. Most people do well with a single kidney. Donors may even live longer than nondonors, although this observation may simply reflect the careful selection of living donors from among very healthy candidates.

But long-term risks are not the same for every donor. Kidney function normally declines with age, and kidney donors have an age-related decline in renal function consistent with that observed in the general population. Renal failure has gradually developed in a small proportion of donors, and according to the United Network for Organ Sharing, 56 of more than 50,000 previous kidney donors have ultimately been listed for transplants themselves.[1] Although hypertension has developed in some donors over time, hypertension is so common in the developed world that ascribing it to kidney donation is probably not warranted. Occasionally, microalbuminuria develops in donors, and it is worth considering whether renoprotective therapy might be tailored to prevent this complication.

The decision to donate a kidney is inherently emotional. When all goes well, the donor almost always benefits. But what of the emotional stress that can occur when donation fails? A family friend donated a kidney to her brother some 20 years ago, and the kidney failed immediately because of unanticipated hyperacute rejection, for which no treatment was yet available. My friend was devastated, and it took many months for her emotional equanimity to be restored, although her physical recovery was rapid. Nevertheless, she remained glad throughout that she had donated, saying, "I did everything I could for my brother, and I would do it again."

What of donors in whom a physical illness subsequently develops that eventually results in progressive renal failure? More than 25 years ago, I had a teenage patient with end-stage renal disease who received a kidney from his mother, then in her early 40s. The recipient died a few years later. His mother had been deemed to be in perfect health and had no known risk factors for renal disease. Subsequently, she developed insulin-requiring diabetes, which became complicated by nephropathy, peripheral vascular disease, and peripheral neuropathy. She was close to requiring care for end-stage renal disease when she had a fatal myocardial infarction. Despite all her

health problems, my patient's mother had remained steadfast in her belief that donating her kidney, which had extended her son's life by a number of years, had been "more than worth it." And indeed most donors, when asked, claim that they would do it all again.[4]

But is it always worth it, physically and emotionally, to be a donor? And who should decide? Some transplantation programs inform potential donors, on their Web sites and in their informational material, that donation is relatively safe; others do not. Web sites such as Living Donors Online (www.livingdonorsonline.org/kidney/kidney.htm) and that of the National Kidney Foundation (www.kidney.org/recips/livingdonors/) provide information about donation by living persons and generally encourage people to consider it, while also posting reports of the experiences of donors, not all of which are entirely positive.

Fifty years ago, the first transplantation team suggested that organs from living donors should be used only when the likelihood of success for the recipient was high, the risk to the donor was low, and true voluntary consent was obtained from all involved; the Live Organ Donor Consensus Group has largely supported this viewpoint.[5] This triple principle is still the standard of care and should remain so. The changes in the chances of success, however, have altered the landscape. It is our responsibility as physicians to ensure that, in an era with shifting criteria, we do not stretch the ethics involved in assessing the risks and benefits to the donor.

1. Delmonico F. A report of the Amsterdam Forum on the Care of the Live Kidney Donor: data and medical guidelines. Transplantation 2005;79:Suppl 6:S53-S66.
2. Fehrman-Ekholm I, Duner F, Brink B, Tyden G, Elinder CG. No evidence of accelerated loss of kidney function in living kidney donors: results from a cross-sectional follow-up. Transplantation 2001;72:444-9.
3. Matas AJ, Bartlett ST, Leichtman AB, Delmonico FL. Morbidity and mortality after living kidney donation, 1999-2001: survey of United States transplant centers. Am J Transplant 2003;3:830-4.
4. Fehrman-Ekholm I, Brink B, Ericsson C, Elinder CG, Duner F, Lundgren G. Kidney donors don't regret: follow-up of 370 donors in Stockholm since 1964. Transplantation 2000;69:2067-71.
5. Abecassis M, Adams M, Adams P, et al. Consensus statement on the live organ donor. JAMA 2000;284:2919-26.

THIS WEEK *in the* JOURNAL

ORIGINAL ARTICLE

Screening Blood Donations for West Nile Virus

In 2003 and 2004, routine testing of blood donations for West Nile virus RNA in the United States led to the identification of 540 positive donations, of which 67 percent were IgM-negative and most likely to be infectious. The rates of positive donations decreased from 1.49 per 10,000 in 2003 to 0.44 per 10,000 in 2004.

The rapid implementation of a nucleic acid amplification test to screen for West Nile virus has made the blood supply safer. There were no confirmed cases of West Nile virus infection among recipients of tested blood.

SEE P. 451; EDITORIAL, P. 516; CME, P. 539

ORIGINAL ARTICLE

Testing for West Nile Virus

In 2003, nucleic acid amplification screening of 677,603 blood donations for West Nile virus with the use of "minipools" of 16 samples led to the identification of 1 positive donation for every 3703 analyzed (0.027 percent). In 2004, a strategy of testing individual donations in selected regions led to a 32 percent increase in the identification of donations with West Nile virus.

This study shows that screening of pooled blood samples for West Nile virus prevented hundreds of infections but missed donations with a low level of viremia.

SEE P. 460; EDITORIAL, P. 516

ORIGINAL ARTICLE

An Exercise Nomogram for Women

Exercise capacity is a predictor of the risk of death among both women and men. Whereas there are extensive data on expected functional capacity for age among men, normative values for women have not been well established. In a study of 5721 asymptomatic women who underwent symptom-limited exercise testing, a nomogram was developed to compare a woman's exercise performance with levels predicted for age.

SEE P. 468; EDITORIAL, P. 517

ORIGINAL ARTICLE

Modafinil for Excessive Sleepiness Associated with Shift-Work Sleep Disorder

About 1 of 10 night-shift workers suffers from severe excessive sleepiness on the job. This realization has led to the development of specific diagnostic criteria for people affected with shift-work sleep disorder. In this multicenter study, patients meeting this case definition were treated with placebo or modafinil. Although there was a significant improvement in laboratory measures of sleep, treated patients were still quite sleepy.

Modafinil improves performance among patients with shift-work sleep disorder but does not come close to returning them to a normal level of sleepiness.

SEE P. 476; EDITORIAL, P. 519

DRUG THERAPY

Adherence to Medication

The full benefit of many effective medications will be achieved only if patients adhere to prescribed treatment regimens. Unfortunately, applying terms such as "noncompliant" and "nonadherent" to patients who do not consume every pill at the desired time can stigmatize them in their future relationships with health care providers. This article on medication adherence (or compliance) reviews strategies to assess and enhance this important aspect of patient care.

SEE P. 487; CME, P. 538

CURRENT CONCEPTS

Diagnosis from the Blood Smear

Even in the age of automated analysis, the examination of the blood smear remains an important diagnostic tool. An expert examination of the blood smear can identify errors, establish a diagnosis, or lead to a useful fortuitous finding. An atlas of instructive blood smears is included and is available at www.nejm.org as a set of slides.

SEE P. 498; CME, P. 537

CLINICAL PROBLEM-SOLVING

A Fractured Diagnosis

A 44-year-old woman came to the emergency department because of pain in her right thigh shortly after she had a minor fall. A right femoral-neck fracture was diagnosed, and she was admitted to the orthopedic ward to await surgery. Six months before hospitalization, limb pain had developed, which became progressively worse. The patient also reported a weight loss of 30 kg and fatigue.

SEE P. 509

CLINICAL IMPLICATIONS OF BASIC RESEARCH

An Enzyme Critical to Osteoarthritis

The aggrecanases — enzymes that break down cartilage — are attractive candidates as drug targets. Studies of two mouse models implicate a single aggrecanase as a mediator of the erosion of cartilage in osteoarthritis.

SEE P. 522

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 AUGUST 4, 2005 VOL. 353 NO. 5

West Nile Virus among Blood Donors in the United States, 2003 and 2004

Susan L. Stramer, Ph.D., Chyang T. Fang, Ph.D., Gregory A. Foster, B.S., Annette G. Wagner, M.S., Jaye P. Brodsky, B.S., and Roger Y. Dodd, Ph.D.

ABSTRACT

BACKGROUND

West Nile virus first appeared in the United States in 1999 and has since spread throughout the contiguous states, resulting in thousands of cases of disease. By 2002, it was clear that the virus could be transmitted by blood transfusion, and by the middle of 2003, essentially all blood donations were being tested for West Nile virus RNA with the use of investigational nucleic acid amplification tests; testing was performed on individual samples or on "minipools" of up to 16 donations.

METHODS

We analyzed data from the West Nile virus testing program of the American Red Cross for 2003 and 2004 to identify geographic and temporal trends. In areas with a high incidence of infection, individual donations were tested to increase the sensitivity of testing. Donors with reactive results participated in follow-up studies to confirm the original reactivity and to assess the natural history of infection.

RESULTS

Routine testing in 2003 and 2004 identified 540 donations that were positive for West Nile virus RNA, of which 362 (67 percent) were IgM-antibody–negative and most likely infectious. Of the 540 positive donations, 148 (27 percent) were detectable only by testing of individual donations, but only 15 of the 148 (10 percent) were negative for IgM antibody. The overall frequencies of RNA-positive donations during the epidemic periods were 1.49 per 10,000 donations in 2003 and 0.44 per 10,000 in 2004. In 2004, 52 percent of the positive donations were from donors in four counties in southern California.

CONCLUSIONS

Rapid implementation of a nucleic acid amplification test led to the prospective identification of 519 donors who were positive for West Nile virus RNA and the removal of more than 1000 potentially infectious related components from the blood supply of the Red Cross. No cases of transfusion-transmitted infection were confirmed among recipients of the tested blood.

From the American Red Cross, Gaithersburg, Md. (S.L.S., G.A.F., A.G.W.), and Rockville, Md. (C.T.F., R.Y.D.), and Quality Analytics, Riverwoods, Ill. (J.P.B.). Address reprint requests to Dr. Stramer at the American Red Cross, 9315 Gaither Rd., Gaithersburg, MD 20877, or at stramers@usa.redcross.org.

N Engl J Med 2005;353:451-9.

Although West Nile virus was first isolated in 1937 from a patient in Uganda,[1] it was not seen in the Western Hemisphere until 1999, when 62 cases of West Nile virus encephalitis were reported.[2,3] Biggerstaff and Petersen estimated that, during the peak of the 1999 outbreak in Queens, New York, the maximal and mean risks of transmission of West Nile virus by blood transfusion were 2.7 and 1.8 per 10,000 units, respectively.[4] In September 2002, the Centers for Disease Control and Prevention (CDC) announced that three of four recipients of transplanted organs from a single donor had acquired meningoencephalitis and were positive for West Nile virus. The fourth recipient was subsequently confirmed to have West Nile virus infection. The organ donor acquired West Nile virus infection through the transfusion of blood from 63 donors two days before organ harvest.[5,6] Subsequently, 23 cases of transfusion-transmitted West Nile virus infection were confirmed in 2002.[7] As a result, on September 20, 2002, major blood organizations, diagnostic companies, the CDC, the American Association of Blood Banks, and the Food and Drug Administration (FDA) agreed that blood-screening tests for West Nile virus RNA were needed by the 2003 season of West Nile virus infection. The FDA provided guidance for the assessment of donor suitability and the safety of blood and blood products in cases of known or suspected West Nile virus infection in October 2002,[8] with revisions in May 2003[9] and April 2005.[10]

In the United States, routine screening of blood donors for West Nile virus RNA started in the summer (June through August) of 2003. We describe the results of the American Red Cross program of laboratory testing in 2003 and 2004 and compare the yield in those years with the preclinical yield obtained in 2002.[11] The dynamics of viral replication, the serologic profiles of seroconverting donors, and the associated clinical symptoms in infected donors are reported in detail elsewhere.[12-14]

METHODS

SELECTION AND QUALIFICATION OF TEST KITS

In 2002, a total of 383 retrieved units of frozen plasma, including plasma corresponding to blood components transfused to 11 of the 23 patients with confirmed transfusion-transmitted infection,[7] were tested with the use of five different investigational and research-based assays for West Nile virus RNA and three research-based or FDA-cleared assays for viral IgM antibody.[15] On the basis of these studies, the Gen-Probe Procleix West Nile virus assay (Gen-Probe and Chiron) was deemed qualified for routine screening of blood donations involving "minipools" of 16 samples. In this test, West Nile virus–specific RNA is amplified by transcription-mediated amplification. Also on the basis of these studies, the qualitative and quantitative polymerase-chain-reaction (PCR) assays (National Genetics Institute) and IgM antibody test (Abbott Laboratories) were selected for use in confirmation and follow-up studies.

SAMPLE COLLECTION AND LABORATORY TESTING

Plasma samples to be screened for West Nile virus RNA were obtained from the collected Plasma-Preparation Tubes (Becton Dickinson) used for routine screening for human immunodeficiency virus type 1 (HIV-1) RNA and hepatitis C virus RNA.[16] Testing for West Nile virus was performed under an FDA-approved Investigational New Drug application. All studies were approved by the institutional review board of the American Red Cross.

Routine screening for West Nile virus RNA of minipools of plasma from 16 donations was implemented on June 23, 2003, and is ongoing at five American Red Cross laboratories. The testing process is identical to that used for routine screening of blood for HIV-1 and hepatitis C virus RNA.[16] Samples from reactive minipools were tested individually to determine which were reactive. Routine nucleic acid amplification testing of individual donations was substituted for minipool testing in areas in which reactive donations exceeded the defined threshold.[17] In 2003, nucleic acid amplification testing of individual donations was also retrospectively performed on samples collected in Nebraska that had been nonreactive on minipool testing to determine whether any reactive donations had gone undetected with the use of minipool testing.

All reactive samples and samples of the corresponding plasma components manufactured from the index donations were further tested for West Nile virus RNA by transcription-mediated amplification and PCR. Samples from the retrieved plasma components from donors with reactive samples on nucleic acid amplification testing in 2003 were tested for IgM antibodies to the virus. Manufacture of the Abbott IgM assay was discontinued in 2004; the use of another test was necessary. Thus, beginning

in 2004, FDA-cleared assays for West Nile virus IgM and IgG antibodies (Focus Diagnostics) were used.[18] Use of this IgM assay required multiplication of the cutoff value by a correction factor of 0.67 and coupling of the assay with the company's IgG assay so that sensitivity was similar to that of the Abbott IgM assay. Donations that were positive for West Nile virus RNA on minipool or individual nucleic acid amplification testing were considered to be confirmed if the index donation sample, retrieved plasma-component sample, or donor follow-up samples were reactive on repeated nucleic acid amplification testing (transcription-mediated amplification or PCR), were positive for West Nile virus–specific antibodies, or met both criteria. The viral loads (expressed as the number of copies of West Nile virus RNA per milliliter) of PCR-positive index or follow-up samples were determined. According to the manufacturers, the 50 percent detection rate of transcription-mediated amplification is 3 to 4 copies per milliliter, and the sensitivity of the qualitative PCR is 5 copies per milliliter; the sensitivity of quantitative PCR is 100 copies per milliliter.

APPROACH TO BLOOD DONORS, COMPONENTS, AND RECIPIENTS

Donors with either confirmed positive or false positive results on nucleic acid amplification tests for West Nile virus RNA were notified, and they were prevented from making further donations according to FDA guidelines.[8,9] Demographic information about these donors (including ZIP Code of residence, sex, and age and whether they were first-time or repeat donors) was collected for analysis. Donors with reactive specimens who provided written informed consent participated in the follow-up study by providing additional blood samples for repeated RNA and antibody testing.

On identification of an RNA-reactive donation, all components associated with the index donation were quarantined and the plasma unit was retrieved for further testing. We traced recipients of transfused components from confirmed positive index donations identified through retrospective nucleic acid amplification testing of individual donations in 2003.

STUDY DESIGN AND ANALYSIS

The authors are jointly responsible for the study design, the integrity and analysis of the data, and the content of the article. Data on nucleic acid amplification testing were collected and verified by an independent clinical-research organization (Medical Marketing Consultants).

RESULTS

PREVALENCE OF WEST NILE VIRUS AMONG BLOOD DONORS, 2002 THROUGH 2004

Studies conducted on samples collected in September 2002 from six areas with a high incidence of West Nile virus infection yielded a confirmed positive rate of 0.095 percent, or 1 in 1057 samples.[11] This rate, even late in the 2002 season, remained high and similar to the average risks of transfusion-transmitted West Nile virus infection estimated by the CDC for 2002 for the same metropolitan areas.[19] Table 1 shows the prevalence rates of West Nile virus infection in 2002 in comparison with the rates in 2003 and 2004.

Table 1. Results of Tests of Blood Donations for West Nile Virus.*

Individual NAT	Minipool NAT	IgM or IgG Antibody	Sept. 3–28, 2002 (N=48,620)†	June 29–Dec. 1, 2003 (N=2,935,249)	June 16–Oct. 16, 2004 (N=2,386,630)
				no. of confirmed positive samples	
+	–	–	0	10	5
+	+	–	16	283	64
+	+	+	0	36	9
+	–	+	30	107	26
	Total		46	436	104
				rate (no./10,000 donations)	
			9.46	1.49	0.44

* NAT denotes nucleic acid amplification test. Plus signs denote a positive result, and minus signs a negative result.
† Retrospective nucleic acid amplification testing was performed on frozen samples of individual donations from six high-incidence regions.[11]

In 2003, the first confirmed positive donation was identified on June 26 in Los Angeles from a donor who had returned from a trip to Colorado on the day before donation. The last positive donation was identified on December 1 in Georgia. Overall, 436 confirmed positive donations were identified from a total of 2,935,249 donations screened, for a rate of 0.015 percent, or 1 in 6732 (Table 1). Of these positive donations, 328 (75 percent) were collected from Kansas and Nebraska residents, for a combined rate of 0.68 percent, or 1 in 147, which was 45 times as high as the systemwide rate.

The first confirmed positive donor in 2004 was identified on June 16 in Phoenix, Arizona, and the last was identified on October 16 in Los Angeles. During this period, 104 confirmed positive donations were identified from a total of 2,386,630 screened, for a rate of 0.004 percent, or 1 in 22,948 (Table 1). Of these positive donations, 54 (52 percent) were identified from residents of four southern California counties (Los Angeles, Orange, Riverside, and San Bernardino), for a rate of 0.064 percent, or 1 in 1566, which was 16 times as high as the systemwide rate.

Figure 1 shows the frequency of confirmed positive donors identified in 2002, 2003, and 2004, according to week of donation. Although all positive donors in both 2003 and 2004 were identified between June and December, the peak season for 2003 was from mid-August to mid-September, whereas for 2004, the peak season started in late July and continued through late September, but at a lower frequency. The absence of any West Nile virus RNA–positive donations during the week before the initiation of prospective screening in 2003 (from retrospective testing of samples retained from the prior week that had been frozen) suggests that the onset of routine screening preceded the 2003 epidemic.

Our data indicate that the areas of highest incidence of confirmed positive donors moved westward from 2002 to 2004. The Cleveland and Detroit metropolitan areas had higher rates in September 2002 than at any time during 2003 and 2004. Kansas and Nebraska had higher rates in 2003 than in 2004, and southern California had the highest rates in 2004. These findings are in agreement with the pattern of clinical cases reported to the CDC.[20]

DEVELOPMENT OF THE TRIGGER FOR NUCLEIC ACID AMPLIFICATION TESTING OF INDIVIDUAL DONATIONS

On the basis of data obtained in 2002, we found that there was one potentially infectious (RNA-pos-

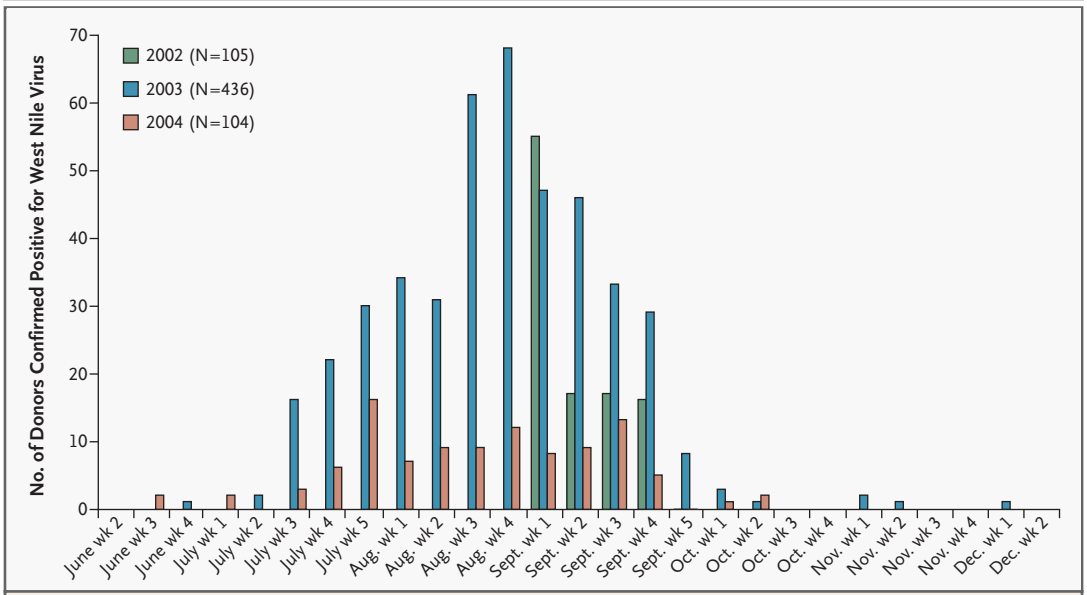


Figure 1. Numbers of U.S. Blood Donors Who Were Confirmed to Be Positive for West Nile Virus RNA in 2002, 2003, and 2004, According to the Week of Collection.

Data on 2002 totals are from Stramer et al.[11] The weekly totals for 2002 were adjusted for the number of donations tested during the four-week study.

itive, IgM-negative) sample detectable only by nucleic acid amplification testing of individual donations for every four such samples detected by minipool testing.[11] For 2003 and 2004, we therefore chose to initiate testing of individual donations in any blood-collection region after the identification of four RNA-positive donations and consequent calculation of a detection frequency of 1 in 1000 (the epidemic frequency documented in 2002[11,19]), on the basis of the date of collection of the first reactive donation. Once nucleic acid amplification testing of individual donations had been initiated, a seven-day period with no RNA-reactive donations was required before a collection region could revert to the use of minipool testing.

NUCLEIC ACID AMPLIFICATION TESTING AND IgM AND IgG TESTING OF INDIVIDUAL DONATIONS

In 2003, 30,501 Red Cross donations from Kansas residents (August 19 to September 27) and Nebraska residents (August 25 to October 4) underwent individual nucleic acid amplification testing. Prospective testing identified 181 confirmed positive donations. Of these, 96 (53 percent) were nonreactive at a 1:16 dilution, of which 88 (92 percent) were IgM-positive and 8 (8 percent) were IgM-negative. In addition, as requested by the FDA, individual nucleic acid amplification testing was performed retrospectively on frozen samples from 18,049 donations collected from July 10 (the date of the first confirmed positive donation identified by minipool testing) to August 22 from donors who lived in Nebraska to determine whether donations that were nonreactive on minipool testing and had therefore been released for transfusion would be identified as reactive on testing of individual donations. This retrospective evaluation identified 21 additional confirmed positive donations: 19 were IgM-positive samples and 2 were IgM-negative samples. During the same period, minipool testing had previously identified 80 confirmed positive donations (or 79 percent of the total detected during this period): 7 were IgM-positive samples and 73 were IgM-negative samples.

Overall, 117 of the 436 confirmed positive donations identified in 2003 (27 percent) were detected only by individual nucleic acid amplification testing (although this may be an underestimate, since not all donations were tested by this method), and of these 117, 10 (9 percent) were IgM-negative (Table 1). Of the remaining 319 donations that were reactive on minipool testing, 283 (89 percent) were IgM-negative (Table 1). For the 143 IgM-positive donations, the median viral load was below 100 copies per milliliter (range, less than 5 to 14,000), as compared with 5800 copies per milliliter (range, less than 5 to 580,000) for the 293 IgM-negative donations ($P<0.001$ by the Wilcoxon rank-sum test). Donations that were positive for West Nile virus RNA and IgM were identified more frequently as the season progressed (Fig. 2A).

In 2004, the trigger for nucleic acid amplification testing of individual donations was reached for collections tested by the Red Cross in three areas: southern California (July 25 to October 8), Kansas (September 12 to 27), and Arkansas (August 28 to September 6), for a total of 92,460 donations tested individually. No reactive donations were identified by individual testing of Arkansas donors. However, 48 of 54 confirmed positive donations from southern California residents (89 percent) and 3 of 7 confirmed positive donations from Kansas residents (43 percent) were identified during the period of individual testing, for a combined positive rate of 0.056 percent, or 1 in 1791 donations. Of the 51 confirmed positive donations identified by nucleic acid amplification testing of individual donations, 31 (61 percent) were nonreactive on minipool testing. These 31 donations represented 30 percent of the total 104 positive donations identified in 2004; 26 (84 percent) were positive for IgM or IgG, and 5 (16 percent) were negative for IgM and IgG (Table 1). Figure 2B shows that the identification of confirmed positive donations with IgM or IgG antibody reactivity increased as the 2004 season progressed; this increase was less than that observed for 2003. For the 35 IgM- or IgG-positive donations, the median viral load was below 100 copies per milliliter (range, less than 5 to 47,000), as compared with 3200 copies per milliliter (range, less than 5 to 160,000) for the 69 IgM- and IgG-negative donations ($P<0.001$ by the Wilcoxon rank-sum test).

In both 2003 and 2004, two thirds of all viremic donors were negative for West Nile virus antibody, according to two different IgM-antibody testing strategies: the Abbott Laboratories IgM assay (Fig. 2A) and the Focus Technologies IgM and IgG assays (Fig. 2B).

DEMOGRAPHIC CHARACTERISTICS OF DONORS

Among the confirmed positive donors, as compared with the group of donors with false positive

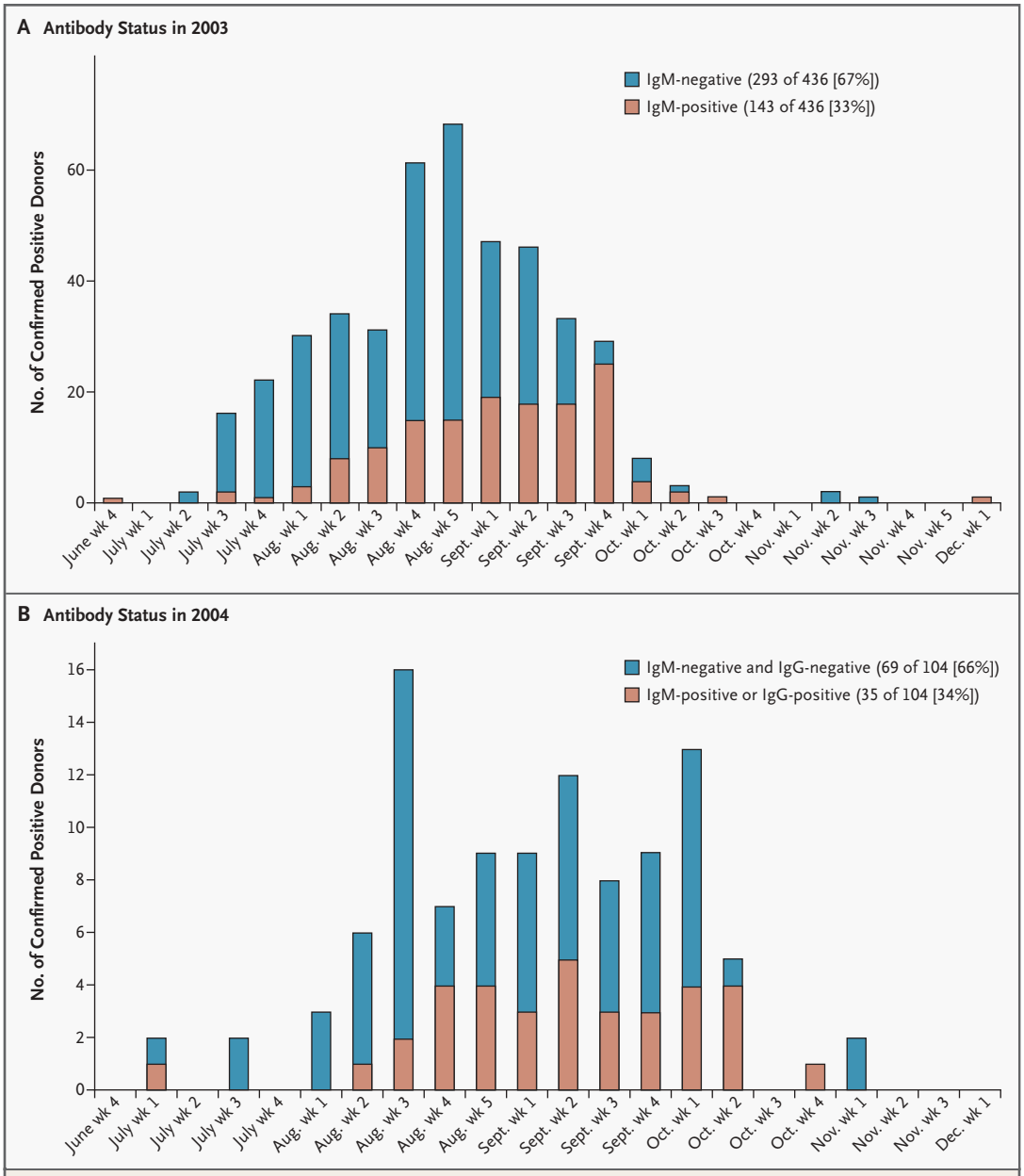


Figure 2. IgM and IgG Antibody Status of Blood Donors Who Were Confirmed to Be Positive for West Nile Virus RNA on Prospective Screening in 2003 and 2004, According to the Week of Collection.

In 2003 (Panel A), IgM antibody testing was performed by Abbott Laboratories; in 2004 (Panel B), IgM and IgG antibody testing was performed by Focus Technologies with the use of a reduced cutoff value (the standard cutoff value was multiplied by a correction factor of 0.67).

results, there were more male than female donors in 2002,[11] 2003, and 2004 (Table 2). In 2003, 50 percent of the overall donor population of the American Red Cross was male. Combining the confirmed positive donors from 2002 through 2004, 13 percent were first-time donors, with a mean age of 46 years (range, 16 to 83). These observations did not differ significantly from those observed for the donors with false positive test results.

FOLLOW-UP OF DONORS AND RECIPIENTS

In 2003, 350 of 415 confirmed positive donors identified prospectively participated in the follow-up study, of whom 335 (96 percent) were IgM-positive

Table 2. Demographic Characteristics of Blood Donors with Reactive Nucleic Acid Amplification Tests.

Period	No. of Donors	Male Donors	1st-Time Donors	Age	
				Mean	Range
		no. (%)		*yr*	
Sept. 3–28, 2002*					
Confirmed positive	46	24 (52)	9 (20)	43.8	17–77
False positive	52	23 (44)	10 (19)	42.7	17–73
June 29–Dec. 1, 2003					
Confirmed positive	436	258 (59)†	43 (10)†	46.5	17– 83
False positive	382	187 (49)	58 (15)	47.5	13–89‡
June 16–Oct. 16, 2004					
Confirmed positive	104	62 (60)	24 (23)	44.9	16–83
False positive	73	37 (51)	12 (16)	41.8	17–75

* Retrospective nucleic acid amplification testing was performed on frozen samples of individual donations from six high-incidence regions.[11]
† $P<0.05$ for the comparison with the false positive group during the same period.
‡ Frozen plasma was obtained from a 13-year-old repeat autologous donor.

or seroconverted during follow-up. Of 186 donors who participated in long-term follow-up, 166 (89 percent) retained specific IgM reactivity for 100 days or longer. However, of the 17 of 46 confirmed positive donors identified in 2002 for whom follow-up data were available,[11] 10 (59 percent) had IgM reactivity for more than 398 days, consistent with the observations of Roehrig and coworkers.[21] Of 104 confirmed positive donors identified in 2004, 82 (79 percent) participated in follow-up studies.

Only two recipients of donations confirmed to be positive on retrospective testing of individual donations in 2003 could be identified and consented to follow-up studies. Both were seronegative for West Nile virus and had had no reported symptoms associated with West Nile virus infection during the year after transfusion. In the case of both recipients, the transfused component was IgM-positive and had an RNA level that was too low to quantitate. In contrast, a 2002 recipient of an IgM-negative unit with a viral load of 6300 copies per milliliter had West Nile virus–related symptoms and antibody seroconversion.[11] Although these numbers are small, the data are consistent with reports of seroconversion and disease related to West Nile virus infection only among recipients of viremic, IgM-negative blood components.[7,22-25]

DISCUSSION

In 2002, transfusion-transmitted West Nile virus infection was confirmed in 23 recipients of blood components, with the true number of transmissions believed to be much higher. By early summer of 2003, blood-collection agencies had implemented blood-donor testing for West Nile virus RNA, identifying and reporting a total of 1041 RNA-positive donations to the CDC through Arbonet for the 2003 and 2004 seasons.[19,26] The Red Cross program identified 540 of these viremic donations, of which 362 (67 percent) were negative for West Nile virus antibodies and likely infectious. The major epidemic focus in 2003 was the upper Plains states, moving to the Southwest in 2004. Most important, on the basis of prospective screening for West Nile virus RNA performed in 2003 and 2004 in our program, 1023 components manufactured from 519 prospectively screened viremic donations were not released for use and therefore not transfused.

Screening of blood donations for West Nile virus RNA was initiated in minipools of 16 samples, leading to concern that donations with low-level viremia might escape detection.[11,17,23] Our data from the 2002 West Nile virus season suggested that, at the height of the epidemic, there might be one viremic donation undetectable by minipool testing for every four that were detected. Accordingly, in areas with a high prevalence of RNA-positive donations (i.e., more than 1 in 1000 samples), we initiated nucleic acid amplification testing of individual donations after identifying a total of four RNA-positive donations on minipool testing in any given blood-collection region. This policy was supported by the observation of West Nile virus infec-

tions associated with the transfusion of blood units with RNA levels that could not be detected by minipool testing.[22-25] In addition, the effectiveness of this evidence-based trigger is demonstrated by the absence of any confirmed cases of transfusion-transmitted West Nile virus infection associated with blood components from our system in 2003 and 2004. An assessment of the effectiveness of various trigger strategies has been published elsewhere.[27] The continued use of a trigger strategy for the screening of blood donations for West Nile virus appears justified in order to focus available resources at times and locations of peak incidence. In contrast, at times and locations in which there are few or no identified viremic donations, minipool screening provides adequate safety.

Through careful follow-up studies of all RNA-reactive donors, we were able to establish the natural history of West Nile virus infection, finding that IgM antibodies against the virus were detectable about 11 days after the detection of viral RNA on minipool testing (13 days after the detection of viral RNA on testing of individual samples), followed rapidly by the appearance of IgG antibodies.[12,13] The transmission of West Nile virus through transfusion has not been linked to an RNA-positive component that is also positive for IgM or IgG antibodies against the virus. Although we identified 148 viremic donations that were detectable only by nucleic acid amplification testing of individual donations, only 15 of them (10 percent), or 1 in 9400 samples, were IgM-negative and thus represent the earliest stages of donor infection. Therefore, in programs dependent on trigger strategies, careful, real-time monitoring is critical; in the absence of timely system readiness and monitoring, breakthrough infection has been documented.[24]

The vast majority of the yield of nucleic acid amplification testing of individual donations was IgM-positive, demonstrating the long duration of positivity for IgM antibody in the presence of low-level viremia. It is likely that such donations would be noninfectious, especially in the presence of high titers of IgM and IgG, but studies to confirm this possibility have not been performed. Studies tracing recipients of blood components generally have a low yield, and our study is no exception; however, two recipients of IgM-positive, viremic blood components had no evidence of West Nile virus infection, in contrast to recipients who received IgM-negative, viremic units.[7,11,22-25]

In future years, will the need to screen blood donors for West Nile virus continue? We have now seen three West Nile virus epidemic seasons in the United States, with an expanding geographic range of viremic blood donors and clinical infections. The number of reported cases of West Nile virus neuroinvasive disease peaked in 2002 and 2003 (2946 and 2866 cases, respectively[20]) and decreased (to 1108 cases) in 2004.[26] This same trend was observed by our identification of 436 viremic blood donors in 2003 and of 104 in 2004, as well as by the total number of viremic blood donors reported to the CDC during these years (818 and 223, respectively).[20,26] West Nile virus infection may eventually follow the same pattern as St. Louis encephalitis, with only infrequent, localized recurrences.

To our knowledge, the implementation of nationwide testing for West Nile virus RNA has been the most rapidly instituted test-based intervention in the history of transfusion safety, taking only about nine months from the decision to develop a test to its implementation. The program is a model of cooperation among public health agencies, regulators, manufacturers, and the blood-supply system.[28] It is to be hoped that this process will be effectively replicated should there be a similar outbreak in the future.

Dr. Stramer reports holding stock in Abbott Laboratories and having received consulting fees from Gen-Probe, Chiron, and Abbott Laboratories. Dr. Fang reports holding stock in and having received consulting fees from Chiron. Ms. Brodsky reports holding stock in Abbott Laboratories. Dr. Dodd reports having received consulting fees from Gen-Probe, Chiron, and Abbott Laboratories.

We are indebted to staff members at all participating blood-collection regions and the American Red Cross National Testing Laboratories, especially those in the Midwest, Central Plains, and Southern California regions and the St. Louis National Testing Laboratory; to J. Davis and staff members at Medical Marketing Consultants; to R. Lanciotti and members of the technical staff at the CDC laboratories in Fort Collins, Colo.; to J. Linnen and members of the technical staff at Gen-Probe and Chiron; to R. Smith and members of the technical staff at the National Genetics Institute; to H. Prince and members of the technical staff at Focus Technologies; to G. Dawson and members of the technical staff at Abbott Laboratories; and to A. Fisher for assistance in the assembly and submission of the manuscript.

REFERENCES

1. Smithburn KC, Hughes TP, Burke AW, Paul JH. A neurotropic virus isolated from the blood of a native of Uganda. Am J Trop Med 1940;20:471-92.
2. Outbreak of West Nile viral encephalitis — New York, 1999. MMWR Morb Mortal Wkly Rep 1999;48:845-9.
3. Mostashari F, Bunning ML, Kitsutani PT, et al. Epidemic West Nile encephalitis, New York, 1999: results of a household-based seroepidemiological survey. Lancet 2001;358:261-4.
4. Biggerstaff BJ, Petersen LR. Estimated risk of West Nile virus transmission through blood transfusion during an epidemic in

Queens, New York City. Transfusion 2002; 42:1019-26.
5. Update: investigations of West Nile virus infection in recipients of organ transplantation and blood transfusion. MMWR Morb Mortal Wkly Rep 2002;51:833-6.
6. Iwamoto M, Jernigan DB, Guasch A, et al. Transmission of West Nile virus from an organ donor to four transplant recipients. N Engl J Med 2003;348:2196-203.
7. Pealer LN, Marfin AA, Petersen LR, et al. Transmission of West Nile virus through blood transfusion in the United States in 2002. N Engl J Med 2003;349:1236-45.
8. Department of Health and Human Services. Guidance for industry: recommendations for the assessment of donor suitability and blood and blood product safety in cases of known or suspected West Nile virus infection. Rockville, Md.: Food and Drug Administration, October 2002.
9. *Idem*. Guidance for industry: revised recommendations for the assessment of donor suitability and blood and blood product safety in cases of known or suspected West Nile virus infection. Rockville, Md.: Food and Drug Administration, May 2003.
10. *Idem*. Guidance for industry: revised recommendations for the assessment of donor suitability and blood and blood product safety in cases of known or suspected West Nile virus infection. Rockville, Md.: Food and Drug Administration, April 2005.
11. Stramer SL, Beyers ML, Wagner AG, et al. West Nile virus (WNV) RNA prevalence in blood donations during the 2002 US epidemic. Transfusion 2003;43:Suppl:3A. abstract.
12. Stramer SL, Brodsky JP, Dodd RY, et al. Dynamics of WNV infection and resolution among blood donors. Presented at the Sixth National Conference on West Nile Virus in the United States, San Jose, Calif., February 8–9, 2005.
13. Busch MP, Stramer SL. Viremia and antibody studies in WNV infected blood donors. (Accessed July 8, 2005, at www.cdc.gov/ncidod/dvbid/westnile/conf/February_2005.htm.)
14. Orton S, Stramer SL, Dodd RY. Self-reported symptoms associated with West Nile virus infection in RNA-positive blood donors. Transfusion (in press).
15. Stramer SL, Wagner AG, Paolillo JE, et al. West Nile virus (WNV) tests identify units in transfusion transmission. Transfusion 2003;43:Suppl:2A. abstract.
16. Stramer SL, Glynn SA, Kleinman SH, et al. Detection of HIV-1 and HCV infections among antibody-negative blood donors by nucleic acid–amplification testing. N Engl J Med 2004;351:760-8.
17. Sazama K, Shoos-Lipton K. Update on WNV-related activities and considerations, 2004: a summary of the WNV Task Force meeting. AABB bulletin. No. 04-03. Bethesda, Md.: American Association of Blood Banks, 2004:1-3.
18. Prince HE, Hogrefe WR. Detection of West Nile virus (WNV)-specific immunoglobulin M in a reference laboratory setting during the 2002 WNV season in the United States. Clin Diagn Lab Immunol 2003;10:764-8.
19. Biggerstaff BJ, Petersen LR. Estimated risk of transmission of the West Nile virus through blood transfusion in the US, 2002. Transfusion 2003;43:1007-17.
20. Petersen LR, Hayes EB. Westward ho? — the spread of West Nile virus. N Engl J Med 2004;351:2257-9.
21. Roehrig JT, Nash D, Maldin B, et al. Persistence of virus-reactive serum immunoglobulin M antibody in confirmed West Nile virus encephalitis cases. Emerg Infect Dis 2003;9:376-9.
22. Detection of West Nile virus in blood donations — United States, 2003. MMWR Morb Mortal Wkly Rep 2003;52:769-72.
23. Update: West Nile virus screening of blood donations and transfusion-associated transmission — United States, 2003. MMWR Morb Mortal Wkly Rep 2004;53:281-4.
24. Transfusion-associated transmission of West Nile virus — Arizona, 2004. MMWR Morb Mortal Wkly Rep 2004;53:842-4.
25. Macedo de Oliveira A, Beecham BD, Montgomery SP, et al. West Nile virus blood transfusion-related infection despite nucleic acid testing. Transfusion 2004;44:1695-9.
26. 2003 and 2004 West Nile virus activity in the U.S. as of April 29, 2005. (Accessed July 8, 2005, at http://www.cdc.gov/ncidod/dvbid/westnile/index.htm.)
27. Custer B, Tomasulo P, Murphy EL, et al. Triggers for switching from minipool testing by nucleic acid technology to individual-donation nucleic acid testing for West Nile virus: analysis of 2003 data to inform 2004 decision making. Transfusion 2004;44:1547-54.
28. Dodd RY. Emerging infections, transfusion safety, and epidemiology. N Engl J Med 2003;349:1205-6.

ORIGINAL ARTICLE

Screening the Blood Supply for West Nile Virus RNA by Nucleic Acid Amplification Testing

Michael P. Busch, M.D., Ph.D., Sally Caglioti, M.T.(A.S.C.P.), Eugene F. Robertson, Ph.D., Joan D. McAuley, M.T.(A.S.C.P.), Leslie H. Tobler, Dr.P.H., Hany Kamel, M.D., Jeffrey M. Linnen, Ph.D., Venkatakrishna Shyamala, Ph.D., Peter Tomasulo, M.D., and Steven H. Kleinman, M.D.

ABSTRACT

BACKGROUND

The use of nucleic acid amplification tests of "minipools" of 16 samples to screen blood donors for West Nile virus RNA began in July 2003. We report the yield and characteristics of positive donations and the incremental yield and safety of nucleic acid amplification tests of individual donations.

METHODS

Reactive minipools were analyzed to identify the individual reactive donations. For the regions with the highest yield on minipool testing, retrospective nucleic acid amplification testing was performed on individual donations that were negative on minipool testing. Reactive donations were confirmed by alternative nucleic acid amplification tests and IgM and IgG tests, and donors were followed to document seroconversion.

RESULTS

From July 1 through October 31, 2003, 677,603 donations were prospectively screened for West Nile virus by minipool testing, yielding 183 confirmed viremic donations (0.027 percent, or 1 in 3703 donations). Retrospective individual testing of 23,088 donations from high-prevalence regions that were negative on minipool testing yielded 30 additional units with a low level of viremia, with 14 additional viremic units detected by prospective testing of individual donations late in the 2003 transmission season. Of all the viremic units detected, 5 percent were detected only by individual testing and were negative for IgM antibody, 29 percent were detected by individual testing after IgM seroconversion, and 66 percent were detected by minipool testing. West Nile virus infection was confirmed in both recipients of IgM-negative units that were reactive on individual testing, whereas neither recipient of antibody-positive blood components that were reactive on individual testing was infected. In 2004, prospective testing of individual donations in regions that yielded donations that were reactive on minipool testing resulted in a 32 percent incremental yield of units with a low level of viremia that would have been missed by minipool testing.

CONCLUSIONS

Although nucleic acid amplification testing of minipools of blood donations prevented hundreds of cases of West Nile virus infection in 2003, it failed to detect units with a low level of viremia, some of which were antibody-negative and infectious. These data support the use of targeted nucleic acid amplification testing of individual donations in high-prevalence regions, a strategy that was implemented successfully in 2004.

From the Blood Systems Research Institute, San Francisco (M.P.B., L.H.T., S.H.K.); the Department of Laboratory Medicine, University of California, San Francisco (M.P.B.); Blood Systems Laboratories, Tempe, Ariz. (S.C., E.F.R., J.D.M.); Blood Systems, Scottsdale, Ariz. (H.K., P.T.); Gen-Probe, San Diego, Calif. (J.M.L.); Chiron, Emeryville, Calif. (V.S.); and the University of British Columbia, Victoria, B.C., Canada (S.H.K.).

N Engl J Med 2005;353:460-7.

West Nile virus, a mosquito-borne flavivirus, emerged as a cause of meningoencephalitis in the United States in 1999, and infections reached epidemic proportions in 2002.[1,2] In 2002, West Nile virus was shown to be transmissible by transfusion, when 23 cases were documented.[3,4] In late 2002, the Food and Drug Administration (FDA), U.S. blood-collecting organizations, and test-kit manufacturers accelerated programs to develop nucleic acid amplification tests to screen blood donors for West Nile viremia in an effort to implement such programs before the 2003 transmission season.[4-7] The resulting assays involved the testing of pools of 6 to 24 samples, or "minipools," an approach that is now also routinely used to screen blood donors for human immunodeficiency virus (HIV) and hepatitis C virus (HCV).[8]

We report the results of a large, multicenter testing program conducted during the summer and fall of 2003 and of studies that compared minipool results with those obtained by nucleic acid amplification tests of individual (undiluted) donations. Our results allow an estimation of the number of infections averted in 2003 by the implementation of screening with minipool nucleic acid amplification tests and an estimation of the additional benefit of screening individual donations in regions with seasonal epidemics of West Nile virus infections. We also summarize experience with implementation of a targeted screening strategy involving nucleic acid amplification testing of individual donations in 2004, a strategy that successfully identified units with a low level of viremia that would have been missed by minipool testing.

METHODS

Blood Systems Laboratories screens blood donations in two FDA-licensed laboratories in Tempe, Arizona, and Bedford, Texas. Clients include 18 blood-collection facilities owned by Blood Systems (Scottsdale, Ariz.) and 59 other community-based and hospital-based blood-collection programs. The geographic areas covered by this client base include much of the Southwest, the South, the Central Plains, and parts of California. The annual testing volume is approximately 2.2 million donations, which represents approximately 20 percent of the U.S. blood supply.

The West Nile virus Transcription-Mediated Amplification system (Procleix WNV Assay, Gen-Probe and Chiron) was used for nucleic acid amplification tests. This technique involves lysis of viral particles in plasma, either from individual donations or from a minipool of plasma specimens from 16 donations and the isolation of West Nile virus RNA with the use of probes bound to magnetic beads, amplification with the use of RNA transcription, and subsequent detection by a chemiluminescent probe.[9,10] All samples within a reactive minipool are then tested individually. The assay has an analytical sensitivity of approximately 4 RNA copies per milliliter when used for individual donations (50 percent limit of detection by probit analysis of dilutions of West Nile virus standards) and a sensitivity of approximately 45 copies per milliliter when used for minipool testing.[10]

Individual specimens identified as reactive were evaluated by means of a confirmatory algorithm with the use of multiple assays for West Nile virus.[10-12] These tests included an alternative nucleic acid amplification test (either a modification of a TaqMan polymerase-chain-reaction [PCR] assay or another primer-based transcription-mediated amplification assay, performed by Bayer Reference Laboratory) and assays of plasma for viral IgM and IgG antibody (Focus Diagnostics).[10,13] All donors with reactive tests were promptly asked to enroll in a follow-up study, involving return visits approximately every week. Follow-up specimens were tested for West Nile virus RNA and for IgM and IgG antibodies against West Nile virus. A confirmed positive result was defined by the detection of viral IgM in either the index specimen or a follow-up specimen, the detection of viral RNA in the index specimen by means of the alternative nucleic acid amplification test, or the detection of viral RNA in a follow-up specimen by means of a transcription-mediated amplification assay.[11] Since the actual times at which donors returned for follow-up varied (resulting in intermittent blood collection) and since seroconversion would have occurred in the interval between the last seronegative and the first seropositive result, median times (and interquartile ranges) to IgM and IgG seroconversion were estimated with the use of an analysis in which data were censored in the intervals between visits.[14] The viral load in confirmed positive index donations for which frozen plasma components were available for analysis was evaluated by a kinetic PCR assay based on target-capture TaqMan techniques (Chiron).[10]

Blood Systems Laboratories also conducted geo-

graphically and temporally targeted nucleic acid amplification testing of individual donations in 2003 according to two protocols. The first protocol involved testing of individual frozen specimens from donations previously found to be negative on nucleic acid amplification testing of minipools of specimens from regions with a high prevalence of West Nile virus infections. Donors whose specimens were retrospectively determined to be reactive on individual testing were asked to enroll in the follow-up study. In addition, in-stock blood products from these donors were retrieved, and in collaboration with the Centers for Disease Control and Prevention (CDC), recipients of blood components from donors confirmed to be positive for West Nile virus on nucleic acid amplification testing of individual donations were evaluated to ascertain whether transmission of West Nile virus had occurred.[15] The second protocol was conducted in selected blood-collection regions that had had a high number of reactive minipools on nucleic acid amplification testing in the previous weeks. Individual donations were prospectively screened exclusively by nucleic acid amplification testing (i.e., minipool testing was not performed). Reactive units were subsequently diluted 1:16 and retested individually to determine what the results of minipool testing would have been. No blood components that were reactive on nucleic acid amplification testing of individual donations were transfused during this prospective study.

In June 2004 Blood Systems Laboratories implemented a targeted screening strategy involving nucleic acid amplification testing of individual donations (described elsewhere in detail[16]) and real-time tracking of the results of minipool testing. Prospective nucleic acid amplification testing of individual donations was implemented in geographically defined zones if there were two or more reactive donations on minipool testing and a rate of more than 1 reactive minipool per 1000 tested. Testing reverted to the minipool format when regions had had no individual donations with reactive tests for at least seven consecutive days and had a weekly rate of reactivity of fewer than 1 per 1000 donations.

All studies were approved by the FDA and the relevant institutional review boards. All donors and recipients gave written informed consent to undergo screening and follow-up testing for West Nile virus. Drs. Busch, Tomasulo, and Kleinman and Ms. Caglioti designed the studies; supervised data collection, management, and analyses; and drafted and revised the manuscript. Drs. Robertson, Tobler, Linnen, and Shyamala and Ms. McAuley supervised testing. Dr. Kamel supervised donor follow-up and participated in "look-back" activities involving recipients of blood components. All authors approved the manuscript, which was written primarily by Drs. Busch and Kleinman.

RESULTS

The results of minipool nucleic acid amplification testing during the active West Nile virus season in 2003 are shown in Figure 1. Of 677,603 donations tested between July 1 and October 31, 2003, 183 were confirmed to be positive, for an aggregate rate of 0.027 percent (1 in 3703 donations). The highest rates occurred during a six-week period from mid-July through mid-August.

Forty-seven additional viremic units were detected on nucleic acid amplification testing of individual donations. Retrospective testing of 23,088 individual donations that had been negative on minipool testing, collected from donor centers in Texas, North Dakota, and South Dakota (i.e., centers with high rates of reactivity on minipool testing during the summer of 2003), identified 30 additional confirmed positive specimens. Prospective testing of individual donations was subsequently performed on 3964 donations collected in North Dakota and South Dakota in September 2003 and identified 17 confirmed positive donations, of which 14 tested negative for West Nile virus with the use of the transcription-mediated amplification system at a 1:16 dilution, indicating they would have been missed by minipool testing. Thus, 186 units were detectable by minipool testing and 44 units were detectable only by nucleic acid amplification testing of individual donations.

West Nile virus antibody status was determined for 41 of the 44 donations that were negative on minipool testing and confirmed positive by individual testing (3 samples had insufficient volume for testing). Thirty-one specimens (76 percent) had detectable West Nile virus antibody: 10 were IgM-positive, and 21 were positive for both IgM and IgG. In contrast, only 16 of the 183 confirmed positive donations detected by minipool testing (9 percent) were positive for IgM antibody at the time of donation ($P<0.001$ by the chi-square test). Among 145 viremic donors who were initially seronegative and who enrolled in the follow-up assessment, West

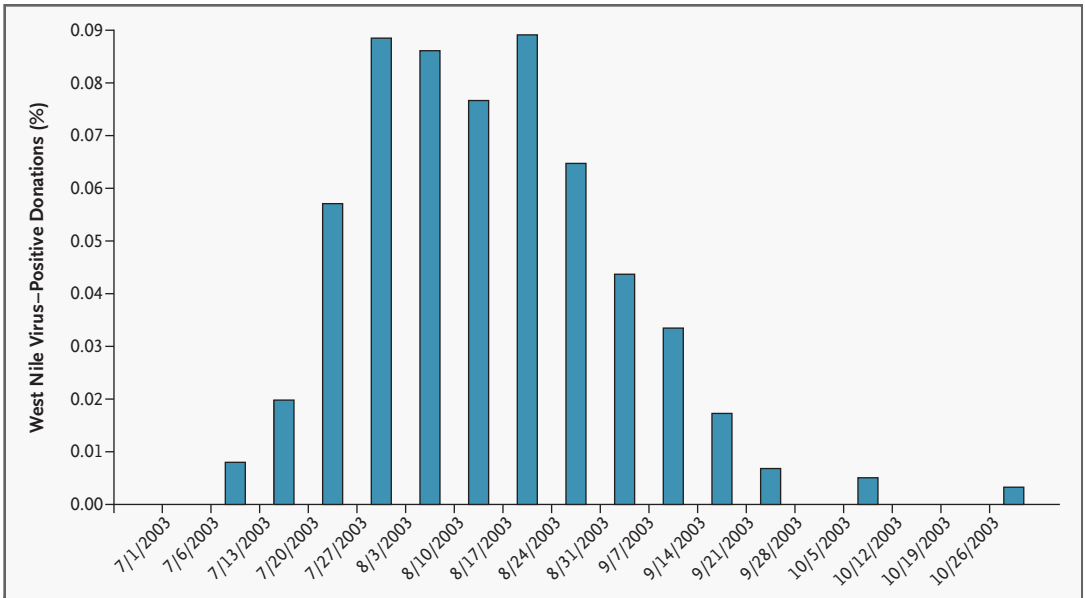


Figure 1. Results of Minipool Nucleic Acid Amplification Testing during the 2003 West Nile Virus Season.

Nile virus–specific IgM antibody appeared a median of four days after donation (interquartile range, one to six) and virus-specific IgG antibody appeared a median of two days later (interquartile range, one to five).

The median viral load for 143 of the 183 specimens confirmed to be positive on minipool testing with sufficient volume for quantitative PCR analysis was 3519 copies per milliliter (range, less than 50 to 690,159). Twelve IgM-positive donations identified by minipool testing had significantly lower viral loads than 131 IgM-negative donations identified by minipool testing (median, less than 50 and 5325 copies per milliliter, respectively; $P<0.001$ by the two-sample Wilcoxon rank test). As expected, viral loads were very low in donations detected only by nucleic acid amplification testing of individual donations: 21 of these 44 samples (48 percent) were reactive on only one of two replicate tests, indicating that the viral load was near the limit of detection of the transcription-mediated amplification assay, and 16 of 22 evaluated by TaqMan PCR (73 percent) had RNA levels that were below the limit of quantitation (i.e., fewer than 50 copies per milliliter).

We performed a subanalysis of 113 confirmed viremic donations identified from July 1 through September 30, 2003, in North Dakota and South Dakota from a donor population that had been screened with the use of both minipool and individual nucleic acid amplification testing. Table 1 shows the overall rates of detection of units confirmed positive by minipool testing and by individual testing alone, as well as the incremental rate of detection of viremia by nucleic acid amplification testing of individual donations throughout the epidemic, with adjustment to account for the proportion of units tested individually. Minipool testing detected 66 percent of viremic units detected by nucleic acid amplification testing of individual donations. Five percent of viremic donations were detectable by individual testing alone and were negative for West Nile virus antibody; 7 percent were reactive on individual testing alone and were positive for IgM but negative for IgG; and 22 percent were reactive on individual testing alone and were positive for both IgM and IgG (Table 1).

On the basis of previous studies of inoculation of West Nile virus in humans[17] and animals,[18,19] acute-phase infection is thought to be characterized by a brief period of very-low-level viremia shortly after inoculation (as reflected by the viral-load data presented above and the finding of IgM-negative specimens that were reactive on nucleic acid amplification testing of individual donations). This period is followed by a longer interval (approximately seven days) with an increasing and then decreasing viral load, which makes the viremia de-

Table 1. Yield of Minipool and Individual Nucleic Acid Amplification Testing of Donations from North Dakota and South Dakota, July 1 through September 30, 2003.

Variable	Total No. of Donations	Minipool Testing	Individual Nucleic Acid Amplification Testing Alone			
			Total No.	IgM–	IgM+, IgG–	IgM+, IgG+
No. of donations tested	27,009	27,009	22,641	22,641	22,641	22,641
No. of confirmed positive units	113	79	34*	5	6	21
Rate of detection (no. of confirmed positive units/100 donations)	0.443	0.292	0.150	0.024†	0.028†	0.099†
Adjusted no. of confirmed positive units‡	119	79	40	6	8	26
Proportional yield (%)	100	66	34	5	7	22

* Two units had insufficient volume for serologic testing.
† The value was adjusted to account for serologic classification of 32 of 34 units that were identified by individual testing alone.
‡ The value was adjusted to account for individual testing of 22,641 of the 27,009 units that underwent minipool testing.

tectable by minipool testing. Finally, as IgM and IgG seroconversion evolves, the viral load decreases to a level detectable only by testing of individual donations. Given this natural history, Figure 2 presents the biweekly yield data for North Dakota and South Dakota according to the results of individual and minipool tests and antibody tests. The figure shows that although testing of individual donations identified additional donations with a low level of viremia throughout the epidemic, the characteristics of these units shifted from primarily antibody-negative early in the epidemic to predominantly IgM-positive and IgG-positive late in the season ($P<0.001$ by Fisher's exact test).

Case investigations were initiated for 17 recipients of blood components from 14 donations that were negative on minipool testing but identified as having a low level of viremia on retrospective nucleic acid amplification testing of individual donations. On the basis of clinical symptoms and serologic analysis for West Nile virus, two recipients of seronegative donations with low-level viremia were infected with West Nile virus, probably as a result of transfusion, whereas two recipients of components from one donation that was reactive on individual testing alone and was positive for IgM and IgG were not infected.[12] The evaluation of the remaining 13 recipients was deemed inconclusive owing to a lack of follow-up laboratory data to support or rule out West Nile virus infection (Montgomery S and Brown J, CDC: personal communication).

During the 2004 epidemic (from May 1 to October 23, 2004), analysis of 1,065,212 donations by minipool testing yielded 71 confirmed viremic donations (Table 2). An additional 58,679 donations (5 percent of all donations tested) were prospectively tested individually with the use of previously described triggers,[16] and 54 donations were confirmed to be viremic. Sufficient volume was available to test 48 of these 54 donations at a 1:16 dilution: 27 were negative (and thus classified as reactive on individual testing alone), and 21 were positive (and thus classified as detectable by minipool testing). Of the 27 donations identifiable by individual testing alone, 23 were IgM-positive and 4 were IgM-negative. Thus, as seen in Table 2, targeted testing of individual donations in the regions of the 2004 epidemic yielded percentages of units detectable by minipool testing (76 percent) and by individual testing alone (24 percent) and a serologic profile for units detectable by individual testing alone that were similar to the percentages and profile observed in North Dakota and South Dakota during the 2003 epidemic (66 percent and 34 percent, respectively).

DISCUSSION

The implementation of nucleic acid amplification testing for West Nile virus RNA in 2003 resulted in the identification of 183 confirmed viremic units, with 47 additional infected units detected by targeted testing of individual donations. Nationally, the combination of minipool testing and targeted testing of individual donations resulted in the identification of approximately 1000 viremic donations.[11,15] Since, on average, each unit is made into 1.45 transfusable components,[20] the transfusion

of almost 1500 viremic components (most of which lacked antibody and would be expected to be infectious) was averted in 2003.

Our study confirms previous data from clinical cases and experimental inoculation studies indicating that infected persons would probably have low titers of West Nile virus.[2,3,12,15,17-19] Unlike seronegative donors with HIV and HCV infection and positive results on minipool testing, who usually have viral titers of 10^5 to 10^7 copies per milliliter,[6] in our study, the median number of copies of West Nile virus RNA was only about 3500. This observation of a low viral load combined with data indicating that proven transfusion-transmitted cases of West Nile virus infection occurred from donors with low viral titers[3,15] prompted us to study retrospectively the incremental value of individual nucleic acid amplification testing as compared with minipool testing and to implement individual testing prospectively in selected high-prevalence regions in late 2003 and 2004. We found that individual testing identified up to 50 percent more viremic donors than were detected by minipool testing. Five percent of all viremic donations were identifiable by individual testing alone and were antibody-negative and thus were donations that have been shown to be infectious. (Both recipients of units that were antibody-negative and reactive on individual nucleic acid amplification testing alone in our study were infected.) These donations were detected at a fairly constant rate throughout the epidemic. In contrast, the additional yield of antibody-positive donations identified by means of individual testing alone was minimal when tests were performed during the early weeks of the epidemic, but toward the end of the epidemic it increased to levels greater than those observed with minipool testing.

The incremental safety to be achieved by the use of individual testing over minipool testing is difficult to quantify because of the unknown risk of transmission by donations with low-level viremia that contain West Nile virus antibody. The absence of transmission of West Nile virus from two antibody-positive components identified by individual testing alone in this study is consistent with the observation that no documented case of post-transfusion infection has been attributed to a seroreactive donation, despite the relatively high frequency of such donations during the later stages of West Nile virus epidemics.[5,10,15] This is also consistent with in vitro and animal infectivity experiments suggesting that IgM and IgG antibodies neutralize infectivity.[19] Similar experience with hepatitis A virus, an acute infection that also leads to the production of neutralizing IgM antibody, has indicated that the concurrence of viremia and IgM production does not result in infectivity.[7,21] Thus, one working hypothesis is that viremic units that contain IgM (and particularly those that also contain IgG) are not infectious when transfused.

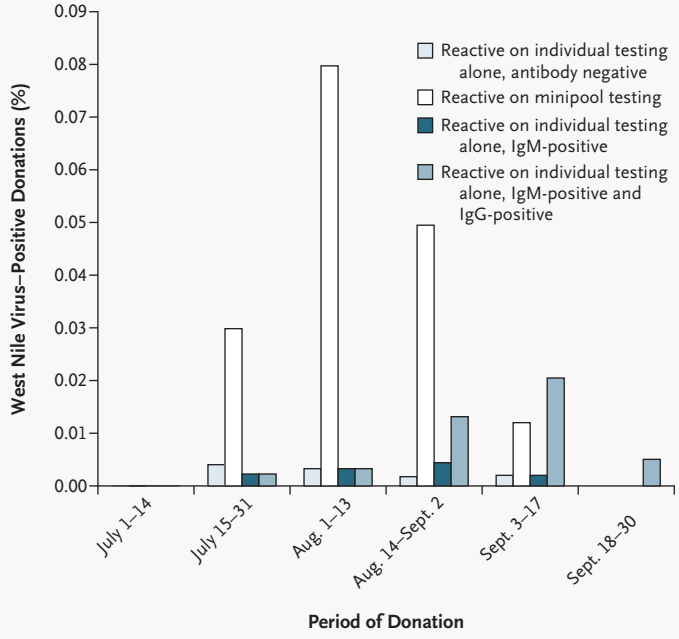


Figure 2. Rate of West Nile Virus–Positive Donations Detected in North Dakota and South Dakota by Minipool and Individual Nucleic Acid Amplification Testing.

The pattern of observed viremia (i.e., the percentage of West Nile virus RNA–positive donations detected by minipool testing as compared with the percentage detected by individual testing) and seroreactivity (presence of IgM and IgG) of viremic units changed dramatically through the 10 weeks of the epidemic, strongly suggesting that the greatest benefit of individual testing can be obtained by implementing such screening early in the epidemic when new infections are on the rise. This observation led Blood Systems Laboratories to adopt a new nucleic acid amplification testing strategy during the 2004 epidemic. A designated level of reactive minipool tests in a defined geographic region was used to determine when sufficient risk existed to implement individual testing.[16] This strategy

Table 2. Actual and Adjusted Systemwide Yield of Minipool and Individual Nucleic Acid Amplification Testing of Donations Screened from May 1 through October 23, 2004.

Type of Nucleic Acid Amplification Test	Total Tested	Confirmed Positive Results		Adjusted Confirmed Positive Results*	
	no. (%)	*no. (%)*	*no./100 donations*	*no. (%)*	*no./100 donations*
Minipool	1,065,212 (95)	71 (57)	0.007	95 (76)	0.009
Individual	58,679 (5)	54 (43)*	0.092	30 (24)	0.051
Total	1,123,891 (100)	125	0.012	125	0.012

* Of the 54 samples confirmed to be positive by nucleic acid amplification testing of individual donations, 48 were available for retesting at a 1:16 dilution; 27 of the latter (56 percent) were found to be negative and therefore classified as reactive on individual testing alone. This percentage was applied to the 54 positive units detected by individual testing to project that 30 of these donations would be positive with the use of this method alone and that 24 would have been detected by minipool testing. The reclassification is reflected in the column that provides adjusted confirmed positive results.

was designed to balance the residual risk of transfusion-transmitted West Nile virus infection accompanying the use of minipool testing against the limited capacity for individual testing, given the current limitations of automation.[5] This strategy proved highly effective, with the identification and removal of at least 27 units that would have been missed by minipool testing in regions with West Nile virus epidemic activity. They included at least four units that were IgM-negative and hence likely to transmit West Nile virus to recipients. Moreover, the similarity of the systemwide distribution of the yields of the two tests in 2004 (an increase in the rate of detection by approximately 32 percent with the use of individual over minipool testing) with that observed in a region of epidemic infection in 2003 indicates that our targeting of individual testing was appropriately directed to regions with an increased yield.

Before the initiation of blood-donor screening, information from experimental studies of the inoculation of West Nile virus conducted in the early 1950s in patients with advanced cancer indicated that the duration of viremia (as assayed by intracerebral injection of virus into mice) was approximately six or seven days.[17] Contemporary data from primate and murine models of West Nile virus infection are consistent with this estimate.[7,18,19] On the basis of the proportional rates of detection of viremia in the early phases of infection in asymptomatic viremic blood donors, we estimate that the duration of antibody-negative viremia detectable only by nucleic acid amplification testing of individual donations is shorter (one or two days). Our data cannot be used to estimate the length of the phase in which specimens are reactive on individual testing and antibody-positive, since a cross-sectional analysis would be biased: the frequency of donation in this convalescent phase is reduced as a result of the signs and symptoms of West Nile virus infection.[2] As compared with other transfusion-transmissible infections (e.g., HIV and HCV), West Nile virus has a similar interval in which it is detectable only by nucleic acid amplification testing of individual donations.[22] In contrast, the duration of viremia detectable by minipool testing is much shorter for West Nile virus infections than it is for HIV and HCV infections, in which high-titer viremia is detected for weeks or months before seroconversion and usually persists for many years after seroconversion.

In conclusion, although the use of minipool screening in 2003 prevented hundreds of West Nile virus infections, it failed to detect donations with a low level of viremia, some of which were antibody-negative and infectious. Our 2003 data supported the use of targeted nucleic acid amplification testing of individual donations in high-prevalence regions, a strategy that was successfully implemented in 2004. On the basis of the price of reagents for previously licensed nucleic acid amplification tests, the costs of performing minipool and individual screening at Blood Systems Laboratories, and the observed yields of minipool testing, the cost of minipool screening was $120,000 per unit intercepted in 2003 and $232,000 per unit intercepted in 2004. The reduced cost utility in 2004 reflects the decreased rate of viremic donations detected, illustrating the close inverse relationship between yield and cost-effectiveness. The targeted individual screening of approximately 60,000 donations in the summer of 2004 resulted in a 33 percent increment in yield and cost only $32,000 per incremen-

tal case detected. In contrast, had individual testing been performed for West Nile virus during all of 2004, the cost per viremic donation detected would have been $281,000. This approach of performing targeted testing of individual donations on the basis of real-time monitoring of the yield of minipool testing may prove to be a rational and cost-effective donor-screening paradigm for other agents similar to West Nile virus that cause seasonal and regional epidemics.

Supported in part by a grant (R01-CI-000214) from the Centers for Disease Control and Prevention (CDC) and the Blood Systems Foundation.

Dr. Busch is an employee of Blood Systems, a not-for-profit company that collects and tests donated blood for West Nile virus and other infections as described in this article; is a member of Scientific Advisory Boards of Chiron and Gen-Probe, the manufacturers of the West Nile virus assays that were evaluated in this study; and reports having received an unrestricted research grant from Chiron and speaking honoraria from Chiron and Gen-Probe. Ms. Caglioti, Dr. Robertson, Ms. McAuley, Dr. Tobler, Dr. Kamel, and Dr. Tomasulo are employees of Blood Systems. Dr. Linnen is an employee of Gen-Probe and reports owning equity stock and stock options in Gen-Probe. He is named on patents filed by Gen-Probe for the West Nile virus nucleic acid amplification assay described in this article. Dr. Shyamala is an employee of Chiron and is named on patents filed by Chiron for the West Nile virus nucleic acid amplification assay described in this article.

Reagents for performing the retrospective and prospective West Nile virus transcription-mediated amplification testing on collections in 2003 were supplied to Blood Systems at no charge by Gen-Probe. The costs for reagents to perform minipool screening in 2003 and 2004 and individual screening in 2005 were paid to Chiron and Gen-Probe on a cost-reimbursement basis, as detailed in an FDA Investigational New Drug (IND) application. All other sample-collection, shipping, and labor costs related to West Nile virus screening and acquisition of follow-up specimens were funded by Blood Systems, with reimbursement from hospitals or other blood centers. The cost of follow-up serologic and PCR testing was covered by Gen-Probe and Chiron as part of the IND application.

We are indebted to the staff at Blood Systems Laboratories and the United Blood Services and other donor centers for their effort in support of this study and to Susan Montgomery and Jennifer Brown at the CDC for coordination of look-back investigations of recipients exposed to West Nile virus.

REFERENCES

1. Provisional surveillance summary of the West Nile virus epidemic — United States, January–November 2002. MMWR Morb Mortal Wkly Rep 2002;51:1129-33.
2. Petersen LR, Marfin AA, Gubler DJ. West Nile virus. JAMA 2003;290:524-8.
3. Pealer LN, Marfin AA, Petersen LR, et al. Transmission of West Nile virus through blood transfusion in the United States in 2002. N Engl J Med 2003;349:1236-45.
4. Dodd RY. Emerging infections, transfusion safety, and epidemiology. N Engl J Med 2003;349:1205-6.
5. Epstein JS. Insights on donor screening for West Nile virus. Transfusion 2005;45:460-2.
6. Food and Drug Administration, Center for Biologics Evaluation and Research. Workshop on development of donor screening assays for West Nile Virus, Bethesda, Md., November 4-5, 2002. (Accessed July 8, 2005, at http://www.fda.gov/cber/meetings/wnv110402.htm.)
7. Hollinger FB, Kleinman S. Transfusion transmission of West Nile virus: a merging of historical and contemporary perspectives. Transfusion 2003;43:992-7.
8. Stramer SL, Glynn SA, Kleinman SH, et al. Detection of HIV-1 and HCV infections among antibody-negative blood donors by nucleic acid–amplification testing. N Engl J Med 2004;351:760-8.
9. Giachetti C, Linnen JM, Kolk DP, et al. Highly sensitive multiplex assay for detection of human immunodeficiency virus type 1 and hepatitis C virus RNA. J Clin Microbiol 2002;40:2408-19.
10. Busch MP, Tobler LH, Saldanha J, et al. Analytical and clinical sensitivity of West Nile virus RNA screening and supplemental assays available in 2003. Transfusion 2005;45:492-9.
11. Kleinman S, Glynn SA, Busch M, et al. The 2003 West Nile virus United States epidemic: the America's Blood Centers experience. Transfusion 2005;45:469-79.
12. Update: detection of West Nile virus in blood donations — United States, 2003. MMWR Morb Mortal Wkly Rep 2003;52:916-9. [Erratum, MMWR Morb Mortal Wkly Rep 2003;52:942.]
13. Prince HE, Hogrefe HR. Detection of West Nile virus (WNV)-specific immunoglobulin M in a reference laboratory setting during the 2002 WNV season in the United States. Clin Diagn Lab Immunol 2003;10:764-8.
14. Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. J R Stat Soc B 1976;38:290-5.
15. Update: West Nile virus screening of blood donations and transfusion-associated transmission — United States, 2003. MMWR Morb Mortal Wkly Rep 2004;53:281-4.
16. Custer BS, Tomasulo PA, Murphy EL, et al. Triggers for switching from minipool testing by nucleic acid technology to individual-donation nucleic acid testing for West Nile virus: analysis of 2003 data to inform 2004 decision making. Transfusion 2004;44:1547-54.
17. Southam CH, Moore AE. Induced virus infections in man by the Egypt isolates of West Nile virus. Am J Trop Med Hyg 1954;3:19-50.
18. Ratterree MS, Gutierrez RA, Travassos da Rosa AP, et al. Experimental infection of rhesus macaques with West Nile virus: level and duration of viremia and kinetics of the antibody response after infection. J Infect Dis 2004;189:669-76.
19. Diamond MS, Sitati EM, Friend LD, Higgs S, Shrestha B, Engle M. A critical role for induced IgM in the protection against West Nile virus infection. J Exp Med 2003;198:1853-62.
20. Marshall DA, Kleinman SH, Wong JB, et al. Cost-effectiveness of nucleic acid test screening of volunteer blood donations for hepatitis B, hepatitis C, and human immunodeficiency virus in the United States. Vox Sang 2004;86:28-40.
21. Krugman S, Ward R, Giles JP, Bodansky O, Jacobs AM. Infectious hepatitis: detection of virus during the incubation period and in clinically inapparent infection. N Engl J Med 1959;261:729-34.
22. Busch MP, Glynn SA, Stramer SL, et al. A new strategy for estimating risks of transfusion-transmitted viral infections based on rates of detection of recently infected donors. Transfusion 2005;45:254-64.

ORIGINAL ARTICLE

The Prognostic Value of a Nomogram for Exercise Capacity in Women

Martha Gulati, M.D., Henry R. Black, M.D., Leslee J. Shaw, Ph.D., Morton F. Arnsdorf, M.D., C. Noel Bairey Merz, M.D., Michael S. Lauer, M.D., Thomas H. Marwick, M.D., Ph.D., Dilip K. Pandey, Ph.D., Roxanne H. Wicklund, R.N., and Ronald A. Thisted, Ph.D.

ABSTRACT

BACKGROUND

Recent studies have demonstrated that exercise capacity is an independent predictor of mortality in women. Normative values of exercise capacity for age in women have not been well established. Our objectives were to construct a nomogram to permit determination of predicted exercise capacity for age in women and to assess the predictive value of the nomogram with respect to survival.

METHODS

A total of 5721 asymptomatic women underwent a symptom-limited, maximal stress test. Exercise capacity was measured in metabolic equivalents (MET). Linear regression was used to estimate the mean MET achieved for age. A nomogram was established to allow the percentage of predicted exercise capacity to be estimated on the basis of age and the exercise capacity achieved. The nomogram was then used to determine the percentage of predicted exercise capacity for both the original cohort and a referral population of 4471 women with cardiovascular symptoms who underwent a symptom-limited stress test. Survival data were obtained for both cohorts, and Cox survival analysis was used to estimate the rates of death from any cause and from cardiac causes in each group.

RESULTS

The linear regression equation for predicted exercise capacity (in MET) on the basis of age in the cohort of asymptomatic women was as follows: predicted MET=14.7−(0.13×age). The risk of death among asymptomatic women whose exercise capacity was less than 85 percent of the predicted value for age was twice that among women whose exercise capacity was at least 85 percent of the age-predicted value (P<0.001). Results were similar in the cohort of symptomatic women.

CONCLUSIONS

We have established a nomogram for predicted exercise capacity on the basis of age that is predictive of survival among both asymptomatic and symptomatic women. These findings could be incorporated into the interpretation of exercise stress tests, providing additional prognostic information for risk stratification.

From the Department of Preventive Medicine, Rush University Medical Center, Chicago (M.G., H.R.B.); the Division of Cardiology, Department of Medicine, Cedars–Sinai Research Institute, Cedars–Sinai Medical Center, Los Angeles (L.J.S., C.N.B.M.); the Departments of Medicine (M.F.A.) and Health Studies (R.A.T.), University of Chicago, Chicago; the Department of Cardiology, Cleveland Clinic Foundation, Cleveland (M.S.L.); the Department of Medicine, Princess Alexandra Hospital, Brisbane, Australia (T.H.M.); the Department of Neurology, University of Illinois, Chicago (D.K.P.); and the Division of Cardiology, St. James Hospital and Health Centers, Chicago Heights, Ill. (R.H.W.). Address reprint requests to Dr. Gulati at Northwestern University, Bluhm Cardiovascular Institute, 201 E. Huron St., Chicago, IL 60611, or at mgulati@nmff.org.

N Engl J Med 2005;353:468-75.

Exercise capacity is an independent predictor of the risk of death and cardiac events among asymptomatic women and men.[1-5] Exercise capacity may be defined as the maximal oxygen uptake for a given workload[6,7] and can be expressed in metabolic equivalents (MET), or multiples of the basal rate of oxygen consumption when a person is at rest (3.5 ml per kilogram of body weight per minute for an average adult).[8] Exercise capacity can be estimated by performing a symptom-limited stress test.

Exercise capacity varies with age, sex, and health. A number of studies have established that there is a negative linear relationship between exercise capacity and age in men, and a nomogram has been developed for men that estimates the percentage of predicted exercise capacity for a given age.[6,7,9-11] Few studies have evaluated exercise capacity in women, and to date, no standard for age-related declines in physical fitness has been established for women.

We had two goals in conducting this study. The first was to create a simple nomogram for women to allow the conversion of the MET value achieved on a stress test into a percentage of the predicted exercise capacity for any age, on the basis of findings in a population of asymptomatic women. The second goal was to assess the usefulness of the nomogram in predicting survival among both asymptomatic women and a referral population of women with cardiovascular symptoms in order to determine its usefulness in clinical practice.

METHODS

ASYMPTOMATIC POPULATION

The asymptomatic population came from the St. James Women Take Heart Project. This cohort has previously been described.[1] Briefly, in 1992 volunteers were solicited from the Chicago metropolitan area to participate in a study of heart disease in women. The project was approved by the institutional review board of St. James Hospital and Rush–Presbyterian–St. Luke's Medical Center, with written informed consent obtained from all study participants.

Inclusion criteria for the study cohort were an age of at least 35 years and the ability to walk on a treadmill at a moderate pace. Women were excluded if they were pregnant, had typical anginal symptoms or any history of cardiac disease (including previous myocardial infarction, documented coronary artery disease, heart failure, or valvular heart disease), weighed more than 148 kg (325 lb), had a baseline blood pressure of 170/110 mm Hg or higher, or had incomplete data concerning cardiac risk factors. Members of the cohort were classified as sedentary or active on the basis of their response to one question: Do you have a regular (exercise) training program?

REFERRAL POPULATION

The referral population came from the Economics of Noninvasive Diagnosis Study, which has been described previously.[12,13] Briefly, this cohort was composed of consecutive women from six medical centers who were referred between 1990 and 1995 for a symptom-limited exercise stress test with the use of the Bruce protocol[14] for the evaluation of suspected coronary disease. Women were excluded if they had recently been hospitalized for unstable angina, myocardial infarction, or coronary revascularization. The study was approved by the institutional review board at each of the six participating centers. For all but one site, written informed consent for follow-up was obtained at the time of the initial procedure. For the remaining site, the requirement for informed consent was waived by the institutional review board because the data were from a previously approved database.

EXERCISE TREADMILL TESTING

All participants underwent a symptom-limited treadmill test according to the Bruce protocol.[14] The test was discontinued in the event of limiting symptoms (angina, dyspnea, or fatigue), abnormalities of rhythm or blood pressure, or marked and progressive ST-segment deviation. Target heart rates were not used as a predetermined end point.

In the asymptomatic group, some participants underwent stress testing according to a modified Bruce protocol. These women were excluded from the analysis, because submaximal exercise testing does not accurately reflect physical fitness in the way that maximal exercise testing does.

EXERCISE CAPACITY

The estimated exercise capacity was measured in MET[6,7] as defined above. The estimate was based on the speed and grade of the treadmill.[15]

FOLLOW-UP DATA

The number of deaths from any cause in the asymptomatic cohort was determined by searching the

National Death Index to identify deaths and causes of death from the time of the baseline evaluation in 1992 through the end of 2000. Follow-up information was obtained on the referral (symptomatic) population through 2000 during a clinic visit or telephone interview. Deaths were identified and the cause of death was classified after a review of death certificates by an independent reviewer who was unaware of the women's clinical history and stress-testing data.

STATISTICAL ANALYSIS

The MET achieved was determined from the final speed and grade of the treadmill, as defined for the Bruce protocol.[15] Using the asymptomatic population, we calculated the linear regression of exercise capacity (in MET) on age; no evidence of nonlinearity was found. The calculated value from the regression equation for age was defined as 100 percent of the age-predicted exercise capacity.

For each participant in both groups, the percentage of the predicted exercise capacity achieved for age was then calculated with the use of the following equation: percentage of predicted exercise capacity achieved for age=(observed MET÷age-predicted MET)×100. A nomogram to determine the percentage of predicted exercise capacity for age was constructed with the use of the linear regression equation for the asymptomatic cohort. Similarly, separate nomograms were also created for both the active and sedentary groups.

The correlation of the percentage of predicted exercise capacity for any given age with subsequent survival was calculated in both groups. The rates of death from any cause and from cardiac causes were analyzed for both populations with the use of univariate Cox proportional-hazards models, on the basis of the deviation from the predicted normal value of exercise capacity for age. From the Cox models, the predicted rates of death from cardiac causes were plotted against the ratio of observed exercise capacity to expected exercise capacity, and an r^2 statistic was calculated. Annualized death rates and rates of death from cardiac causes were calculated by dividing the predicted death rates from the Cox model by the length of follow-up. All analyses were performed with the use of STATA software (version 8.0) or SPSS software (version 12.0). A two-sided P value of less than 0.05 was considered to indicate statistical significance.

Table 1. Characteristics of the Study Cohorts.*

Characteristic	Asymptomatic Women (N=5721)	Symptomatic Women (N=4471)
Age — yr		
Mean	52±11	61±12
Range	35–86	34–93
Race — %†		
White	85	62
Black	9	28
Other	6	10
Risk factors — %†		
Hypertension	17	49
Diabetes	5	22
Current or former smoker	21	25
Family history of CAD	44	35
Hypercholesterolemia	17	12
Typical angina — %	—	58
Atypical or nonanginal CP — %	—	25
Dyspnea or symptoms of CHF — %	—	17
Exercise capacity — MET		
Peak	8.0±2.7	6.9±3.4
Range	1.4–20.0	1.2–17.4
Death — no. (%)		
Any cause	180 (3)	537 (12)
Cardiac causes	58 (1)	45 (1)
Duration of follow-up — yr	8.4±0.7	5.3±2.1

* Plus–minus values are means ±SD. CAD denotes coronary artery disease, CP chest pain, and CHF congestive heart failure.
† Race and risk factors were self-reported.

RESULTS

A total of 5721 asymptomatic women and 4471 symptomatic women met the study-specific inclusion criteria.[1,12,13,16] The characteristics of the two cohorts are shown in Table 1. The asymptomatic cohort was younger than the symptomatic cohort and was predominantly white (85 percent). Although the majority of the symptomatic women were white, almost a third were black. Fewer asymptomatic women had a history of hypertension or diabetes. Atypical symptoms accounted for only a quarter of the symptoms in the symptomatic cohort. The asymptomatic women had a higher mean exercise capacity and a lower overall mortality rate during the follow-up period than did the symptomatic women.

PREDICTED EXERCISE CAPACITY

The relationship between exercise capacity and age in the cohort of 5721 asymptomatic women was linear. Regression analysis of exercise capacity for age yielded the following equation: predicted MET=14.7−(0.13×age), with an age-adjusted SD of 2.3 (r= −0.51, P<0.001).

The same regression analysis was stratified according to the level of reported activity. A total of 212 women did not respond to the exercise question and were excluded from this analysis. For the active subgroup of 866 women, the regression equation was as follows: predicted MET=17.9−(0.16×age), with an age-adjusted SD of 2.4 (r= −0.59, P<0.001). For the sedentary subgroup of 4643 women, the regression equation was as follows: predicted MET=14.0−(0.12×age), with an age-adjusted SD of 2.2 (r= −0.49, P<0.001).

The nomogram for the entire asymptomatic population (the first equation) is shown in Figure 1, with the previously reported nomogram for asymptomatic men provided for comparison.[6] The nomograms based on self-reported activity level (from the second and third equations) are shown in Figure 2. The active women had a greater predicted exercise capacity for any given age than their more sedentary counterparts.

Using the first regression equation to predict normal exercise capacity for age and the equation for the percentage of predicted exercise capacity achieved for age, we determined the percentage of predicted exercise capacity achieved for each participant in both populations. The results ranged from 20 percent to 150 percent of the predicted value for age.

EXERCISE CAPACITY AND PROGNOSIS

In the asymptomatic cohort, there were 180 deaths overall (3 percent) and 58 deaths from cardiac causes (1 percent) during a mean follow-up of 8.4 years (Table 1). Women in this cohort were assumed to be alive if they were not identified as having died by

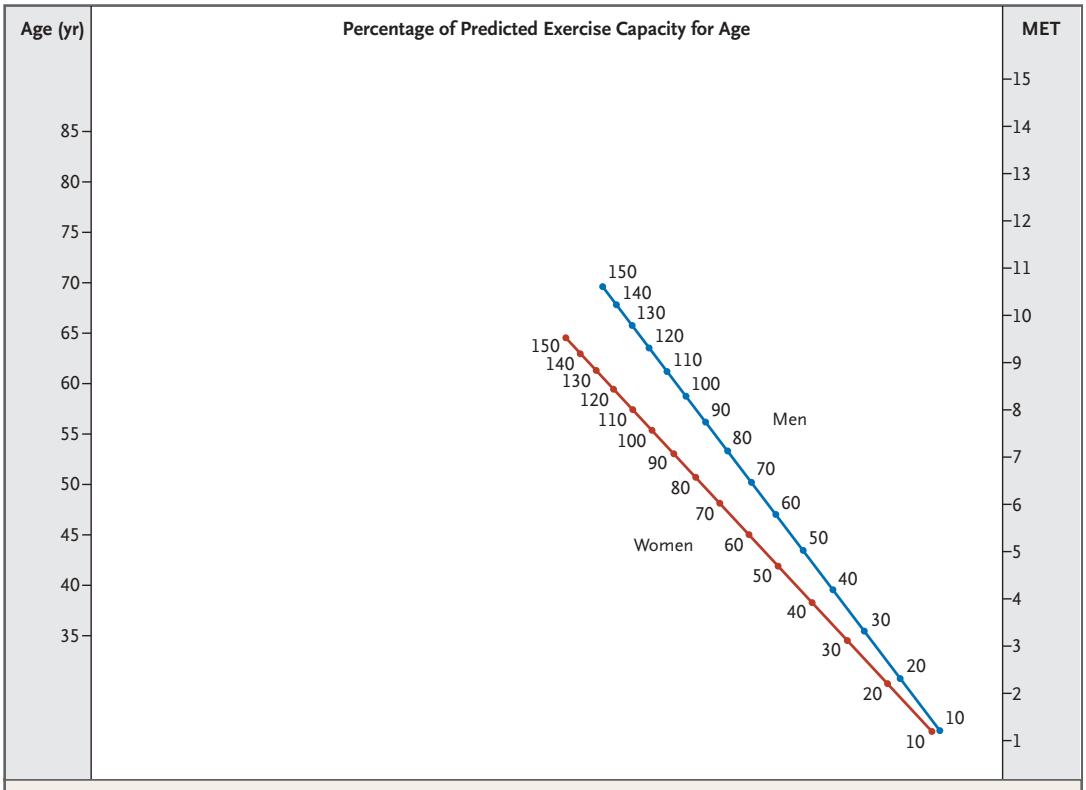


Figure 1. Nomogram of the Percentage of Predicted Exercise Capacity for Age in Asymptomatic Men and Women.

A line drawn from the patient's age on the lefthand scale to the MET value on the righthand scale will cross the percentage line at the point corresponding to the patient's percentage of predicted exercise capacity for age. The nomogram for men was modified from Morris et al.[6]

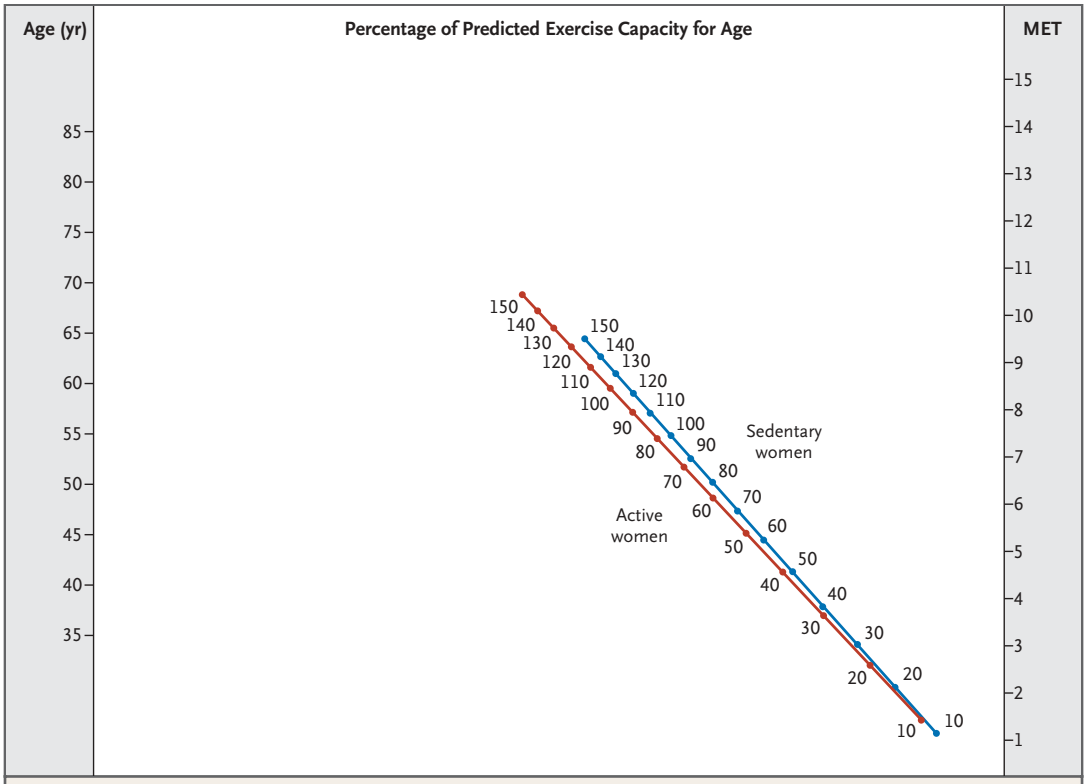


Figure 2. Nomogram of the Percentage of Predicted Exercise Capacity for Age in Sedentary and Active Women in the Asymptomatic Cohort.

A line drawn from the patient's age on the lefthand scale to the MET value on the righthand scale will cross the percentage line at the point corresponding to the patient's percentage of predicted exercise capacity for age.

a search of the National Death Index. In the symptomatic cohort, there were 537 deaths overall (12 percent) and 45 deaths from cardiac causes (1 percent) during a mean follow-up of 5.3 years (Table 1); 3 percent of this cohort was lost to follow-up.

In the asymptomatic cohort, the women whose exercise capacity was less than 85 percent of the age-predicted value had a hazard ratio for death from any cause of 2.03 ($P<0.001$) and a hazard ratio for death from cardiac causes of 2.44 ($P<0.001$), as compared with the women whose exercise capacity was at least 85 percent of the age-predicted value (Table 2). In the asymptomatic cohort, as compared with women whose exercise capacity exceeded the age-predicted value by more than 3 MET, women whose exercise capacity was less than that predicted for age (for whom the observed exercise capacity minus the predicted exercise capacity was less than 0 MET) had a hazard ratio for death from any cause of 2.63 ($P=0.005$) and for death from cardiac causes of 4.27 ($P=0.045$) (Table 2). The amount of deviation from one's age-predicted exercise capacity was correlated with the risk of both death from any cause and death from cardiac causes.

In the symptomatic population, as compared with the women whose exercise capacity was at least 85 percent of the age-predicted value, women whose exercise capacity was less than 85 percent of the age-predicted value had a hazard ratio for death from any cause of 2.37 ($P<0.001$) and for death from cardiac causes of 2.02 ($P<0.001$) (Table 2). As compared with women whose exercise capacity exceeded the age-predicted value by more than 3 MET, women whose exercise capacity was less than that predicted for age (i.e., a deviation of less than 0 MET) had a hazard ratio for death from any cause of 3.28 ($P<0.001$) and for death from cardiac causes of 3.80 ($P<0.001$) (Table 2). The relationship between exercise capacity and the risk of death from cardiac causes was remarkably similar for all age groups in the symptomatic cohort, with two

Table 2. Hazard Ratios for Death from Any Cause and from Cardiac Causes among Women, According to the Deviation from the Expected Exercise Capacity for Age.

Exercise Capacity	Death from Any Cause		Death from Cardiac Causes	
	Asymptomatic Women	Symptomatic Women	Asymptomatic Women	Symptomatic Women
	hazard ratio (95 percent confidence interval)			
Exercise capacity <85% of predicted value for age*	2.03 (1.51–2.71)	2.37 (1.90–2.97)	2.44 (1.46–4.09)	2.02 (1.43–2.85)
Observed exercise capacity minus predicted exercise capacity				
>3 MET†	1.0	1.0	1.0	1.0
0–3 MET	1.70 (0.84–3.44)	1.89 (1.44–2.48)	2.02 (0.46–8.82)	2.21 (1.31–3.27)
<0 MET	2.63 (1.33–5.19)	3.28 (2.47–4.35)	4.27 (1.03–17.6)	3.80 (2.26–6.38)

* The reference group is women whose exercise capacity was at least 85 percent of that predicted for age.
† This group served as the reference group.

exceptions. Young women (less than 55 years of age) with a poor exercise capacity (for whom the observed exercise capacity minus the predicted exercise capacity was less than −2 MET) had an especially high mortality rate, as did the oldest women (older than 70 years) for whom the observed exercise capacity minus the predicted exercise capacity was less than 1 MET (Fig. 3).

USE OF THE NOMOGRAM

Use of the nomogram for the percentage of predicted exercise capacity for age (Fig. 1) requires only the woman's age and exercise capacity achieved (in MET) on the exercise stress test. Drawing a straight line between the age and exercise capacity will allow the determination of the percentage of predicted exercise capacity for age; a value of 100 percent is the mean for any given age. Any result greater than 100 percent indicates better-than-average performance. Any result lower than 100 percent indicates some degree of functional impairment for age. For example, a 60-year-old woman whose exercise capacity was 7 MET on a Bruce-protocol exercise test would have achieved 100 percent of the predicted exercise capacity for her age. In contrast, a 35-year-old woman whose exercise capacity was also 7 MET would have achieved 69 percent of her age-predicted exercise capacity.

DISCUSSION

The first goal of this study was to define the mean age-predicted exercise capacity for women, as depicted by the nomogram in Figure 1. Although such a nomogram has been established for men and is routinely used in clinical practice, no such nomogram has been established for women, nor have the previous findings in men been validated in the female population.

A number of regression equations for predicting exercise capacity in a variety of male populations have been described.[6,7,9-11] Three of these studies examined the relationship of exercise capacity to age in healthy men.[6,7,9] For the asymptomatic women in our study, the regression equation for predicting exercise capacity for age was similar to the regression equations established for healthy men.[6,7,9] In particular, it was very similar to the equation derived from the 244 healthy men in the study by Morris et al.[6]: predicted MET=14.7−(0.11×age), with an age-adjusted SD of 2.5 ($r=-0.53$, $P<0.001$). Both equations for asymptomatic men and women share the same constant in the regression equation, and the coefficient differs by only 0.02. Although this number appears small at first glance, the coefficient is multiplied by age. This means that the difference in the predicted exercise capacity for age between men and women will increase as age increases. The nomograms for men and women, when shown side by side, clearly demonstrate the difference between the sexes (Fig. 1).

Within our cohort of asymptomatic women, the regression equation differed on the basis of self-reported physical-activity status. Women were classified as either sedentary or active, on the basis of one question regarding their participation in a regular activity program. Although the validity of this

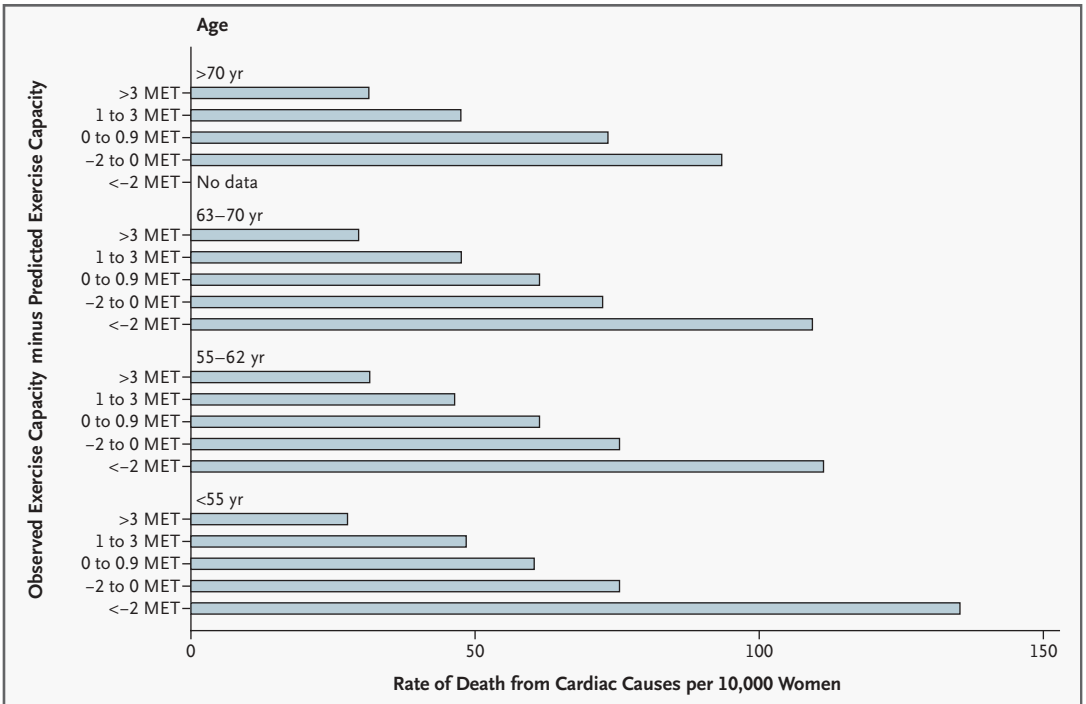


Figure 3. Rate of Death from Cardiac Causes as a Function of the Deviation from the Expected Exercise Capacity for Age among Symptomatic Women.

The rate of death from cardiac causes was stratified according to the difference between the observed and predicted exercise capacity (in MET) in the various age groups.

question has not been established, we have shown a difference in achieved exercise capacity for any age group on the basis of activity status; the more physically active women had a greater exercise capacity at all ages.

The nomogram developed in the cohort of asymptomatic women was also used to determine the percentage of predicted exercise capacity in a referral cohort of racially diverse women from six institutions. Deviation from the established normal values was a significant predictor of the risk of death from any cause and from cardiac causes in both the symptomatic and asymptomatic groups. The further the deviation below the predicted normative value, the greater the risk of death.

The use of the women's nomogram results in a more accurate assessment of prognosis among the women than does the use of the men's nomogram in this group. The sensitivity and specificity of our survival model for predicting the risk of death from any cause among the asymptomatic women are 70 percent and 47 percent, respectively, when the women's nomogram is used. In contrast, the sensitivity and specificity are 55 percent and 64 percent, respectively, when the men's nomogram is used. Use of the men's nomogram in our cohort would have resulted in 800 more false positive results as compared with the 30 fewer false negative results.

The chief limitation of our study is that the nomogram was created from data on a volunteer cohort of asymptomatic, mostly white women. The referral population differed from the asymptomatic cohort. Traditional cardiac risk factors were expected to be more prevalent in the symptomatic population than in the asymptomatic population, and indeed, these women were older and were more likely to have hypertension and diabetes. The symptomatic women were also more racially diverse, with a stronger representation of black women. Prior studies have suggested that clinicians should use a nomogram for the particular population from which it was created. However, the nomogram we have developed was predictive of the risk of death in the symptomatic group, with hazard ratios similar to those in the asymptomatic group from which it was derived. Whether another nomogram derived

from a symptomatic cohort or from a population with more black women would be a better predictor of the risk of death in our referral cohort is open to question. Such an approach could, however, imply the need for a different nomogram for every clinical population tested, an approach that is not likely to be practical and that might in any case result in findings not very different from ours.

We estimated exercise capacity on the basis of the speed and degree of incline of the treadmill. In contrast, in the study by Morris et al., ventilatory gas exchange was measured directly in asymptomatic men during the stress test.[6] This distinction is important, since determination of MET levels from a stress test has been demonstrated to overestimate the exercise capacity.[17-19] If oxygen consumption had been measured directly in our study population, the nomograms for men and women would presumably have differed even more.

Despite extensive research on the role of exercise stress testing and exercise capacity, there has been a paucity of data on women, particularly asymptomatic women. Thus, what is normal or expected for healthy women has not been well established. We have developed a nomogram for women that can be used to predict a woman's expected exercise capacity at any given age and have demonstrated that the resulting measure is a predictor of the risk of death in both asymptomatic and symptomatic cohorts.

Supported by grants from AstraZeneca, DuPont Pharmaceuticals, the Irwin Foundation, Merck, Pfizer/Pharmacia, Siemans Gammasonics, and St. James Hospital and Health Centers (Chicago Heights, Ill.).

Drs. Gulati and Thisted have a patent pending for the nomogram described in this article. The rights to this patent are owned by Rush University Medical Center and the University of Chicago. Dr. Arnsdorf reports being paid for his role as co-editor-in-chief (cardiovascular medicine) at UpToDate. Dr. Merz reports having received consulting fees from Pfizer, Bayer, Fujisawa, and Merck; lecture fees from Pfizer, AM Medica Communications, Merck, and KOS Pharmaceuticals; and medical-education grants from Pfizer, Wyeth, Procter & Gamble, Novartis, AstraZeneca, and Bristol-Myers Squibb and owning stock in Boston Scientific, IVAX, Lilly, Medtronic, Johnson & Johnson, SCIPIE Insurance, ATS Medical, and Biosite.

This article is dedicated to the memory of Dr. Arfan Al-Hani, who designed the St. James Women Take Heart Project. Without his foresight, enthusiasm, and dedication, this study would not exist.

We are indebted to the participants of the St. James Women Take Heart Project and the Economics of Noninvasive Diagnosis Study Group.

REFERENCES

1. Gulati M, Pandey DK, Arnsdorf MF, et al. Exercise capacity and the risk of death in women: the St James Women Take Heart Project. Circulation 2003;108:1554-9.
2. Mora S, Redberg RF, Cui Y, et al. Ability of exercise testing to predict cardiovascular and all-cause death in asymptomatic women: a 20-year follow-up of the Lipid Research Clinics Prevalence Study. JAMA 2003;290:1600-7.
3. Blair SN, Kohl HW III, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: a prospective study of healthy men and women. JAMA 1989;262:2395-401.
4. Ekelund LG, Haskell WL, Johnson JL, Whaley FS, Criqui MH, Sheps DS. Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men: the Lipid Research Clinics Mortality Follow-up Study. N Engl J Med 1988;319:1379-84.
5. Wei M, Kampert JB, Barlow CE, et al. Relationship between low cardiorespiratory fitness and mortality in normal-weight, overweight, and obese men. JAMA 1999;282:1547-53.
6. Morris CK, Myers J, Froelicher VF, Kawaguchi T, Ueshima K, Hideg A. Nomogram based on metabolic equivalents and age for assessing aerobic exercise capacity in men. J Am Coll Cardiol 1993;22:175-82.
7. Bruce RA, Kusumi F, Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. Am Heart J 1973;85:546-62.
8. Jette M, Sidney K, Blumchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol 1990;13:555-65.
9. Froelicher VF Jr, Allen M, Lancaster MC. Maximal treadmill testing of normal USAF aircrewmen. Aerosp Med 1974;45:310-5.
10. Dehn MM, Bruce RA. Longitudinal variations in maximal oxygen intake with age and activity. J Appl Physiol 1972;33:805-7.
11. Wolthuis RA, Froelicher VF Jr, Fischer J, Triebwasser JH. The response of healthy men to treadmill exercise. Circulation 1977;55:153-7.
12. Marwick TH, Shaw LJ, Lauer MS, et al. The noninvasive prediction of cardiac mortality in men and women with known or suspected coronary artery disease. Am J Med 1999;106:172-8.
13. Marwick TH, Case C, Vasey C, Allen S, Short L, Thomas JD. Prediction of mortality by exercise echocardiography: a strategy for combination with the Duke treadmill score. Circulation 2001;103:2566-71.
14. Bruce RA. Exercise testing of patients with coronary heart disease: principles and normal standards for evaluation. Ann Clin Res 1971;3:323-32.
15. Foster C, Jackson AS, Pollock ML, et al. Generalized equations for predicting functional capacity from treadmill performance. Am Heart J 1984;107:1229-34.
16. Shaw LJ, Hendel R, Cerquiera M, et al. Ethnic differences in the prognostic value of stress technetium-99m tetrofosmin gated single-photon emission computed tomography myocardial perfusion imaging. J Am Coll Cardiol 2005;45:1494-504.
17. Myers J, Buchanan N, Walsh D, et al. Comparison of the ramp versus standard exercise protocols. J Am Coll Cardiol 1991;17:1334-42.
18. Sullivan M, McKirnan MD. Errors in predicting functional capacity for postmyocardial infarction patients using a modified Bruce protocol. Am Heart J 1984;107:486-92.
19. Roberts JM, Sullivan M, Froelicher VF, Genter F, Myers J. Predicting oxygen uptake from treadmill testing in normal subjects and coronary artery disease patients. Am Heart J 1984;108:1454-60.

ORIGINAL ARTICLE

Modafinil for Excessive Sleepiness Associated with Shift-Work Sleep Disorder

Charles A. Czeisler, Ph.D., M.D., James K. Walsh, Ph.D., Thomas Roth, Ph.D., Rod J Hughes, Ph.D., Kenneth P. Wright, Ph.D., Lilliam Kingsbury, Ph.D., Sanjay Arora, Ph.D., Jonathan R.L. Schwartz, M.D., Gwendolyn E. Niebler, D.O., and David F. Dinges, Ph.D., for the U.S. Modafinil in Shift Work Sleep Disorder Study Group*

ABSTRACT

BACKGROUND

Patients with shift-work sleep disorder chronically have excessive sleepiness during night work and insomnia when attempting to sleep during the day. We evaluated the use of modafinil for treating sleepiness in patients with this disorder.

METHODS

In a three-month, double-blind trial, we randomly assigned 209 patients with shift-work sleep disorder to receive either 200 mg of modafinil or placebo before the start of each shift. Assessments were performed with the use of the nighttime Multiple Sleep Latency Test, the Clinical Global Impression of Change, the Psychomotor Vigilance Test, diaries of patients, and daytime polysomnography. After randomization, we conducted monthly assessments.

RESULTS

Treatment with modafinil, as compared with placebo, resulted in a modest improvement from baseline in mean (±SEM) nighttime sleep latency (the interval between the time a person attempts to fall asleep and the onset of sleep) (1.7±0.4 vs. 0.3±0.3 minutes, respectively; P=0.002), and more patients had improvement in their clinical symptoms (74 percent vs. 36 percent, respectively; P<0.001). Patients who were receiving modafinil also had a reduction in the frequency and duration of lapses of attention during nighttime testing of their performance on the Psychomotor Vigilance Test (change from baseline, a reduction in lapse frequency of 2.6 vs. an increase of 3.8, respectively; P<0.001), and proportionally fewer patients reported having had accidents or near accidents while commuting home (29 percent vs. 54 percent, respectively; P<0.001). Despite these benefits, patients treated with modafinil continued to have excessive sleepiness and impaired performance at night. Modafinil did not adversely affect daytime sleep as compared with placebo. Headache was the most common adverse event.

CONCLUSIONS

Treatment with 200 mg of modafinil reduced the extreme sleepiness that we observed in patients with shift-work sleep disorder and resulted in a small but significant improvement in performance as compared with placebo. However, the residual sleepiness that was observed in the treated patients underscores the need for the development of interventions that are even more effective.

From the Division of Sleep Medicine, Harvard Medical School, and the Division of Sleep Medicine, Department of Medicine, Brigham and Women's Hospital — both in Boston (C.A.C., K.P.W.); the Sleep Medicine Research Center, St. Luke's Hospital, Chesterfield, Mo. (J.K.W.); the Sleep Disorders and Research Center, Henry Ford Hospital, Detroit (T.R.); Cephalon, Frazer, Pa. (R.J.H., L.K., S.A., G.E.N.); the Sleep and Chronobiology Laboratory, Department of Integrative Physiology, University of Colorado, Boulder (K.P.W.); Integris Southwest and Baptist Medical Centers, Oklahoma City (J.R.L.S.); and the Division of Sleep and Chronobiology, Department of Psychiatry and Center for Sleep and Respiratory Neurobiology, University of Pennsylvania School of Medicine, Philadelphia (D.F.D.). Address reprint requests to Dr. Czeisler at Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Ave., Suite 438, Boston, MA 02115, or at caczeisler@hms.harvard.edu.

Drs. Walsh, Roth, and Hughes contributed equally to this article.

*Other members of the study group are listed in the Appendix.

N Engl J Med 2005;353:476-86.

NEARLY 6 MILLION AMERICANS WORK at night on a permanent or rotating basis.[1] Night-shift work disrupts both sleep and waking because of the misalignment of circadian regulation and sleep–wake behavior.[2,3] In about 5 to 10 percent of night-shift workers, the sleep–wake disturbance is severe enough to warrant diagnosis as shift-work sleep disorder,[4-6] which is characterized by a level of excessive sleepiness during night work and insomnia when attempting to sleep in the daytime that is judged to be clinically significant.[4,5] Persons with shift-work sleep disorder miss family and social activities more frequently and have higher rates of ulcers, sleepiness-related accidents, absenteeism, and depression than do night-shift workers without the disorder[6] — conditions long known to affect a subgroup of shift workers.[7] We conducted a study to evaluate the efficacy and safety of 200 mg of modafinil in patients with excessive sleepiness associated with chronic shift-work sleep disorder. This agent has shown efficacy in the treatment of narcolepsy and the residual excessive sleepiness in patients with obstructive sleep apnea.[8-11]

METHODS

PATIENTS

Adults between the ages of 18 and 60 years were eligible if they worked each month at least five night shifts for 12 hours or less, with 6 hours or more worked between 10 p.m. and 8 a.m. and at least three shifts occurring consecutively. Patients were diagnosed with shift-work sleep disorder in accordance with criteria stipulated in the *International Classification of Sleep Disorders*.[5] Our diagnostic criteria included a primary symptom of excessive sleepiness on the night shift and insomnia during opportunities for daytime sleep and the absence of other primary sleep disorders, other medical conditions, and medications that might cause sleepiness. Patients had to have reported chronic excessive sleepiness (≥3 months) during night shifts; a Clinical Global Impression of Severity[12] rating of moderately ill or worse for sleepiness on work nights, including the commute home from work; an average latency to sleep onset of 6 minutes or less during 20-minute nap opportunities at 2-hour intervals during the night, as measured by the Multiple Sleep Latency Test[13,14]; and a sleep efficiency of 87.5 percent or less as determined by daytime polysomnography. Participants provided written informed consent.

STUDY DESIGN AND CONDUCT

The three-month randomized, double-blind, placebo-controlled study was conducted in the United States between December 2001 and September 2002. At each of 28 centers, an institutional review board approved the informed-consent statement and protocol. The study included a screening visit to assess eligibility; a baseline visit on the night after the patient had worked three or more consecutive night shifts in order to establish pretreatment levels of alertness and performance, the severity of sleepiness, and results of daytime polysomnography; and a randomization visit for the administration of the study drug. Thereafter, patients were evaluated monthly during an overnight laboratory shift after having worked for three or more consecutive nights.

Patients were randomly assigned (in a 1:1 ratio) to receive 200 mg of modafinil (Provigil, Cephalon), formulated as 100-mg tablets, or an identical-appearing placebo, taken 30 to 60 minutes before the start of each night shift. Assessment of treatment adherence was performed at visits after the initial baseline visit.

MEASURES OF EFFICACY

Sleep latency during laboratory night shifts was measured by polysomnography at two-hour intervals, starting at 2 a.m., with the use of the Multiple Sleep Latency Test (see the Supplementary Appendix, available with the full text of this article at www.nejm.org).[13,14] To assess alertness with the use of a performance measure, a 20-minute Psychomotor Vigilance Test[15] was administered every two hours, starting at 1 a.m. The level of sleepiness as reported by patients was assessed hourly using the Karolinska Sleepiness Scale,[16] which ranges from 1 (very alert) to 9 (very sleepy). The investigator-rated Clinical Global Impression of Change,[12] the scoring of which ranges from 1 (very much improved) to 7 (very much worse), was used to assess changes from baseline in the severity of sleepiness during night shifts, including the commute to and from work. Patients also completed electronic diaries containing questions about sleepiness, sleep, and caffeine use during the night shift and the commute home.

There were two prespecified primary efficacy variables. The first was the rating on the Clinical Global Impression of Change test for sleepiness during the night shift, including the commute to and from work, at the final visit. This test assessed the extent to which treatment effects could be rec-

ognized by patients and physicians. The second prespecified primary efficacy variable was the change between baseline and the final visit (i.e., at the third month or at withdrawal from the study) in overall mean sleep latency on the basis of results of the nighttime Multiple Sleep Latency Test. This test has been validated as a measure of sleepiness during the day but not at night, so questions remain regarding which results on this test indicate pathological sleepiness during the night and what level of improvement at night is clinically meaningful. Therefore, as recommended,[17] one of our secondary outcome measures — the frequency and duration of lapses of attention during performance on the Psychomotor Vigilance Test[15] — served as a validated and objective measure of alertness at night.

SAFETY ASSESSMENTS

Adverse events were monitored throughout the study. Blood pressure and heart rate were monitored, and clinical laboratory tests (including chemical and hematologic studies) were conducted at each visit. Physical examination and electrocardiography were performed at the screening visit and the final visit.

OTHER ASSESSMENTS

Polysomnography was conducted for eight hours, starting at 10 a.m. after baseline and after final laboratory night shifts. Melatonin concentrations were measured from saliva samples.[18]

STATISTICAL ANALYSIS

A total of 204 patients were selected for enrollment in the study (see the Supplementary Appendix for calculations regarding the sample size). Randomization was performed with the use of a central randomization process and stratified by center with the use of permuted blocks of two. On the basis of the sample that was available for efficacy analysis, there was at least 80 percent power (with the use of post hoc power calculations) for the statistical inference on both prespecified primary end points (assuming an alpha level of 0.05).

Comparisons of continuous demographic variables between groups were conducted with the use of analysis of variance, with treatment as a factor. Discrete categorical demographic variables were compared with the use of the chi-square test or Fisher's exact test. Included in efficacy analyses were patients who had been randomly assigned to treatment and received at least one dose of study drug and who had had a baseline assessment and at least one assessment after baseline on either the Multiple Sleep Latency Test or the Clinical Global Impression of Change. For the final-visit efficacy analysis, data from the patients' last visit on or before the third month were used.

Comparisons between the groups were conducted on the change between the baseline visit and the final visit with regard to variables on the Multiple Sleep Latency Test, scores on the Karolinska Sleepiness Scale, polysomnographic measures, and data from patients' diaries related to sleepiness ratings, unintentional and intentional sleep episodes during the night shift, the consumption of caffeinated drinks, and sleep efficiencies. All of these variables were analyzed with the use of analysis of variance, with treatment and site as factors. Data from the Clinical Global Impression of Change test were evaluated with the use of a Cochran–Mantel–Haenszel chi-square test, with adjustment made for site and modified ridit scores used to account for ordered categories.

Comparisons were made with the use of a chi-square test or Fisher's exact test on data from diaries that were related to the percentage of patients reporting mistakes, near accidents, or accidents during the night shift; unintentional sleep episodes, accidents, or near accidents during the commute home; and rates of adverse events. Comparisons of performance on the Psychomotor Vigilance Test and of melatonin phase were performed with the use of the Wilcoxon nonparametric rank test. There were no interim analyses of the data. All reported P values are two-sided and not adjusted for multiple testing.

Six academic investigators and four representatives of the sponsor designed the study and analyzed the data. Drs. Czeisler, Walsh, Roth, and Dinges conceived of and designed the study in collaboration with representatives of the corporate sponsors, Drs. Hughes, Niebler, Arora, and Kingsbury. Data were fully accessible to all group members, with the study sponsor placing no limits on interpretation or publication. The study designers vouch for the completeness and accuracy of the analyses. All authors were involved with the preparation of the manuscript.

RESULTS

DISPOSITION AND BASELINE CHARACTERISTICS OF PATIENTS

A total of 4533 shift workers were prescreened through telephone calls placed to a central agency;

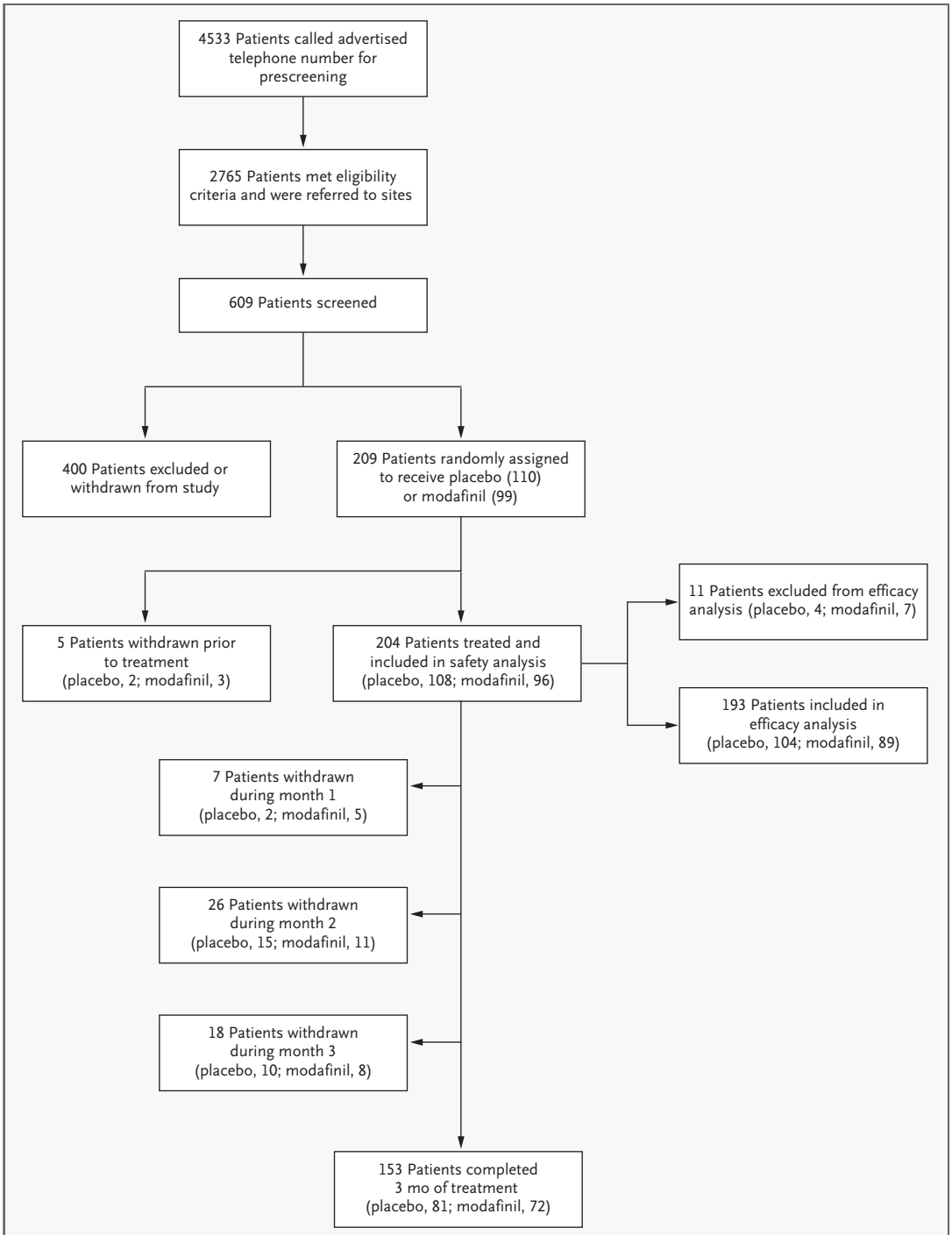


Figure 1. Enrollment and Status of Patients in the Study.

Patients who were initially screened to participate in the study were recruited by advertisements or were referred by investigators. Patients who were randomly assigned to treatment and received at least one dose of study drug and who had a baseline assessment and at least one assessment after baseline for any given variable were included in the efficacy analysis (see the Supplementary Appendix for more details).

2765 workers were referred to study sites. Of 609 patients undergoing laboratory testing, 400 were judged ineligible or withdrew (Fig. 1 and Supplementary Appendix). Most commonly, ineligible patients did not meet inclusion criteria for polysomnography (107 patients) or sleep latency (53) or either withdrew their consent or were lost to follow-up (118). Of 209 patients who were randomly assigned to receive the study drug, 204 patients received the drug, and 153 patients completed the study. At baseline, there were no significant differences in demographic variables, shift-work type, sleepiness, performance, and results on polysomnography between the group that received modafinil and the one that received placebo (Table 1). The patients who completed the trial and those who did not had similar baseline values for the primary outcome variables (as measured by the Multiple Sleep Latency Test and the Clinical Global Impression of Severity test) and similar results on polysomnography. Patients were severely sleepy at baseline, with overall mean (±SD) sleep latencies of 2.0±1.8 minutes and 2.1±1.5 minutes for the placebo and modafinil groups, respectively.

EFFICACY MEASURES

Seventy-four percent of patients in the modafinil group were rated as at least minimally improved on the Clinical Global Impression of Change test at the final visit, as compared with 36 percent in the placebo group (P<0.001) (Fig. 2A, and Table 2 in the Supplementary Appendix). Overall mean (±SEM) sleep latency, as measured by the Multiple Sleep Latency Test, increased from 2.1 minutes at baseline to 3.8 minutes at the final visit with modafinil (change, 1.7±0.4 minutes; P<0.001) but not with placebo (2.04 at baseline vs. 2.37 at the final visit; change, 0.3±0.3; P=0.24) (Fig. 2B). Sleep latency was significantly greater in the modafinil group than in the placebo group (P=0.002). This improvement in sleep latency with modafinil versus placebo was found at 2 a.m. (P=0.02) and 4 a.m. (P<0.001) (Fig. 2C), but not at 6 a.m. (P=0.45) or 8 a.m. (P=0.17) (Fig. 2D). A higher proportion of patients receiving modafinil had a positive change in the sleep-latency score from pretreatment to the final visit (Fig. 1 in the Supplementary Appendix). Notwithstanding this improvement, sleep latencies during the night shift averaged less than six minutes, which is below the level considered normal during the daytime.

Significant differences between the modafinil group and the placebo group were also found for performance on the Psychomotor Vigilance Test. The median number of lapses of attention in 20-minute tests during the night was 12.50 at baseline and 10.25 at the final visit for the modafinil group (median change from baseline, −2.6; P=0.012). In the placebo group, the median number of lapses per test bout was 16.13 at baseline and 23.75 at the final visit (median change from baseline, 3.8; P=0.008). The groups did not differ significantly at baseline (P=0.797), but they did differ significantly at the final visit (P=0.005), and the change in lapses of attention during performance of the Psychomotor Vigilance Test from baseline to the final visit was significant for modafinil versus placebo (P<0.001) (Fig. 2E).

The duration of lapses showed a similar result, decreasing from baseline (780 msec) to the final visit (669 msec) for patients receiving modafinil and increasing from baseline (852 msec) to the final visit (1235 msec) for those receiving placebo. This resulted in a significant difference at the final visit (P=0.004) and in the change from baseline to the final visit in favor of modafinil versus placebo (P=0.019). Sleepiness levels on the Karolinska Sleepiness Scale were also significantly reduced for patients receiving modafinil (baseline mean, 7.3; final visit mean, 5.8; change, −1.5±0.2), as compared with placebo (baseline, 7.1; final visit, 6.7; change, −0.4±0.2) (P<0.001) (Fig. 2F). In general, the results of efficacy measures at the final visit were observed at the first visit after baseline and sustained throughout subsequent visits (see the Supplementary Appendix).

DATA DERIVED FROM ELECTRONIC DIARIES

There were significant effects for three of the seven efficacy variables in the patients' diaries (Table 2). As compared with placebo, 200 mg of modafinil reduced the maximum level of sleepiness during night-shift work (P<0.001 for the change from baseline vs. placebo) and the level of sleepiness during the commute home (P=0.01), and 25 percent fewer patients receiving modafinil reported having had accidents or near accidents during the commute home (P<0.001). Modafinil treatment during night shifts had no statistically significant effects on unintentional or intentional sleep episodes, mistakes, accidents or near accidents, or caffeine consumption (Table 2). During days following nights off, there were no significant differences in caffeine use and sleep efficiency between the modafinil

Table 1. Baseline Characteristics, Test Scores, and Severity of Sleepiness among Patients with Shift-Work Sleep Disorder Treated with Modafinil or Placebo.*

Characteristic	Placebo (N=108)	Modafinil (N=96)	P Value
Mean age — yr	38.8±9.1	37.5±9.2	0.31
Sex — no. (%)			0.81
Male	67 (62)	58 (60)	
Female	41 (38)	38 (40)	
Race or ethnic background — no. (%)†			0.63
White	75 (69)	62 (65)	
Black	27 (25)	25 (26)	
Asian	0	1 (1)	
Other	6 (6)	8 (8)	
Type of shift work — no. (%)			0.35
Permanent night shift	95 (88)	89 (93)	
Rotating shift	13 (12)	7 (7)	
Night shifts worked per mo — no. (%)			0.02
5–10	11 (10)	2 (2)	
>10	97 (90)	94 (98)	
Mean sleep latency — min‡	2.0±1.8	2.1±1.5	0.89
No. of lapses of attention§	24.3±26.4	22.5±23.0	0.69
Patient-estimated sleepiness¶	7.1±1.2	7.3±1.0	0.43
Sleepiness severity — no. (%)			0.91
Moderately ill	53 (49)	49 (51)	
Markedly ill	34 (31)	29 (30)	
Severely ill	17 (16)	16 (17)	
Among the most extremely ill	4 (4)	2 (2)	
Daytime polysomnographic measures‖			
Sleep latency — min	8.0±10.9	7.8±10.4	0.91
Sleep efficiency — %**	74.1±12.6	73.7±11.7	0.83

* Plus–minus values are means ±SD.
† Race or ethnic background was self-reported.
‡ A mean sleep latency of less than 5 minutes indicates a pathological level of daytime sleepiness. In control groups, sleep latencies for adult volunteers without sleepiness occur in the range of 10 to 20 minutes.[13] Multiple Sleep Latency Test data were available for 96 patients receiving placebo and for 86 patients receiving modafinil at baseline.
§ For the Psychomotor Vigilance Test, patients were instructed to respond to randomly occurring visual stimuli appearing in the window of a portable Psychomotor Vigilance Test-192 device (Ambulatory Monitoring) by pushing a button as quickly as possible. Reaction times were collected from four 20-minute sessions conducted at 2-hour intervals. Psychomotor Vigilance Test data were available for 64 patients receiving placebo and for 60 patients receiving modafinil at baseline.
¶ Ratings on the Karolinska Sleepiness Scale range from 1 (very alert) to 9 (very sleepy, great effort to keep awake, fighting sleep). Patients rated their level of sleepiness at the time of testing, not retrospectively. Karolinska Sleepiness Scale data were available for 95 patients receiving placebo and for 85 patients receiving modafinil at baseline.
‖ Polysomnographic data were available for 78 patients receiving placebo and for 72 patients receiving modafinil at baseline.
** Sleep efficiency was calculated as the sleep duration divided by the time spent in bed multiplied by 100.

group and the placebo group (Table 2). The use of sleeping pills was not specifically monitored, although concomitant use of medications was queried at each visit. One of 96 patients in the modafinil group reported the use of a prescription hypnotic agent, whereas none of the 108 patients in the placebo group did. Five of the 96 patients in the modafinil group reported the use of over-the-counter sleep aids versus 1 of the 108 patients in the placebo group (P=0.102).

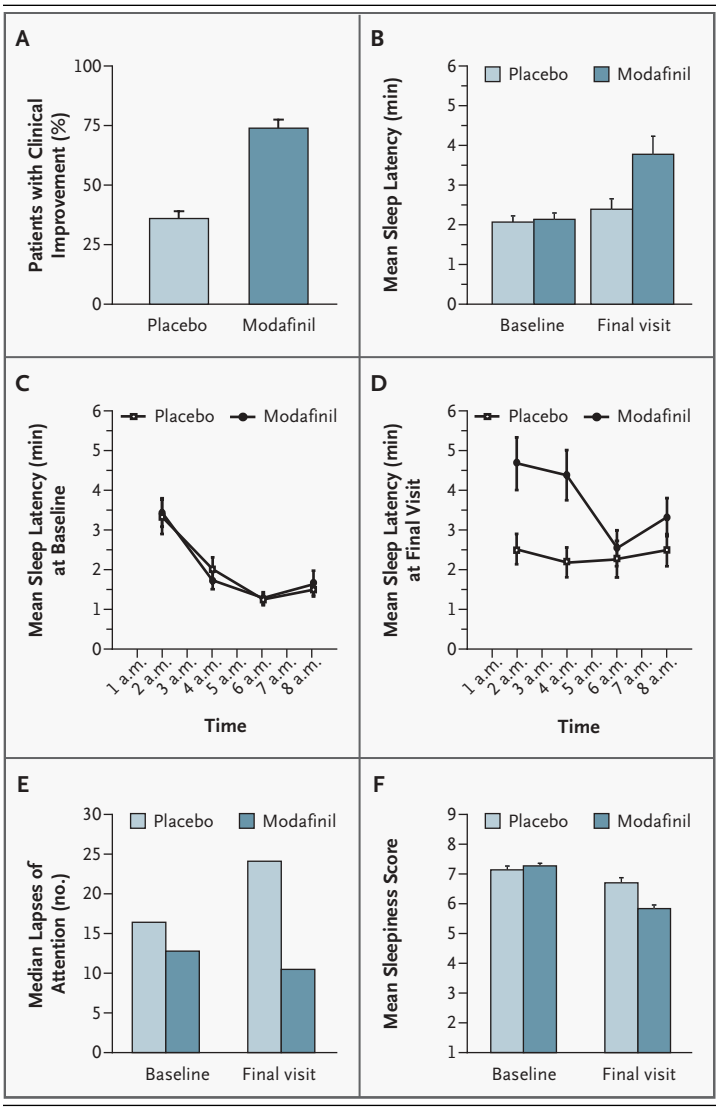


Figure 2. Efficacy Measures Used to Assess the Effects of Modafinil versus Placebo in Patients with Shift-Work Sleep Disorder.

Panel A shows the percentage of patients receiving placebo (104 patients) and modafinil (89 patients) whose symptoms were rated as clinically improved during night shifts (including the commute to and from work) on the basis of the results of the Clinical Global Impression of Change test at the final visit (P<0.001). Patients had to have undergone a baseline assessment and at least one assessment after baseline in order to be included in the analysis. T bars represent 1 SEM. In Panel B, the mean (±SEM) sleep latency, as measured by the Multiple Sleep Latency Test, during the night shift for the placebo group (96 patients at both the baseline visit and the final visit) was 2.04±0.2 minutes at baseline and 2.37±0.3 minutes at the final visit (P=0.24 for the within-treatment comparison). For the modafinil group (86 patients at both the baseline visit and the final visit), the overall mean sleep latency was 2.07±0.2 at baseline and 3.77±0.5 at the final visit (P<0.001 for the within-treatment comparison). The difference in the change in score on the Multiple Sleep Latency Test from baseline to the final visit for modafinil versus placebo was statistically significant (P=0.002). Panel C and Panel D show the mean sleep latency values at each Multiple Sleep Latency Test from 2 a.m. to 8 a.m. for 96 patients receiving placebo and 86 patients receiving modafinil during the baseline and final laboratory night shift, respectively. Patients had to have undergone a baseline assessment and at least one assessment after baseline in order to be included in the analysis. The difference in change from baseline to the final visit for modafinil versus placebo was statistically significant at 2 a.m. (P=0.02) and at 4 a.m. (P<0.001). In Panel E, the median number of lapses of attention in performance of the Psychomotor Vigilance Test for the placebo group (baseline, 64 patients; final visit, 69 patients) per 20-minute test bout was 16.13 at baseline and 23.75 at the final visit (median change from baseline, +3.75; P=0.008). For the modafinil group (baseline, 60 patients; final visit, 66 patients), the median number of lapses of attention was 12.50 at baseline and 10.25 at the final visit (median change from baseline, −2.63; P=0.012). The modafinil group and placebo group did not differ significantly at baseline (P=0.8) but did at the final visit (P=0.005). The difference in change from baseline to the final visit for modafinil versus placebo was statistically significant (P<0.001). In Panel F, the mean sleepiness rating on the Karolinska Sleepiness Scale for the placebo group (baseline, 95 patients; final visit, 97 patients) was 7.1±0.1 at baseline and 6.7±0.2 at the final visit (P=0.01 for the within-treatment comparison). For the modafinil group (baseline, 85 patients; final visit, 86 patients), the overall mean sleepiness score was 7.3±0.1 at baseline and 5.8±0.2 at the final visit (P<0.001 for the within-treatment comparison). The difference in change from baseline to the final visit for modafinil versus placebo was statistically significant (P<0.001).

SAFETY OUTCOMES

Headache was the most common adverse event associated with treatment in both groups (Table 3). No serious adverse events were reported for patients in the modafinil group. More patients in the modafinil group than in the placebo group had insomnia (6 percent vs. 0 percent, respectively; P=0.01). Adverse events that were not serious but resulted in the inability to carry out usual activities were defined as severe. Eleven patients reported such events (six in the modafinil group and five in the placebo group) (see Table 3 of the Supplementary Appendix). No clinically meaningful differences in vital signs, clinical laboratory measures, phys-

Table 2. Variables for Patients with Shift-Work Sleep Disorder Treated with Modafinil or Placebo, as Derived from Diaries.*

Variable	Placebo (N=108)			Modafinil (N=96)			P Value
	Baseline	After Baseline	Change	Baseline	After Baseline	Change	
During night shift							
Maximum level of sleepiness — score†	7.4±1.0	6.6±1.3	−0.9±1.0	7.3±0.9	5.4±1.5	−1.9±1.4	<0.001‡
No. of unintentional sleep episodes†	1.2±1.3	0.6±0.7	−0.6±1.0	1.0±1.1	0.2±0.4	−0.8±0.9	0.20
No. of intentional sleep episodes†	0.5±0.8	0.4±0.5	−0.1±0.5	0.4±0.5	0.2±0.4	−0.2±0.4	0.13
No. of caffeinated drinks consumed†	1.3±1.1	1.1±0.9		1.3±1.2	1.0±1.0		0.10
Patients reporting mistakes, accidents, or near accidents — no. (%)§		59 (55)			46 (48)		0.34
During the commute home							
Level of sleepiness — score†	5.9±1.8	5.4±1.7	−0.6±1.2	5.5±1.8	4.4±1.6	−1.1±1.5	0.012‡
Patients reporting unintentional sleep episodes — no. (%)§		47 (44)			34 (35)		0.24
Patients reporting accidents or near accidents — no. (%)§		58 (54)			28 (29)		<0.001¶
During days after night shift							
No. of caffeinated drinks consumed‖**	1.0±1.3	0.6±0.7	−0.4±1.0	0.9±1.1	0.7±0.8	−0.2±1.0	0.61
Sleep efficiency — %**††	78.0±20.7	87.5±14.1	9.5±18.3	80.3±19.9	87.5 ±14.4	7.3±18.5	0.55

* Plus–minus values are means ±SD. Patients recorded responses in electronic diaries on actual work nights. Sleepiness scores were obtained with the use of the Karolinska Sleepiness Scale. Analysis includes patients with baseline values and values after baseline. For each patient, baseline values and values after baseline are average values calculated before and after the start of double-blind treatment.
† Data were available for 84 patients receiving placebo and for 79 patients receiving modafinil.
‡ P value is for the change from baseline for modafinil versus placebo.
§ Values are for the number of patients with a value after baseline. Patients were counted once.
¶ P value is for modafinil versus placebo.
‖ Data were available for 85 patients receiving placebo and for 78 patients receiving modafinil.
** The time interval was from the end of the night shift until 60 minutes after waking up from the last sleep episode.
†† Data were available for 84 patients receiving placebo and for 78 patients receiving modafinil. Sleep efficiency was calculated as the sleep duration divided by the time spent in bed multiplied by 100 so that scores could range from 0 to 100 percent.

ical examinations, or electrocardiographic findings were observed between treatment groups.

OTHER ASSESSMENTS

There were no significant differences between modafinil and placebo with respect to any measurement of daytime sleep, including sleep duration, latency, and efficiency and the proportion and distribution of sleep stages (Table 4). Patients receiving modafinil did not differ significantly from those receiving placebo in the mean change in melatonin phase from baseline to the final visit (0.4 hour and −0.1 hour, respectively).

DISCUSSION

Circadian-rhythm sleep disorders have long been recognized as important disruptions of sleep–wake behaviors in a subgroup of people who are substantially more impaired than others with similar schedules.[19-21] Such differential vulnerability regarding cognitive impairment that is induced by extended wakefulness at night is a stable characteristic of these persons.[22,23] Estimates of the proportion of night-shift workers who meet the clinical criteria of both excessive sleepiness and daytime insomnia that we used to diagnose shift-work sleep

Table 3. Adverse Events in Patients Diagnosed with Shift-Work Sleep Disorder Treated with Modafinil or Placebo.*

Adverse Event	Placebo (N=108)	Modafinil (N=96)
	no. (%)	
Headache	21 (19)	25 (26)
Infection	11 (10)	6 (6)
Nausea	3 (3)	9 (9)
Rhinitis	7 (6)	3 (3)
Accidental injury	9 (8)	6 (6)
Abdominal pain	2 (2)	6 (6)
Nervousness	1 (<1)	6 (6)
Insomnia	0 (0)	6 (6)†
Dry mouth	4 (4)	5 (5)
Tooth disorder	1 (<1)	5 (5)

* Patients could report more than one event. Adverse events that were associated with treatment included untoward medical occurrences of all causes that developed or worsened in severity during the course of double-blind treatment. The adverse events that are listed are those that occurred in 5 percent or more of patients in either the modafinil group or the placebo group.
† P=0.01 for the comparison between modafinil and placebo.

disorder range from 5 to 10 percent.[4-6,24] The burden of illness in persons with shift-work sleep disorder is substantial, as compared with shift workers without the disorder.[6,7,24,25] Thus, shift-work sleep disorder is more than simply being tired on the night shift.

In this randomized, placebo-controlled study — the first such trial in the investigation of shift-work sleep disorder — improvements in alertness and performance were found with 200 mg of modafinil in measures of sleep latency, clinical-impression rating, sustained-attention performance, and patient-estimated sleepiness. Consistent with this profile were reductions in patient-estimated sleepiness on work nights and during the morning commute home. Despite these benefits, patients treated with modafinil continued to have high levels of sleepiness and impaired performance at night.

Although patients receiving 200 mg of modafinil continued to have lapses in performance on the Psychomotor Vigilance Test,[26-29] there were twice as many lapses per night at the final visit in the placebo group as there were in the modafinil group, and the mean duration of these lapses in the modafinil group was nearly twice as long as that in the placebo group. It is likely that the effects of modafinil on sustained-attention performance derive, at least in part, from its effects on reducing the instability of wakefulness caused by brief episodes of sleep intruding into waking performance.[15,30,31] Although lapses of attention were reduced, they remained at a high level in the treatment group. This suggests that although modafinil improves the measured levels of performance, it is far from what is needed for these patients to function at a normal level.

The results of this study also suggest that 200 mg of modafinil does not affect circadian adaptation to night-work schedules. Thus, the ability of modafinil to treat symptoms of excessive sleepiness in patients diagnosed with shift-work sleep disorder is a result of an improvement in wakefulness during the nocturnal work shift, similar to the improved alertness shown in other disorders of sleep and wakefulness,[8-11] and not an improvement in the alignment between internal circadian rhythms and the work–sleep schedule.

Several considerations limit the interpretation and applicability of the findings. There remains a need for validated criteria and clinical instruments for assessing excessive sleepiness in shift-work sleep disorder. Although the Multiple Sleep Latency Test is sensitive to changes in sleepiness during nighttime hours[32,33] and is recommended for assessing sleepiness at night in this population,[5] it has not been specifically validated as a clinical instrument for measuring nighttime sleepiness, particularly in the absence of objectively monitored sleep in the laboratory on the day before testing. As recommended in the literature,[17] we therefore used a validated performance measure — the Psychomotor Vigilance Test — to assess alertness at night, the results of which were consistent with the nighttime data from the Multiple Sleep Latency Test. Because patients worked in a variety of industries, actual work performance was not evaluated. We do not know how the laboratory sleep and performance variables that were used in the study may apply to actual on-the-job performance, although we do show concordance of results for measures of alertness, performance on the Psychomotor Vigilance Test, and diary data that collectively suggest a positive effect on personal and public safety. Although the study was open to both permanent and rotating night-shift workers with shift-work sleep disorder, the vast majority of study participants (90 percent) were permanent night-shift workers. Thus, it is not appropriate to generalize the findings of the study to patients who work on other types of shifts that include nighttime hours. The patients who met the criteria of having shift-work sleep dis-

Table 4. Daytime Polysomnographic Measures for Patients with Shift-Work Sleep Disorder Treated with Modafinil or Placebo.*

Polysomnographic Variable	Placebo (N=78)			Modafinil (N=72)			P Value
	Baseline	Final	Change	Baseline	Final	Change	
Time in bed (min)	479.7±2.3	478.0±9.1	−1.7±9.3	479.2±5.3	476.2±2.8	−3.0±27.4	0.38
Time awake (min)	118.9±59.9	110.1±75.0	−8.8±79.0	120.5±570	108.8±71.5	−11.7±75.0	0.60
Time asleep (min)	355.4±60.4	360.0±79.5	4.6±81.2	352.9±55.6	354.3±91.3	1.4±88.8	0.82
Sleep latency (min)	8.0±10.9	9.3±10.1	1.3±12.4	7.8±10.4	10.7±17.1	2.9±17.9	0.26
Sleep efficiency (%)	74.1±12.6	75.3±16.4	1.2±17.0	73.7±11.7	75.0±16.2	1.4±16.7	0.85
No. of patients who required >30 sec to awaken	19.1±10.4	16.9±10.9	−2.2±11.7	21.1±10.6	18.4±11.5	−2.7±12.1	0.85
No. of patients who required >2 attempts to awaken	7.2±4.1	6.1±3.8	−1.2±4.2	7.6±3.5	6.3±4.4	−1.3±4.2	0.70
Stage of sleep (%)							
Non–rapid-eye-movement sleep†							
Stage 1	11.9±6.4	11.5±5.7	−0.3±6.6	12.5±6.4	13.9±8.2	1.4±8.6	0.16
Stage 2	54.7±11.6	53.2±11.3	−1.5±11.2	53.1±10.0	52.7±11.0	−0.4±11.9	0.72
Stages 3 and 4	13.2±11.0	12.8±10.6	−0.4±9.6	14.2±9.6	12.7±10.9	−1.4±8.8	0.65
Rapid-eye-movement sleep	20.2±6.2	22.4±8.1	2.3±8.8	20.2±6.5	20.7±6.5	0.5±8.9	0.20

* Plus–minus values are means ±SD. Analysis includes patients with values for both baseline and final visits.

† Stage 1 is a transitional state between waking and sleeping (light sleep); stage 2 is an intermediate stage of sleep that normally accounts for half the total sleep time; and stages 3 and 4 are deep, slow-wave sleep characterized by high-amplitude delta waves on electroencephalography.

order are only a subgroup of shift workers,[7] a fact that limits the applicability of the findings to the broader shift-work population in whom the safety and efficacy of modafinil have not been evaluated. Our study was 12 weeks in duration; the effects of long-term modafinil use in this population are unknown.

In summary, we found that patients with shift-work sleep disorder had excessive sleepiness during night work, similar to that seen during the day in patients with narcolepsy. Even after treatment with modafinil, these patients still showed evidence of excessive sleepiness during the night shift. Although modafinil did not restore sleepiness to normal daytime levels, treatment was associated with improvements in symptoms of sleepiness, as well as objective measures of sleep propensity and performance. Modafinil is of some value in the clinical management of sleepiness associated with shift-work sleep disorder. Concern remains that even with treatment with 200 mg of modafinil, the excessive sleepiness observed in this underrecognized population requires the development of yet more effective therapies.

Supported by Cephalon, Frazer, Pa.

Drs. Czeisler and Dinges report having received consulting fees and speaker's fees from Cephalon and having received clinical trial research contracts, investigator-initiated research grants, and unrestricted research and education funds from Cephalon. Dr. Czeisler serves as the incumbent of an endowed professorship provided to Harvard University by Cephalon. Drs. Walsh, Roth, and Wright report having received consulting fees and speaker's fees from Cephalon and having received clinical trial research contracts and investigator-initiated research grants from Cephalon. Drs. Hughes, Kingsbury, Arora, and Niebler are employees of and report having an equity interest in Cephalon. Dr. Schwartz reports having received consulting fees and clinical trial research contracts from Cephalon and serving on the speakers bureau for Cephalon.

We are indebted to Mark Riotto, Angela Kaya, and Richard Malone for assistance in the preparation of the manuscript.

APPENDIX

In addition to the authors, members of the U.S. Modafinil in Shift Work Sleep Disorder Study Group are as follows: J.E. Black, Stanford, Calif.; R.K. Bogan, Columbia, S.C.; M.H. Bonnet, Dayton, Ohio; M.A. Carskadon, Providence, R.I.; J.S. Cook, Danville, Ind.; B.C. Corser, Cincinnati; M.K. Erman, La Jolla, Calif.; N.T. Feldman, St. Petersburg, Fla.; J.M. Ferguson, Salt Lake City; Y. Furman, Los Angeles; S.C. Hardy, Charlotte, N.C.; J.R. Harsh, Hattiesburg, Miss.; M. Hirshkowitz, Houston; S.G. Hull, Overland Park, Kans.; V.K. Mahajan, Toledo, Ohio; G.V. Pegram, Jr., Birmingham, Ala.; J. Pinto, Las Vegas; G.S. Richardson, Detroit; R. Rosenberg, Atlanta; M.H. Rosenthal, San Diego, Calif.; M.H. Schmidt, Dublin, Ohio; P.K. Schweitzer, Chesterfield, Mo.; D. Seiden, Pembroke Pines, Fla.; D.R. Wagner, White Plains, N.Y.; C.C. Wells, Jr., Macon, Ga.; J.K. Wyatt, Chicago; and G.K. Zammit, New York.

REFERENCES

1. Beers TM. Flexible schedules and shift work: replacing the '9-to-5' workday? Monthly Labor Rev 2000;123:33-40.
2. Åkerstedt T. Shift work and disturbed sleep/wakefulness. Occup Med (Lond) 2003; 53:89-94.
3. Ohayon MM, Lemoine P, Arnaud-Briant V, Dreyfus M. Prevalence and consequences of sleep disorders in a shift worker population. J Psychosom Res 2002;53:577-83.
4. Diagnostic and statistical manual of mental disorders, 4th ed. Washington, D.C.: American Psychiatric Association, 2000.
5. ICSD — international classification of sleep disorders: diagnostic and coding manual. 2nd ed. Westchester, Ill.: American Academy of Sleep Medicine, 2005.
6. Drake CL, Roehrs T, Richardson G, Walsh JK, Roth T. Shift work sleep disorder: prevalence and consequences beyond that of symptomatic day workers. Sleep 2004;27: 1453-62.
7. Knutsson A. Health disorders of shift workers. Occup Med (Lond) 2003;53:103-8.
8. US Modafinil in Narcolepsy Multicenter Study Group. Randomized trial of modafinil as a treatment for the excessive daytime somnolence of narcolepsy. Neurology 2000; 54:1166-75.
9. *Idem.* Randomized trial of modafinil for the treatment of pathological somnolence in narcolepsy. Ann Neurol 1998;43:88-97.
10. Pack AI, Black JE, Schwartz JR, Matheson JK. Modafinil as adjunct therapy for daytime sleepiness in obstructive sleep apnea. Am J Respir Crit Care Med 2001;164:1675-81.
11. Black JE, Hirshkowitz M. Modafinil for treatment of residual excessive sleepiness in nasal continuous positive airway pressure-treated obstructive sleep apnea/hypopnea syndrome. Sleep 2005;28:464-71.
12. Guy W. ECDEU assessment manual for psychopharmacology. Rockville, Md.: National Institute of Mental Health, 1976.
13. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep 1986;9:519-24.
14. Carskadon MA, Dement WC. Daytime sleepiness: quantification of a behavioral state. Neurosci Biobehav Rev 1987;11: 307-17.
15. Dorrian J, Rogers NL, Dinges DF. Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In: Kushida CA, ed. Sleep deprivation: clinical issues, pharmacology and sleep loss effects. New York: Marcel Dekker, 2005:39-70.
16. Åkerstedt T, Torsvall L, Gillberg M. Sleepiness and shift work: field studies. Sleep 1982;5:Suppl 2:S95-S106.
17. Mitler MM, Carskadon MA, Hirshkowitz M. Evaluating sleepiness. In: Kryger MH, Roth T, Dement WC, eds. Principles and practice of sleep medicine. 3rd ed. Philadelphia: W.B. Saunders, 2000:1251-7.
18. Wright KP Jr, Hughes RJ, Kronauer RE, Dijk DJ, Czeisler CA. Intrinsic near-24-h pacemaker period determines limits of circadian entrainment to a weak synchronizer in humans. Proc Natl Acad Sci U S A 2001; 98:14027-32.
19. Richardson GS, Malin HV. Circadian rhythm sleep disorders: pathophysiology and treatment. J Clin Neurophysiol 1996; 13:17-31.
20. Zisapel N. Circadian rhythm sleep disorders: pathophysiology and potential approaches to management. CNS Drugs 2001; 15:311-28.
21. Mitler MM, Miller JC, Lipsitz JJ, Walsh JK, Wylie CD. The sleep of long-haul truck drivers. N Engl J Med 1997;337:755-61.
22. Leproult R, Colecchia EF, Berardi AM, Stickgold R, Kosslyn SM, Van Cauter E. Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. Am J Physiol Regul Integr Comp Physiol 2003;284:R280-R290.
23. Van Dongen HPA, Baynard MD, Maislin G, Dinges DF. Systemic inter-individual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. Sleep 2004;27:423-33.
24. Åkerstedt T. Shift work and sleep disorders. Sleep 2005;28:9-11.
25. Axelsson J, Åkerstedt T, Kecklund G, Lowden A. Tolerance to shift work — how does it relate to sleep and wakefulness? Int Arch Occup Environ Health 2004;77:121-9.
26. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003;26:117-26. [Erratum, Sleep 2004;27:600.]
27. Dinges DF, Pack F, Williams K, et al. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep 1997;20:267-77.
28. Powell NB, Riley RW, Schechtman KB, Blumen MB, Dinges DF, Guilleminault C. A comparative model: reaction time performance in sleep-disordered breathing versus alcohol-impaired controls. Laryngoscope 1999;109:1648-54.
29. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. J Sleep Res 2003;12:1-12.
30. Doran SM, Van Dongen HP, Dinges DF. Sustained attention performance during sleep deprivation: evidence of state instability. Arch Ital Biol 2001;139:253-67.
31. Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. Trends Neurosci 2001;24: 726-31.
32. Walsh JK, Muehlbach MJ, Humm TM, Dickins QS, Sugerman JL, Schweitzer PK. Effect of caffeine on physiological sleep tendency and ability to sustain wakefulness at night. Psychopharmacology (Berl) 1990; 101:271-3.
33. Muehlbach MJ, Walsh JK. The effects of caffeine on simulated night-shift work and subsequent daytime sleep. Sleep 1995;18: 22-9.

CLINICAL TRIAL REGISTRATION

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

CORRECTION

Modafinil for Excessive Sleepiness Associated with Shift-Work Sleep Disorder

Modafinil for Excessive Sleepiness Associated with Shift-Work Sleep Disorder . On page 484, in the left-hand column, lines 23 through 26 should have read, ``. . . and the mean duration of these lapses in the placebo group was nearly twice as long as that in the modafinil group,'' rather than ``. . . and the mean duration of these lapses in the modafinil group was nearly twice as long as that in the placebo group,'' as printed.

N Engl J Med 2005;353:1078-a

REVIEW ARTICLE

DRUG THERAPY

Adherence to Medication

Lars Osterberg, M.D., and Terrence Blaschke, M.D.

Drugs don't work in patients who don't take them.

— C. Everett Koop, M.D.

From the General Medicine Division, Veterans Affairs Palo Alto Health Care System, Palo Alto (L.O.); and the Division of Clinical Pharmacology, Stanford University Medical Center, Stanford (T.B.) — both in California. Address reprint requests to Dr. Osterberg at the VA Palo Alto Health Care System, 3801 Miranda Ave., Palo Alto, CA 94304, or at larso@stanford.edu.

N Engl J Med 2005;353:487-97.

Adherence to (or compliance with) a medication regimen is generally defined as the extent to which patients take medications as prescribed by their health care providers. The word "adherence" is preferred by many health care providers, because "compliance" suggests that the patient is passively following the doctor's orders and that the treatment plan is not based on a therapeutic alliance or contract established between the patient and the physician. Both terms are imperfect and uninformative descriptions of medication-taking behavior. Unfortunately, applying these terms to patients who do not consume every pill at the desired time can stigmatize these patients in their future relationships with health care providers. The language used to describe how patients take their medications needs to be reassessed, but these terms are still commonly used.[1] Regardless of which word is preferred, it is clear that the full benefit of the many effective medications that are available will be achieved only if patients follow prescribed treatment regimens reasonably closely.

Rates of adherence for individual patients are usually reported as the percentage of the prescribed doses of the medication actually taken by the patient over a specified period. Some investigators have further refined the definition of adherence to include data on dose taking (taking the prescribed number of pills each day) and the timing of doses (taking pills within a prescribed period). Adherence rates are typically higher among patients with acute conditions, as compared with those with chronic conditions; persistence among patients with chronic conditions is disappointingly low, dropping most dramatically after the first six months of therapy.[2-4] For example, approximately half of patients receiving hydroxymethylglutaryl–coenzyme A reductase inhibitor therapy will discontinue their medication within six months of starting the therapy.[5]

The average rates of adherence in clinical trials can be remarkably high, owing to the attention study patients receive and to selection of the patients, yet even clinical trials report average adherence rates of only 43 to 78 percent among patients receiving treatment for chronic conditions.[3,6,7] There is no consensual standard for what constitutes adequate adherence. Some trials consider rates of greater than 80 percent to be acceptable, whereas others consider rates of greater than 95 percent to be mandatory for adequate adherence, particularly among patients with serious conditions such as infection with the human immunodeficiency virus (HIV). Although data on adherence are often reported as dichotomous variables (adherence vs. nonadherence), adherence can vary along a continuum from 0 to more than 100 percent, since patients sometimes take more than the prescribed amount of medication.[8-10]

The ability of physicians to recognize nonadherence is poor, and interventions to improve adherence have had mixed results. Furthermore, successful interventions gener-

ally are substantially complex and costly.[11-14] Poor adherence to medication regimens accounts for substantial worsening of disease, death, and increased health care costs in the United States.[15-19] Of all medication-related hospital admissions in the United States, 33 to 69 percent are due to poor medication adherence, with a resultant cost of approximately $100 billion a year.[15,17,20,21] Participants in clinical trials who do not follow medication regimens or placebo regimens have a poorer prognosis than subjects in the respective groups who do.[22-24] Adherence to medication and placebo regimens, therefore, both predict better outcomes, and collecting adherence data from subjects is now considered an essential part of clinical trials.[25,26] Given the magnitude and importance of poor adherence to medication regimens, the World Health Organization has published an evidence-based guide for clinicians, health care managers, and policymakers to improve strategies of medication adherence.[27]

MEASURES OF ADHERENCE

Adherence to medication regimens has been monitored since the time of Hippocrates, when the effects of various potions were recorded with notations of whether the patient had taken them or not. Even today, patients' self-reports can simply and effectively measure adherence.[28,29] The methods available for measuring adherence can be broken down into direct and indirect methods of measurement (Table 1). Each method has advantages and disadvantages, and no method is considered the gold standard.[30,31]

Directly observed therapy, measurement of concentrations of a drug or its metabolite in blood or urine, and detection or measurement in blood of a biologic marker added to the drug formulation are examples of direct methods of measures of adherence. Direct approaches are expensive, burdensome to the health care provider, and susceptible to distortion by the patient. However, for some drugs, measuring these levels is a good and commonly used means of assessing adherence. For instance, the serum concentration of antiepileptic drugs such as phenytoin or valproic acid will probably reflect adherence to regimens with these medications, and subtherapeutic levels will probably reflect poor adherence or suboptimal dose strengths.

Indirect methods of measurement of adherence include asking the patient about how easy it is for him or her to take prescribed medication, assessing clinical response, performing pill counts, ascertaining rates of refilling prescriptions, collecting patient questionnaires, using electronic medication monitors, measuring physiologic markers, asking the patient to keep a medication diary, and assessing children's adherence by asking the help of a caregiver, school nurse, or teacher. Questioning the patient (or using a questionnaire), patient diaries, and assessment of clinical response are all methods that are relatively easy to use, but questioning the patient can be susceptible to misrepresentation and tends to result in the health care provider's overestimating the patient's adherence.

The use of a patient's clinical response as a measure is confounded by many factors other than adherence to a medication regimen that can account for clinical outcome. The most common method used to measure adherence, other than patient questioning, has been pill counts (i.e., counting the number of pills that remain in the patient's medication bottles or vials). Although the simplicity and empiric nature of this method are attractive to many investigators, the method is subject to many problems, because patients can switch medicines between bottles and may discard pills before visits in order to appear to be following the regimen. For these reasons, pill counts should not be assumed to be a good measure of adherence.[8,9,32] In addition, this method provides no information on other aspects of taking medications, such as dose timing and drug holidays (i.e., omission of medication on three or more sequential days), both of which may be important in determining clinical outcomes.

Rates of refilling prescriptions are an accurate measure of overall adherence in a closed pharmacy system (e.g., health maintenance organizations, the Department of Veterans Affairs Health Care System, or countries with universal drug coverage), provided that the refills are measured at several points in time.[33-35] A medical system that uses electronic medical records and a closed pharmacy can provide the clinician or research scientist with readily available objective information on rates of refilling prescriptions that can be used to assess whether a patient is adhering to the regimen and to corroborate the patient's responses to direct questions or on questionnaires.

Electronic monitors capable of recording and stamping the time of opening bottles, dispensing drops (as in the case of glaucoma), or activating a

Table 1. Methods of Measuring Adherence.

Test	Advantages	Disadvantages
Direct methods		
Directly observed therapy	Most accurate	Patients can hide pills in the mouth and then discard them; impractical for routine use
Measurement of the level of medicine or metabolite in blood	Objective	Variations in metabolism and "white-coat adherence" can give a false impression of adherence; expensive
Measurement of the biologic marker in blood	Objective; in clinical trials, can also be used to measure placebo	Requires expensive quantitative assays and collection of bodily fluids
Indirect methods		
Patient questionnaires, patient self-reports	Simple; inexpensive; the most useful method in the clinical setting	Susceptible to error with increases in time between visits; results are easily distorted by the patient
Pill counts	Objective, quantifiable, and easy to perform	Data easily altered by the patient (e.g., pill dumping)
Rates of prescription refills	Objective; easy to obtain data	A prescription refill is not equivalent to ingestion of medication; requires a closed pharmacy system
Assessment of the patient's clinical response	Simple; generally easy to perform	Factors other than medication adherence can affect clinical response
Electronic medication monitors	Precise; results are easily quantified; tracks patterns of taking medication	Expensive; requires return visits and downloading data from medication vials
Measurement of physiologic markers (e.g., heart rate in patients taking beta-blockers)	Often easy to perform	Marker may be absent for other reasons (e.g., increased metabolism, poor absorption, lack of response)
Patient diaries	Help to correct for poor recall	Easily altered by the patient
When the patient is a child, questionnaire for caregiver or teacher	Simple; objective	Susceptible to distortion

canister (as in the case of asthma) on multiple occasions have been used for approximately 30 years.[32,36-38] Rather than providing weekly or monthly averages, these devices provide precise and detailed insights into patients' behavior in taking medication, but they are still indirect methods of measuring adherence; they do not document whether the patient actually ingested the correct drug or correct dose. Patients may open a container and not take the medication, take the wrong amount of medication, or invalidate the data by placing the medication into another container or taking multiple doses out of the container at the same time. The cost of electronic monitoring is not covered by insurance, and thus these devices are not in routine use. However, this approach provides the most accurate and valuable data on adherence in difficult clinical situations and in the setting of clinical trials and adherence research[10,39] and has advanced our knowledge of medication-taking behavior.[40] Although certain methods of measuring adherence may be preferred in specific clinical or research settings, a combination of measures maximizes accuracy.[10,41,42]

EPIDEMIOLOGY OF MEDICATION-TAKING BEHAVIOR

Electronic medication-monitoring devices have provided very detailed information about the patterns of medication-taking behavior. Most deviations in medication taking occur as omissions of doses (rather than additions) or delays in the timing of doses.[11,43] Patients commonly improve their medication-taking behavior in the 5 days before and after an appointment with the health care provider, as compared with 30 days after, in a phenomenon known as "white-coat adherence."[44,45] Stud-

ies using these monitors have shown six general patterns of taking medication among patients treated for chronic illnesses who continue to take their medications. Approximately one sixth come close to perfect adherence to a regimen; one sixth take nearly all doses, but with some timing irregularity; one sixth miss an occasional single day's dose and have some timing inconsistency; one sixth take drug holidays three to four times a year, with occasional omissions of doses; one sixth have a drug holiday monthly or more often, with frequent omissions of doses; and one sixth take few or no doses while giving the impression of good adherence.[40,46]

Simple dosing (one pill, once daily) helps to maximize adherence, particularly when combined with frequent reinforcing visits, despite the fact that 10 to 40 percent of patients taking these simple regimens continue to have imperfect dosing.[47,48] In a large systematic review of 76 trials in which electronic monitors were used, Claxton and colleagues[7] found that adherence was inversely proportional to frequency of dose (Fig. 1), and patients taking medication on a schedule of four times daily achieved average adherence rates of about 50 percent (range, 31 to 71 percent).

IDENTIFYING POOR ADHERENCE

Indicators of poor adherence to a medication regimen are a useful resource for physicians to help identify patients who are most in need of interventions to improve adherence.[5,49,50] Table 2 lists major predictors associated with poor adherence. Race, sex, and socioeconomic status have not been consistently associated with levels of adherence.[59,61] When these predictors, listed in Table 2, are present, physicians should have a heightened awareness of the possibility of poor adherence, but even patients in whom these indicators are absent miss taking medications as prescribed. Thus, poor adherence should always be considered when a patient's condition is not responding to therapy.

The simplest and most practical suggestion for physicians is to ask patients nonjudgmentally how often they miss doses. Patients generally want to please their physicians and will often say what they think their doctor wants to hear. It can be reassuring to the patient when the physician tells them, "I know it must be difficult to take all your medications regularly. How often do you miss taking them?" This approach makes most patients feel comfortable in telling the truth and facilitates the identification of poor adherence. A patient who admits to poor adherence is generally being candid.[29,62] Patients should also be asked whether they are having any side effects of their medications, whether they know why they are taking their medications, and what the benefits of taking them are, since these questions can often expose poor adherence to a regimen.[63]

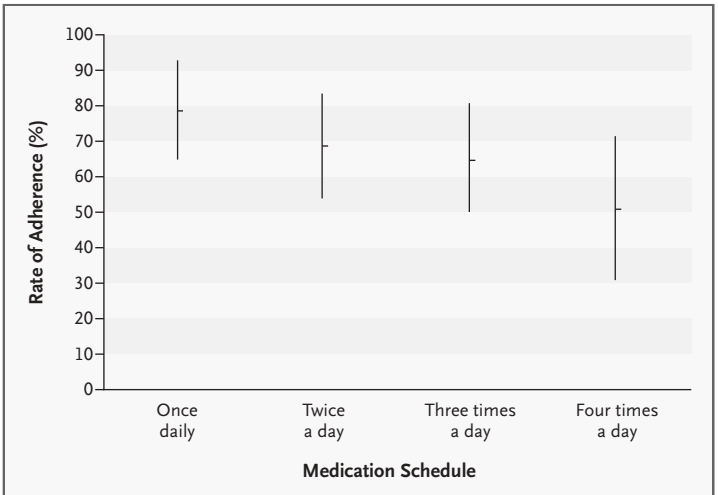


Figure 1. Adherence to Medication According to Frequency of Doses.
Vertical lines represent 1 SD on either side of the mean rate of adherence (horizontal bars). Data are from Claxton et al.[7]

BARRIERS TO ADHERENCE

Research on adherence has typically focused on the barriers patients face in taking their medications. Common barriers to adherence are under the patient's control, so that attention to them is a necessary and important step in improving adherence. In responses to a questionnaire, typical reasons cited by patients for not taking their medications included forgetfulness (30 percent), other priorities (16 percent), decision to omit doses (11 percent), lack of information (9 percent), and emotional factors (7 percent); 27 percent of the respondents did not provide a reason for poor adherence to a regimen.[64] Physicians contribute to patients' poor adherence by prescribing complex regimens, failing to explain the benefits and side effects of a medication adequately, not giving consideration to the patient's lifestyle or the cost of the medications, and having poor therapeutic relationships with their patients.[49,65-67]

More broadly, health care systems create barriers to adherence by limiting access to health care, using a restricted formulary, switching to a different formulary, and having prohibitively high costs for drugs, copayments, or both.[60,68,69] To improve the patient's ability to follow a medication regimen, all potential barriers to adherence need to be considered. An expanded view that takes into account factors under the patient's control as well as interactions between the patient and the health care provider and between the patient and the health care system will have the greatest effect on improving medication adherence (Fig. 2).[70,71]

INTERVENTIONS

Methods that can be used to improve adherence can be grouped into four general categories: patient education; improved dosing schedules; increased hours when the clinic is open (including evening hours), and therefore shorter wait times; and improved communication between physicians and patients. Educational interventions involving patients, their family members, or both can be effective in improving adherence.[72,73] Strategies to improve dosing schedules include the use of pillboxes to organize daily doses, simplifying the regimen to daily dosing, and cues to remind patients to take medications. Patients who miss appointments are often those who need the most help to improve their ability to adhere to a medication regimen; such patients will often benefit from assistance in clinic scheduling and what is called "cue-dose training" to optimize their adherence. Clinic-scheduling strategies to improve adherence include making follow-up visits convenient and efficient for the patient. Delays in seeing patients and problems with transportation and parking can undermine a patient's willingness to comply with a medication regimen and to keep follow-up appointments. Interventions that enlist ancillary health care providers such as pharmacists, behavioral specialists, and nursing staff can improve adherence.[12,74,75] Finally, enhancing communication between the physician and the patient is a key and effective strategy in boosting the patient's ability to follow a medication regimen.[11,18,76,77]

Most methods of improving adherence have involved combinations of behavioral interventions and reinforcements in addition to increasing the convenience of care, providing educational information about the patient's condition and the treatment, and other forms of supervision or attention.[12,78-80] Successful methods are complex and labor intensive, and innovative strategies will need to be developed that are practical for routine clinical use.[12] Given the many factors contributing to poor adherence to medication, a multifactorial approach is required, since a single approach will not be effective for all patients.[81,82] Table 3 lists some simple strategies for optimizing a patient's ability to follow a medication regimen.

Table 2. Major Predictors of Poor Adherence to Medication, According to Studies of Predictors.

Predictor	Study
Presence of psychological problems, particularly depression	van Servellen et al.,[51] Ammassari et al.,[52] Stilley et al.[53]
Presence of cognitive impairment	Stilley et al.,[53] Okuno et al.[54]
Treatment of asymptomatic disease	Sewitch et al.,[55]
Inadequate follow-up or discharge planning	Sewitch et al.,[55] Lacro et al.[56]
Side effects of medication	van Servellen et al.[51]
Patient's lack of belief in benefit of treatment	Okuno et al.,[54] Lacro et al.[56]
Patient's lack of insight into the illness	Lacro et al.,[56] Perkins[57]
Poor provider–patient relationship	Okuno et al.,[54] Lacro et al.[56]
Presence of barriers to care or medications	van Servellen et al.,[51] Perkins[57]
Missed appointments	van Servellen et al.,[51] Farley et al.[58]
Complexity of treatment	Ammassari et al.[52]
Cost of medication, copayment, or both	Balkrishnan,[59] Ellis et al.[60]

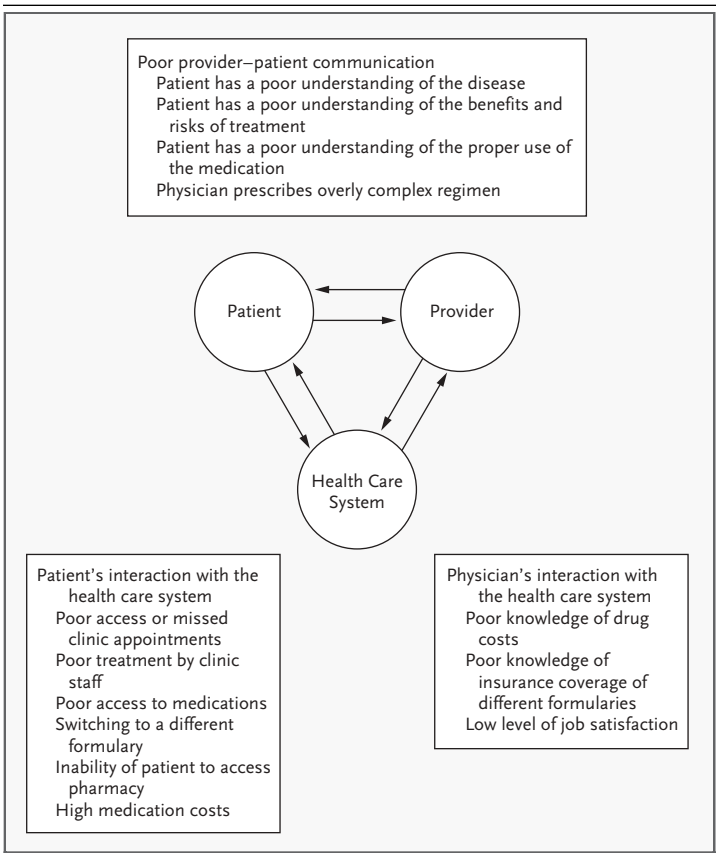


Figure 2. Barriers to Adherence.

The interactions among the patient, health care provider, and health care system depicted are those that can have a negative effect on the patient's ability to follow a medication regimen.

EXAMPLES OF CHALLENGES TO ADHERENCE

HIV INFECTION

In the treatment of patients with HIV infection or the acquired immunodeficiency syndrome, it is essential to achieve more than 95 percent adherence to highly active antiretroviral therapy (HAART) in order to suppress viral replication and avoid the emergence of resistance.[84,85] Achieving such high rates of adherence is very challenging to such patients, because their regimens include multiple, often expensive medications that have complex dosing schedules and may cause food interactions and side effects that result in poor tolerability. In addition, lifestyle factors and issues in the patient–provider relationship may make adherence difficult.[85]

Promising strategies for improving adherence to HAART that have been studied in randomized clinical trials include pharmacist-led individualized interventions, cognitive–behavioral educational interventions based on self-efficacy theory, and cue-dose training in combination with monetary reinforcement.[75,79] Cognitive–behavioral approaches have resulted in more than 90 percent of patients achieving 95 percent adherence, but these approaches require considerable resources, and adherence is typically not sustained after the intervention is withdrawn.[86,87] Federally funded trials of strategies to improve patients' ability to follow treatment regimens are ongoing, including the use of handheld devices, two-way pagers, medication vials equipped with alarms, and the enhancement of social and emotional support.[75]

HYPERTENSION

Consistent control of blood pressure requires that patients with hypertension follow medication and dietary regimens. However, antihypertensive therapy may have untoward side effects and result in little symptomatic relief, since hypertension often causes no symptoms. No matter how effectively the clinician communicates the benefits of antihypertensive therapy, patients are still ultimately responsible for taking their medications. Since adherence is enhanced when patients are involved in medical decisions about their care and in monitoring their care, the traditional model of the authoritarian provider should be replaced by the more useful dynamic of shared decision making by the health care provider and the patient.[78,88,89] The patient must actively participate in the selection and adjustment of drug treatment and in changes in lifestyle in order to maximize the usefulness of the therapeutic regimen. When feasible, self-monitoring of blood pressure can also enhance adherence.[78,90] Simplifying instructions to the patient and medication schedules is essential, and minimizing the total number of daily doses has been found to be more important in promoting adherence than minimizing the total number of medications.[48,91]

When inadequate adherence to medication has been identified and the available strategies for improving adherence have not achieved the target level of blood pressure, selecting "more forgiving" antihypertensive agents that either do not depend on half-life or have a longer half-life — drugs whose efficacy will not be affected by delayed or missed doses — will probably help to maintain a more stable blood pressure, despite imperfect adher-

Table 3. Strategies for Improving Adherence to a Medication Regimen.*

Identify poor adherence
Look for markers of nonadherence: missed appointments ("no-shows"), lack of response to medication, missed refills
Ask about barriers to adherence without being confrontational
Emphasize the value of the regimen and the effect of adherence
Elicit patient's feelings about his or her ability to follow the regimen, and if necessary, design supports to promote adherence
Provide simple, clear instructions and simplify the regimen as much as possible
Encourage the use of a medication-taking system
Listen to the patient, and customize the regimen in accordance with the patient's wishes
Obtain the help from family members, friends, and community services when needed
Reinforce desirable behavior and results when appropriate
Consider more "forgiving" medications when adherence appears unlikely†
Medications with long half-lives
Depot (extended-release) medications
Transdermal medications

* Information in this table was adapted from Osterberg and Rudd.[83]

† Forgiving medications are drugs whose efficacy will not be affected by delayed or missed doses.

ence.[40,46] When choosing among the major classes of antihypertensive agents — calcium-channel blockers, angiotensin-converting–enzyme inhibitors, angiotensin II type 1–receptor antagonists, alpha blockers, and direct vasodilators — the practitioner should consider selecting the agent with the longest half-life in each class. The antihypertensive effect of some drugs, such as the thiazide diuretics, is not related to plasma concentrations or drug half-life, and for these drugs, timing doses and short lapses in adherence are probably clinically unimportant. The most forgiving medications, such as the thiazides or modified formulations such as the transdermal clonidine patch, are more likely than less forgiving drugs to achieve an acceptable therapeutic outcome if they are otherwise tolerated.

Another strategy used by Burnier and colleagues[92] in a study of a highly selected group of patients with refractory hypertension was to monitor adherence objectively with the use of microelectronic monitors. In more than 30 percent of patients initially identified as having refractory hypertension, blood pressure became controlled merely as a result of monitoring, and an additional 20 percent of patients were identified as having lapsed adherence. Further control of blood pressure was achieved in a subgroup of subjects with poor adherence who agreed to continued monitoring and adjustment of their medications.[92]

PSYCHIATRIC ILLNESS

Patients with psychiatric illness typically have great difficulty following a medication regimen, but they also have the greatest potential for benefiting from adherence.[80,93] Half of patients with major depression for whom antidepressants are prescribed will not be taking the drugs three months after the initiation of therapy.[94] Rates of adherence among patients with schizophrenia are between 50 and 60 percent, and among those with bipolar affective disorder the rates are as low as 35 percent.[56,57,95] In a systematic review by Cramer and Rosenheck, among patients with physical disorders, the mean rate of medication adherence was 76 percent (range, 40 to 90 percent), whereas among those with psychoses the mean rate was 58 percent (range, 24 to 90 percent) and among those with depression the mean rate was 65 percent (range, 58 to 90 percent).[96]

A number of interventions to improve adherence to medication regimens among patients with psychiatric illnesses have been tried. Successful approaches include a combination of educational interventions (involving both patient and family), cognitive–supportive interventions, and the periodic use of reinforcement techniques.[73,89,97,98] Educational approaches appear to be most effective when they are combined with behavioral techniques and supportive services.[80] Reinforcements include a wide variety of techniques, such as monetary rewards or vouchers, frequent contact with the patient, and other types of personalized reminders.[79,99-101] Unfortunately, these interventions require trained personnel and repeated sessions if increased adherence is to be maintained; without these resources, adherence falls with time.

New antidepressant drugs and antipsychotic agents generally have fewer side effects than do older medications, and, consequently, their use results in reduced rates of discontinuation.[57,102-105] New agents may be preferred to older agents for a variety of reasons, but factors such as cost and effi-

cacy may be more important for some patients in achieving optimal adherence. Depot neuroleptic agents are often the treatment of choice for patients with schizophrenia who are not adhering to a regimen of oral agents.[106,107] The recent development of atypical depot neuroleptic drugs has the potential to improve adherence, since these agents combine the better efficacy and tolerability of the atypical agents with the reliability of the depot formulation.[106,108]

ILLNESS IN PEDIATRIC PATIENTS

Anyone who has seen a child with clenched teeth and a caregiver struggling desperately to administer the next dose of a medication understands the challenge of adherence to a medication regimen in the treatment of children. Achieving full adherence in pediatric patients requires not only the child's cooperation but also a devoted, persistent, and adherent parent or caregiver. Adolescent patients create even more challenges, given the unique developmental, psychosocial, and lifestyle issues implicit in adolescence.[109-112] Although the factors that contribute to poor adherence in children and adolescents are similar to those affecting adults, an added dimension of the situation is the involvement of patients' families.[113-115] Rates of adherence to medication regimens among children with chronic diseases are similar to those among adults with chronic diseases, averaging about 50 percent, with decrements in adherence occurring with time.[116-118]

Many interventions to improve adherence have been tried in pediatric patients but have had limited success. Most of the successful interventions in patients with chronic childhood illnesses have used behavioral interventions or a combination of behavioral and other interventions. The most common intervention is the token reinforcement system,[119-122] which involves motivating adherence by providing tokens or other rewards for taking medications successfully. The tokens can be used to obtain privileges, access to certain activities, or other rewards. Behavioral strategies often require resources and trained staff, yet simple reinforcement systems are practical for use by parents or other caregivers. The use of a more palatable medication than was initially prescribed has met with some success in improving adherence,[123,124] and the involvement of family members, schools, and other social supports are valuable strategies for maximizing children's ability to adhere to medication regimens.[113,115]

CONCLUSIONS

Poor adherence to medication regimens is common, contributing to substantial worsening of disease, death, and increased health care costs. Practitioners should always look for poor adherence and can enhance adherence by emphasizing the value of a patient's regimen, making the regimen simple, and customizing the regimen to the patient's lifestyle. Asking patients nonjudgmentally about medication-taking behavior is a practical strategy for identifying poor adherence. A collaborative approach to care augments adherence. Patients who have difficulty maintaining adequate adherence need more intensive strategies than do patients who have less difficulty with adherence, a more forgiving medication regimen, or both. Innovative methods of managing chronic diseases have had some success in improving adherence when a regimen has been difficult to follow.[99,125-127] New technologies such as reminders through cell phones and personal digital assistants and pillboxes with paging systems may be needed to help patients who have the most difficulty meeting the goals of a regimen.

Dr. Blaschke reports having received consulting fees from Jazz Pharmaceuticals, Portola Pharmaceuticals, Gilead Sciences, Aerogen, Depomed, Kai Pharmaceuticals, and Pharsight, and reports having shares in Johnson & Johnson and Procter & Gamble.

REFERENCES

1. Steiner JF, Earnest MA. The language of medication-taking. Ann Intern Med 2000;132:926-30.
2. Jackevicius CA, Mamdani M, Tu JV. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. JAMA 2002;288:462-7.
3. Cramer J, Rosenheck R, Kirk G, Krol W, Krystal J. Medication compliance feedback and monitoring in a clinical trial: predictors and outcomes. Value Health 2003;6:566-73.
4. Haynes RB, McDonald HP, Garg AX. Helping patients follow prescribed treatment: clinical applications. JAMA 2002;288:2880-3.
5. Benner JS, Glynn RJ, Mogun H, Neumann PJ, Weinstein MC, Avorn J. Long-term persistence in use of statin therapy in elderly patients. JAMA 2002;288:455-61.
6. Waeber B, Leonetti G, Kolloch R, McInnes GT. Compliance with aspirin or placebo in the Hypertension Optimal Treatment (HOT) study. J Hypertens 1999;17:1041-5.
7. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001;23:1296-310.
8. Rudd P, Byyny RL, Zachary V, et al. Pill

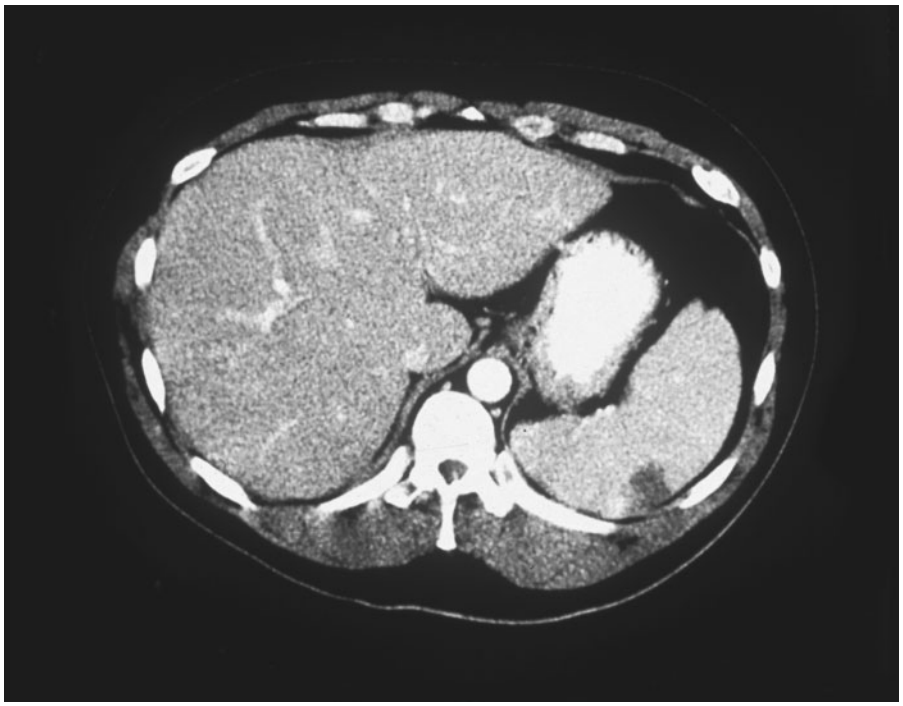
count measures of compliance in a drug trial: variability and suitability. Am J Hypertens 1988;1:309-12.
9. Pullar T, Kumar S, Tindall H, Feely M. Time to stop counting the tablets? Clin Pharmacol Ther 1989;46:163-8.
10. Spilker B. Methods of assessing and improving compliance in clinical trials. In: Cramer JA, Spilker B, eds. Patient compliance in medical practice and clinical trials. New York: Raven Press, 1991:37-56.
11. Burnier M. Long-term compliance with antihypertensive therapy: another facet of chronotherapeutics in hypertension. Blood Press Monit 2000;5:Suppl 1:S31-S34.
12. Haynes RB, McDonald H, Garg AX, Montague P. Interventions for helping patients to follow prescriptions for medications. Cochrane Database Syst Rev 2002;2: CD000011.
13. Miller LG, Liu H, Hays RD, et al. How well do clinicians estimate patients' adherence to combination antiretroviral therapy? J Gen Intern Med 2002;17:1-11.
14. Murri R, Ammassari A, Trotta MP, et al. Patient-reported and physician-estimated adherence to HAART: social and clinic center-related factors are associated with discordance. J Gen Intern Med 2004;19:1104-10.
15. McDonnell PJ, Jacobs MR. Hospital admissions resulting from preventable adverse drug reactions. Ann Pharmacother 2002; 36:1331-6.
16. Schiff GD, Fung S, Speroff T, McNutt RA. Decompensated heart failure: symptoms, patterns of onset, and contributing factors. Am J Med 2003;114:625-30.
17. Senst BL, Achusim LE, Genest RP, et al. Practical approach to determining costs and frequency of adverse drug events in a health care network. Am J Health Syst Pharm 2001; 58:1126-32.
18. Misdrahi D, Llorca PM, Lancon C, Bayle FJ. Compliance in schizophrenia: predictive factors, therapeutical considerations and research implications. Encephale 2002;28: 266-72. (In French.)
19. Rodgers PT, Ruffin DM. Medication nonadherence: Part II — a pilot study in patients with congestive heart failure. Manag Care Interface 1998;11(9):67-9, 75.
20. Levy G, Zamacona MK, Jusko WJ. Developing compliance instructions for drug labeling. Clin Pharmacol Ther 2000;68:586-91.
21. Berg JS, Dischler J, Wagner DJ, Raia JJ, Palmer-Shevlin N. Medication compliance: a healthcare problem. Ann Pharmacother 1993;27:Suppl 9:S1-S24.
22. The Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. N Engl J Med 1980;303:1038-41.
23. LaRosa JC. Poor compliance: the hidden risk factor. Curr Atheroscler Rep 2000;2: 1-4.
24. Horwitz RI, Horwitz SM. Adherence to treatment and health outcomes. Arch Intern Med 1993;153:1863-8.
25. Ickovics JR, Meisler AW. Adherence in AIDS clinical trials: a framework for clinical research and clinical care. J Clin Epidemiol 1997;50:385-91.
26. Kastrissios H, Suarez JR, Hammer S, Katzenstein D, Blaschke TF. The extent of non-adherence in a large AIDS clinical trial using plasma dideoxynucleoside concentrations as a marker. AIDS 1998;12:2305-11.
27. Sabate E. Adherence to long-term therapies: evidence for action. Geneva: World Health Organization, 2003. (Accessed July 11, 2005, at http://www.who.int/chronic_conditions/en/adherence_report.pdf.)
28. Walsh JC, Mandalia S, Gazzard BG. Responses to a 1 month self-report on adherence to antiretroviral therapy are consistent with electronic data and virological treatment outcome. AIDS 2002;16:269-77.
29. Haynes RB, Taylor DW, Sackett DL, Gibson ES, Bernholz CD, Mukherjee J. Can simple clinical measurements detect patient noncompliance? Hypertension 1980;2:757-64.
30. Wagner JH, Justice AC, Chesney M, Sinclair G, Weissman S, Rodriguez-Barradas M. Patient- and provider-reported adherence: toward a clinically useful approach to measuring antiretroviral adherence. J Clin Epidemiol 2001;54:Suppl 1:S91-S98.
31. Alcoba M, Cuevas MJ, Perez-Simon MR, et al. Assessment of adherence to triple antiretroviral treatment including indinavir: role of the determination of plasma levels of indinavir. J Acquir Immune Defic Syndr 2003; 33:253-8.
32. Cramer JA, Mattson RH, Prevey ML, Scheyer RD, Ouellette VL. How often is medication taken as prescribed? A novel assessment technique. JAMA 1989;261:3273-7. [Erratum, JAMA 1989;262:1472.]
33. Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. J Clin Epidemiol 1997;50:105-16.
34. Lau HS, de Boer A, Beuning KS, Porsius A. Validation of pharmacy records in drug exposure assessment. J Clin Epidemiol 1997; 50:619-25.
35. Christensen DB, Williams B, Goldberg HI, Martin DP, Engelberg R, LoGerfo JP. Assessing compliance to antihypertensive medications using computer-based pharmacy records. Med Care 1997;35:1164-70.
36. Spector SL, Kinsman R, Mawhinney H, et al. Compliance of patients with asthma with an experimental aerosolized medication: implications for controlled clinical trials. J Allergy Clin Immunol 1986;77:65-70.
37. Norell SE. Monitoring compliance with pilocarpine therapy. Am J Ophthalmol 1981; 92:727-31.
38. Kass MA, Meltzer DW, Gordon M. A miniature compliance monitor for eyedrop medication. Arch Ophthalmol 1984; 102:1550-4.
39. Paterson DL, Potoski B, Capitano B. Measurement of adherence to antiretroviral medications. J Acquir Immune Defic Syndr 2002;31:Suppl 3:S103-S106.
40. Urquhart J. The electronic medication event monitor: lessons for pharmacotherapy. Clin Pharmacokinet 1997;32:345-56.
41. Liu H, Golin CE, Miller LG, et al. A comparison study of multiple measures of adherence to HIV protease inhibitors. Ann Intern Med 2001;134:968-77. [Erratum, Ann Intern Med 2002;136:175.]
42. Turner BJ, Hecht FM. Improving on a coin toss to predict patient adherence to medications. Ann Intern Med 2001;134: 1004-6.
43. Paes AH, Bakker A, Soe-Agnie CJ. Impact of dosage frequency on patient compliance. Diabetes Care 1997;20:1512-7.
44. Feinstein AR. On white-coat effects and the electronic monitoring of compliance. Arch Intern Med 1990;150:1377-8.
45. Cramer JA, Scheyer RD, Mattson RH. Compliance declines between clinic visits. Arch Intern Med 1990;150:1509-10.
46. Urquhart J. The odds of the three nons when an aptly prescribed medicine isn't working: non-compliance, non-absorption, non-response. Br J Clin Pharmacol 2002;54: 212-20.
47. Greenberg RN. Overview of patient compliance with medication dosing: a literature review. Clin Ther 1984;6:592-9.
48. Eisen SA, Miller DK, Woodward RS, Spitznagel E, Przybeck TR. The effect of prescribed daily dose frequency on patient medication compliance. Arch Intern Med 1990;150:1881-4.
49. Golin CE, Liu H, Hays RD, et al. A prospective study of predictors of adherence to combination antiretroviral medication. J Gen Intern Med 2002;17:756-65.
50. DiMatteo MR, Sherbourne CD, Hays RD, et al. Physicians' characteristics influence patients' adherence to medical treatment: results from the Medical Outcomes Study. Health Psychol 1993;12:93-102.
51. van Servellen G, Chang B, Garcia L, Lombardi E. Individual and system level factors associated with treatment nonadherence in human immunodeficiency virus-infected men and women. AIDS Patient Care STDS 2002;16:269-81.
52. Ammassari A, Trotta MP, Murri R, et al. Correlates and predictors of adherence to highly active antiretroviral therapy: overview of published literature. J Acquir Immune Defic Syndr 2002;31:Suppl 3:S123-S127.
53. Stilley CS, Sereika S, Muldoon MF, Ryan CM, Dunbar-Jacob J. Psychological and cognitive function: predictors of adherence with cholesterol lowering treatment. Ann Behav Med 2004;27:117-24.
54. Okuno J, Yanagi H, Tomura S. Is cognitive impairment a risk factor for poor compliance among Japanese elderly in the community? Eur J Clin Pharmacol 2001;57: 589-94.
55. Sewitch MJ, Abrahamowicz M, Barkun A, et al. Patient nonadherence to medication

in inflammatory bowel disease. Am J Gastroenterol 2003;98:1535-44.
56. Lacro JP, Dunn LB, Dolder CR, Leckband SG, Jeste DV. Prevalence of and risk factors for medication nonadherence in patients with schizophrenia: a comprehensive review of recent literature. J Clin Psychiatry 2002;63:892-909.
57. Perkins DO. Predictors of noncompliance in patients with schizophrenia. J Clin Psychiatry 2002;63:1121-8.
58. Farley J, Hines S, Musk A, Ferrus S, Tepper V. Assessment of adherence to antiviral therapy in HIV-infected children using the Medication Event Monitoring System, pharmacy refill, provider assessment, caregiver self-report, and appointment keeping. J Acquir Immune Defic Syndr 2003;33:211-8.
59. Balkrishnan R. Predictors of medication adherence in the elderly. Clin Ther 1998;20:764-71.
60. Ellis JJ, Erickson SR, Stevenson JG. Suboptimal statin adherence and discontinuation in primary and secondary prevention populations. J Gen Intern Med 2004;19:638-45.
61. Stone VE, Hogan JW, Schuman P, et al. Antiretroviral regimen complexity, self-reported adherence, and HIV patients' understanding of their regimens: survey of women in the HER Study. J Acquir Immune Defic Syndr 2001;28:124-31.
62. Stephenson BJ, Rowe BH, Haynes RB, Macharia WM, Leon G. The rational clinical examination: is this patient taking the treatment as prescribed? JAMA 1993;269:2779-81.
63. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986;24:67-74.
64. Cramer J. Identifying and improving compliance patterns. In: Cramer JA, Spilker B, eds. Patient compliance in medical practice and clinical trials. New York: Raven Press, 1991:387-92.
65. Elliott WJ, Maddy R, Toto R, Bakris G. Hypertension in patients with diabetes: overcoming barriers to effective control. Postgrad Med 2000;107:29-32, 35-6, 38.
66. Black HR. Will better-tolerated antihypertensive agents improve blood pressure control? JNC VI revisited. Am J Hypertens 1999;12:225S-230S.
67. Ickovics JR, Meade CS. Adherence to HAART among patients with HIV: breakthroughs and barriers. AIDS Care 2002;14:309-18.
68. Murphy DA, Sarr M, Durako SJ, Moscicki AB, Wilson CM, Muenz LR. Barriers to HAART adherence among human immunodeficiency virus-infected adolescents. Arch Pediatr Adolesc Med 2003;157:249-55.
69. Stuart B, Zacker C. Who bears the burden of Medicaid drug copayment policies? Health Aff (Millwood) 1999;18(2):201-12.
70. Wilson J, Axelsen K, Tang S. Medicaid prescription drug access restrictions: exploring the effect on patient persistence with hypertension medications. Am J Manag Care 2005;11(Special Issue):SP27-SP34.
71. Zyczynski TM, Coyne KS. Hypertension and current issues in compliance and patient outcomes. Curr Hypertens Rep 2000;2:510-4.
72. Patton K, Meyers J, Lewis BE. Enhancement of compliance among patients with hypertension. Am J Manag Care 1997;3:1693-8.
73. Ran MS, Xiang MZ, Chan CL, et al. Effectiveness of psychoeducational intervention for rural Chinese families experiencing schizophrenia — a randomised controlled trial. Soc Psychiatry Psychiatr Epidemiol 2003;38:69-75.
74. Bouvy ML, Heerdink ER, Urquhart J, Grobbee DE, Hoes AW, Leufkens HG. Effect of a pharmacist-led intervention on diuretic compliance in heart failure patients: a randomized controlled study. J Card Fail 2003;9:404-11.
75. Simoni JM, Frick PA, Pantalone DW, Turner BJ. Antiretroviral adherence interventions: a review of current literature and ongoing studies. Top HIV Med 2003;11:185-98.
76. Maro EE, Lwakatare J. Medication compliance among Tanzanian hypertensives. East Afr Med J 1997;74:539-42.
77. Ross FM. Patient compliance — whose responsibility? Soc Sci Med 1991;32:89-94.
78. Feldman R, Bacher M, Campbell N, Drover A, Chockalingam A. Adherence to pharmacologic management of hypertension. Can J Public Health 1998;89:I16-I18.
79. Rigsby MO, Rosen MI, Beauvais JE, et al. Cue-dose training with monetary reinforcement: pilot study of an antiretroviral adherence intervention. J Gen Intern Med 2000;15:841-7.
80. Zygmunt A, Olfson M, Boyer CA, Mechanic D. Interventions to improve medication adherence in schizophrenia. Am J Psychiatry 2002;159:1653-64.
81. Cramer JA. Optimizing long-term patient compliance. Neurology 1995;45:Suppl 1:S25-S28.
82. Crespo-Fierro M. Compliance/adherence and care management in HIV disease. J Assoc Nurses AIDS Care 1997;8(4):43-54.
83. Osterberg LG, Rudd P. Medication adherence for antihypertensive therapy. In: Oparil S, Weber MA, eds. Hypertension: a companion to Brenner and Rector's The Kidney. 2nd ed. Philadelphia: Elsevier Mosby, 2005:848.
84. Ickovics JR, Cameron A, Zackin R, et al. Consequences and determinants of adherence to antiretroviral medication: results from Adult AIDS Clinical Trials Group protocol 370. Antivir Ther 2002;7:185-93.
85. Chesney M. Adherence to HAART regimens. AIDS Patient Care STDS 2003;17:169-77.
86. Tuldra A, Fumaz CR, Ferrer MJ, et al. Prospective randomized two-arm controlled study to determine the efficacy of a specific intervention to improve long-term adherence to highly active antiretroviral therapy. J Acquir Immune Defic Syndr 2000;25:221-8.
87. Safren SA, Otto MW, Worth JL, et al. Two strategies to increase adherence to HIV antiretroviral medication: life-steps and medication monitoring. Behav Res Ther 2001;39:1151-62.
88. Golin CE, DiMatteo MR, Gelberg L. The role of patient participation in the doctor visit: implications for adherence to diabetes care. Diabetes Care 1996;19:1153-64.
89. Shon KH, Park SS. Medication and symptom management education program for the rehabilitation of psychiatric patients in Korea: the effects of promoting schedule on self-efficacy theory. Yonsei Med J 2002;43:579-89.
90. Vrijens B, Goetghebeur E. Comparing compliance patterns between randomized treatments. Control Clin Trials 1997;18:187-203.
91. Schroeder K, Fahey T, Ebrahim S. How can we improve adherence to blood pressure-lowering medication in ambulatory care? Systematic review of randomized controlled trials. Arch Intern Med 2004;164:722-32.
92. Burnier M, Schneider MP, Chiolero A, Stubi CL, Brunner HR. Electronic compliance monitoring in resistant hypertension: the basis for rational therapeutic decisions. J Hypertens 2001;19:335-41.
93. Nose M, Barbui C, Gray R, Tansella M. Clinical interventions for treatment nonadherence in psychosis: meta-analysis. Br J Psychiatry 2003;183:197-206.
94. Vergouwen AC, van Hout HP, Bakker A. Methods to improve patient compliance in the use of antidepressants. Ned Tijdschr Geneeskd 2002;146:204-7. (In Dutch.)
95. Colom F, Vieta E, Martinez-Aran A, Reinares M, Benabarre A, Gasto C. Clinical factors associated with treatment noncompliance in euthymic bipolar patients. J Clin Psychiatry 2000;61:549-55.
96. Cramer JA, Rosenheck R. Compliance with medication regimens for mental and physical disorders. Psychiatr Serv 1998;49:196-201.
97. Mundt JC, Clarke GN, Burroughs D, Brenneman DO, Griest JH. Effectiveness of antidepressant pharmacotherapy: the impact of medication compliance and patient education. Depress Anxiety 2001;13:1-10.
98. Pekkala E, Merinder L. Psychoeducation for schizophrenia. Cochrane Database Syst Rev 2002;2:CD002831.
99. Molassiotis A, Lopez-Nahas V, Chung WY, Lam SW. A pilot study of the effects of a behavioural intervention on treatment adherence in HIV-infected patients. AIDS Care 2003;15:125-35.
100. Fulmer TT, Feldman PH, Kim TS, et al. An intervention study to enhance medication compliance in community-dwelling el-

derly individuals. J Gerontol Nurs 1999;25:6-14.
101. Dunbar PJ, Madigan D, Grohskopf LA, et al. A two-way messaging system to enhance antiretroviral adherence. J Am Med Inform Assoc 2003;10:11-5.
102. Mitchell PB, Mitchell MS. The management of depression. Part 2: the place of the new antidepressants. Aust Fam Physician 1994;23:1771-3, 1776-81.
103. Nemeroff CB. Improving antidepressant adherence. J Clin Psychiatry 2003;64:Suppl 18:25-30.
104. Rosenheck R, Chang S, Choe Y, et al. Medication continuation and compliance: a comparison of patients treated with clozapine and haloperidol. J Clin Psychiatry 2000;61:382-6.
105. Hoffman L, Enders J, Luo J, Segal R, Pippins J, Kimberlin C. Impact of an antidepressant management program on medication adherence. Am J Manag Care 2003;9:70-80.
106. Love RC. Strategies for increasing treatment compliance: the role of long-acting antipsychotics. Am J Health Syst Pharm 2002;59:Suppl 8:S10-S15.
107. Gillis LS, Trollip D, Jakoet A, Holden T. Non-compliance with psychotropic medication. S Afr Med J 1987;72:602-6.
108. Bhanji NH, Chouinard G, Margolese HC. A review of compliance, depot intramuscular antipsychotics and the new long-acting injectable atypical antipsychotic risperidone in schizophrenia. Eur Neuropsychopharmacol 2004;14:87-92.
109. Tebbi CK, Cummings KM, Zevon MA, Smith L, Richards M, Mallon J. Compliance of pediatric and adolescent cancer patients. Cancer 1986;58:1179-84.
110. Friedman IM, Litt IF. Adolescents' compliance with therapeutic regimens: psychological and social aspects and intervention. J Adolesc Health Care 1987;8:52-67.
111. Friedman IM, Litt IF, King DR, et al. Compliance with anticonvulsant therapy by epileptic youth: relationships to psychosocial aspects of adolescent development. J Adolesc Health Care 1986;7:12-7.
112. Michaud PA, Suris JC, Viner R. The adolescent with a chronic condition. Part II: healthcare provision. Arch Dis Child 2004;89:943-9.
113. Tebbi CK. Treatment compliance in childhood and adolescence. Cancer 1993;71:Suppl 10:3441-9.
114. Loghman-Adham M. Medication noncompliance in patients with chronic disease: issues in dialysis and renal transplantation. Am J Manag Care 2003;9:155-71.
115. Brackis-Cott E, Mellins CA, Abrams E, Reval T, Dolezal C. Pediatric HIV medication adherence: the views of medical providers from two primary care programs. J Pediatr Health Care 2003;17:252-60.
116. Jay S, Litt IF, Durant RH. Compliance with therapeutic regimens. J Adolesc Health Care 1984;5:124-36.
117. Rapoff MA. Compliance with treatment regimens for pediatric rheumatic diseases. Arthritis Care Res 1989;2:S40-S47.
118. Festa RS, Tamaroff MH, Chasalow F, Lanzkowsky P. Therapeutic adherence to oral medication regimens by adolescents with cancer. I. Laboratory assessment. J Pediatr 1992;120:807-11.
119. Rapoff MA, Lindsley CB, Christophersen ER. Improving compliance with medical regimens: case study with juvenile rheumatoid arthritis. Arch Phys Med Rehabil 1984;65:267-9.
120. Greenan-Fowler E, Powell C, Varni JW. Behavioral treatment of adherence to therapeutic exercise by children with hemophilia. Arch Phys Med Rehabil 1987;68:846-9.
121. Epstein LH, Beck S, Figueroa J, et al. The effects of targeting improvements in urine glucose on metabolic control in children with insulin dependent diabetes. J Appl Behav Anal 1981;14:365-75.
122. Magrab PR, Papadopoulou ZL. The effect of a token economy on dietary compliance for children on hemodialysis. J Appl Behav Anal 1977;10:573-8.
123. Tucker CM, Fennell RS, Pedersen T, Higley BP, Wallack CE, Peterson S. Associations with medication adherence among ethnically different pediatric patients with renal transplants. Pediatr Nephrol 2002;17:251-6.
124. Blaiss M. Current concepts and therapeutic strategies for allergic rhinitis in school-age children. Clin Ther 2004;26:1876-89.
125. Weingarten SR, Henning JM, Badamgarav E, et al. Interventions used in disease management programmes for patients with chronic illness — which ones work? Meta-analysis of published reports. BMJ 2002;325:925-8.
126. Ofman JJ, Badamgarav E, Henning JM, et al. Does disease management improve clinical and economic outcomes in patients with chronic diseases? A systematic review. Am J Med 2004;117:182-92.
127. Farris KB, Cote I, Feeny D, et al. Enhancing primary care for complex patients: demonstration project using multidisciplinary teams. Can Fam Physician 2004;50:998-1003.

IMAGES IN CLINICAL MEDICINE

Medical Mystery — Abdominal Pain



Karen M. Kim, M.D.

Tufts University School of Medicine
Boston, MA 02111

Richard I. Kopelman, M.D.

Tufts–New England Medical Center
Boston, MA 02111

A 40-year-old, previously healthy man presented with a four-day history of pain in the left upper quadrant of the abdomen, accompanied by fatigue, fever, sweating, and sore throat. Physical examination revealed exudative pharyngitis, an enlarged lymph node in the left-sided posterior cervical region, and mild tenderness in the left upper quadrant with no hepatosplenomegaly. A computed tomographic scan of the patient's abdomen is shown here. What is the diagnosis?

Editor's note: We invite our readers to submit their answers at www.nejm.org/mystery. We will publish the diagnosis in the Correspondence section of the September 29, 2005, issue and e-mail it to everyone who submits an answer. All answers must be received by August 18, 2005.

CLINICAL PROBLEM-SOLVING

A Fractured Diagnosis

Tali Cukierman, M.D., Moshe E. Gatt, M.D., Nurith Hiller, M.D., and Tova Chajek-Shaul, M.D.

In this Journal *feature, information about a real patient is presented in stages (boldface type) to an expert clinician, who responds to the information, sharing his or her reasoning with the reader (regular type). The authors' commentary follows.*

From the Department of Medicine, Hadassah–Hebrew University Medical Center, Mount Scopus, Jerusalem. Address reprint requests to Dr. Chajek-Shaul at the Department of Medicine, Hadassah–Hebrew University Medical Center, Mount Scopus, P.O. Box 24035, Jerusalem 91240, Israel, or at chajek@hadasssah.org.il.

N Engl J Med 2005;353:509-14.

A 44-year-old woman came to the emergency department because of pain in her right thigh shortly after she had a minor fall. A right femoral-neck fracture was diagnosed, and she was admitted to the orthopedic ward to await surgery. Six months before hospitalization, and before she fell, limb pain had developed, which had become progressively worse. The patient also reported a weight loss of 30 kg and fatigue.

A fractured femoral neck after a minor fall in a patient 44 years of age is very uncommon. The associated limb pain and severe weight loss suggest the possibility of cancer, chronic infection, or autoimmune disease. A fracture may also be secondary to osteosclerosis or osteoporosis caused by immobilization, low intake of calcium and vitamin D or malabsorption, endocrine disorders such as Cushing's syndrome or hyperparathyroidism, renal disease, or adverse effects of drugs such as glucocorticoids or anticonvulsants. The possibility of an eating disorder should also be considered. I would want to know whether the patient has anorexia or difficulty in eating, whether she has other manifestations of systemic disease (such as fever, rash, or neurologic symptoms), and whether she takes any medications.

The patient was a housewife with four healthy children. She reported mild difficulty in swallowing solid food, which caused her to eat less. She said that she had not had nausea, vomiting, diarrhea or constipation, fevers, sweats, itching, or rash in the past several months. She did not smoke and reported no drug use.

On physical examination, the patient appeared cachectic and pale. She had poor oral hygiene and mild cervical lymphadenopathy. She had substantial bone tenderness on palpation of her limbs and chest. The remainder of the physical examination, which included breast, rectal, and neurologic examination, showed no abnormalities.

The hemoglobin level was 11.2 g per deciliter, and the mean corpuscular volume 90 μm³. The white-cell count, platelet count, levels of liver enzymes, and the prothrombin time and partial-thromboplastin time were within normal ranges. The level of blood urea nitrogen was 4.2 mmol per liter (normal range, 3.3 to 6.5), and the level of serum creatinine 67 μmol per liter (normal, 60 to 106). The serum potassium levels were persistently low (2.3 to 3 mmol per liter) and were not corrected by oral supplementation with 9 g of potassium chloride daily for four days. The orthopedic operation was postponed, and the patient was transferred to the internal medicine ward for further evaluation.

A metastatic cancer could explain the dysphagia, bone pain, and reduced food intake. A collagen vascular disease such as scleroderma with esophageal dysmotility or myositis might be an alternative explanation. The absence of fever makes a chronic infection

less likely. Multiple myeloma may explain the bone pain, fatigue, anemia, weight loss, and hypokalemia (which could result from renal tubular dysfunction); serum and urinary protein electrophoresis should be performed. A skeletal survey may aid in the diagnosis of lytic lesions and other bone abnormalities.

The persistent hypokalemia could be caused by renal wasting or, alternatively, could indicate gastrointestinal losses. Measurement of the levels of other electrolytes, such as chloride and phosphate, and determination of the blood pH, partial pressure of carbon dioxide, and bicarbonate concentration are required.

Additional laboratory testing revealed a serum phosphate level of 1.7 mg per liter (0.55 mmol per liter; normal range, 2.5 to 4.5 mg per liter [0.8 to 1.45 mmol per liter]); serum calcium level of 8.2 mg per deciliter (2.0 mmol per liter; normal range, 8 to 10 mg per deciliter [2.0 to 2.5 mmol per liter]); sodium, 144 mmol per liter; potassium, 2.8 mmol per liter; chloride, 119 mmol per liter; magnesium, 2.2 mg per deciliter (0.9 mmol per liter); uric acid, 2.6 mg per deciliter (154.6 μmol per liter; normal range, 2.5 to 6.3 mg per deciliter [148.7 to 374.7 μmol per liter]); alkaline phosphatase, 168 U per liter (normal range, 40 to 130) and albumin, 3.2 g per deciliter. The blood pH was 7.32, the partial pressure of carbon dioxide 37 mm Hg, and the bicarbonate level 18.7 mmol per liter, with a normal value for the serum anion gap.

The results of protein electrophoresis in the blood and urine were normal. Radiographs of the spine, hips, and limbs showed diffuse osteopenia without lytic lesions or fractures. A reevaluation of the right femur radiograph showed osteopenia with a Looser–Milkman pseudofracture (Fig. 1). A computed tomographic (CT) scan of the chest showed several pseudofractures of the ribs (Fig. 2). These findings indicated that the canceled surgery would not be required.

This patient has radiologic findings that are consistent with osteomalacia, with low serum phosphate, normal calcium levels, and raised levels of alkaline phosphatase. A low bone density is also characteristic of osteomalacia, although this finding would not distinguish it from osteoporosis. Whereas the gold standard for the diagnosis of osteomalacia involves bone biopsy with double tetracycline labeling, the radiologic findings combined with the abnormalities in phosphate and alkaline phosphatase levels in this case are sufficient to establish the working diagnosis of osteomalacia. The osteomalacia may be secondary to abnormal vitamin D metabolism (where secondary hyperparathyroidism would be expected), or it could be secondary to a phosphate deficiency, such as in hyperparathyroidism, Fanconi's syndrome, hypophosphatemic rickets, oncogenic osteomalacia, or other conditions that cause mineralization defects (for example, abnormal pH or abnormal alkaline phosphatase bioactivity). In this patient, the coexistence of osteomalacia, persistent hypokalemia, and hyperchloremic normal anion-gap metabolic acidosis suggests a renal tubular disorder (such as renal tubular acidosis) or a malabsorptive process (such as celiac or inflammatory bowel disease) that causes the loss of potassium and bicarbonate and malabsorption of vitamin D. Because the patient did not report diarrhea among her symptoms, the diagnosis is more likely to favor a renal tubular disorder, although the possibility of malabsorption has not been ruled out. Additional testing would be warranted at this point, including measurement of uri-

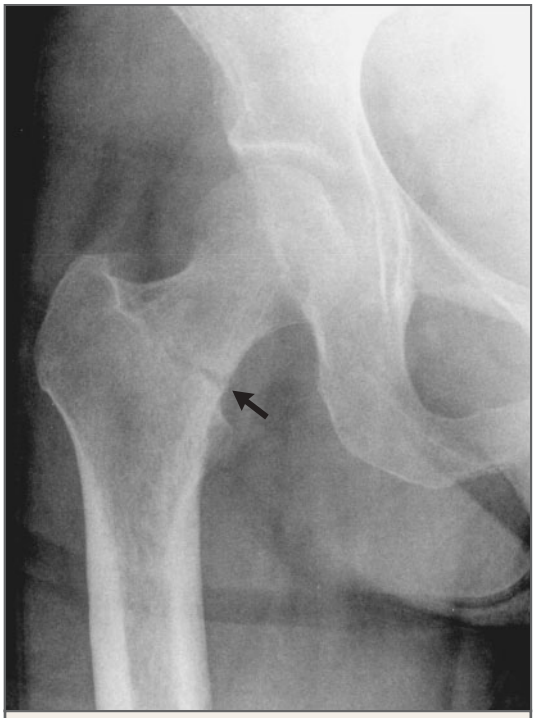


Figure 1. A Radiograph of the Right Hip Joint.
A pseudofracture is visible at the medial aspect of the neck of the femur (arrow).

Figure 2. A Computed Tomographic Scan of the Chest.
A pseudofracture is visible at the anterior aspect of the sixth rib (arrow).

nary pH and urine levels of urea nitrogen and creatinine, the osmolality of a spot-urine sample, and a 24-hour urine collection for electrolytes — including phosphorus and calcium. The presence of glycosuria and aminoaciduria would support the diagnosis of Fanconi's syndrome. Measurement of the level of serum parathyroid hormone could be helpful; the level does not increase in some primary renal-phosphate–wasting syndromes, such as oncogenic osteomalacia. An endoscopy with duodenal biopsy could rule out malabsorption due to celiac disease.

The urinary pH was 8; the daily urine output was 2.2 liters (55 ml per kilogram of body weight per day [normal, less than 50]); the level of urine urea nitrogen was 69.5 mmol per liter; creatinine, 2730 μmol per liter; sodium, 66 mmol per liter; chloride, 79 mmol per liter; potassium, 32.6 mmol per liter; calcium, 7.5 mg per deciliter (1.9 mmol per liter) and daily calcium excretion, 165 mg (4.1 mg per kilogram of body weight per day [normal, less than 4]); phosphate, 16.5 mg per deciliter (5.32 mmol per liter), and daily phosphate excretion, 380 mg; magnesium, 0.8 mmol per liter; urine osmolality, 279 mOsm per kilogram of water; plasma osmolality, 295 mOsm per kilogram of water; range in urinary potassium-to-creatinine ratio, 9.1 to 11.9; urinary anion gap (urine:Na+K–Cl), 19.6 mmol per liter; and urinary osmolal gap ($osm_u=[2Na_u+2K_u+urea_u+glucose_u]$), 12.3 mmol per liter. No glycosuria and aminoaciduria were detected. The fractional excretion of phosphate was 23.7 percent (normal, 5 to 20); tubular reabsorption of phosphorus, 77 percent; and tubular reabsorption of calcium, 97.5 percent. Serum parathyroid hormone levels were 55 to 79 pg per milliliter (normal, 10 to 60), associated with serum calcium levels of 8.2 to 8.4 mg per deciliter (2.0 to 2.1 mmol per liter). Endoscopy revealed no pathologic findings, and a biopsy specimen obtained from the distal part of the duodenum appeared normal.

In the presence of depleted plasma potassium, the appropriate renal response would be to excrete less than 15 mmol of potassium daily, with a spot urinary ratio of potassium to creatinine of 1 to 1.5. This patient's daily urinary potassium level of 71.7 mmol and the high ratio of potassium to creatinine in the spot urinary sample indicate renal loss. Furthermore, the absence of diarrhea and the normal duodenal-biopsy specimen argue against gastrointestinal losses of potassium. The high level of urinary phosphate excretion and low tubular reabsorption of phosphorus indicate that there is also renal phosphate wasting. Hyperparathyroidism is a common cause of phosphate wasting but it is unlikely in this case, given the consistently normal levels of serum calcium. Vitamin D deficiency, another possible cause of phosphate wasting, is typically associated with a higher value of parathyroid hormone (above 100 pg per milliliter) and also could not account for the potassium wasting. More likely causes in this case include primary defects such as X-linked hypophosphatemic rickets or secondary defects such as renal tubular acidosis or oncogenic osteomalacia. Primary or acquired Fanconi's syndrome could also cause renal phosphate wasting, although this would be associated with glycosuria, aminoaciduria, and hypouricemia; how-

ever, an isolated defect in proximal tubular function that led to a decrease in phosphate reabsorption may still be a possibility.

The defective urinary concentrating ability and mild polyuria are most likely due to hypokalemia. Other disorders that may affect the collecting tubules, such as amyloidosis or Sjögren's syndrome, are also possible.

In the absence of diarrhea, the presence of a normal anion-gap metabolic acidosis (also called hyperchloremic acidosis) with hypokalemia most likely indicates a renal tubular acidosis. The positive urinary anion gap and the low urinary osmolal gap (below 100 mmol per liter) implicate reduced urinary ammonium excretion (as occurs in types 1 and 4 renal tubular acidosis); the presence of hypokalemia and a urinary pH that was persistently above 6 indicate that this is a type 1 renal tubular acidosis; the presence of hyperkalemia and a urinary pH that is below 5.5 during acidosis would indicate type 4 renal tubular acidosis. At this stage of the differential diagnosis, I would take the medical history again and ask about possible familial renal or bone disease, hearing impairment, or any other hereditary condition that might cause distal renal tubular acidosis, such as Fabry's disease.

An arginine hydrochloride test was performed (to assess effects on urine pH of acidification of the plasma). After the infusion of an arginine hydrochloride solution over the course of two hours, the blood pH was assessed at 7.34, 7.28, and 7.27 at zero, two, and six hours, respectively, although the urine remained at a high-alkaline pH (pH 8) throughout the test. The patient said she knew of no familial disorders.

The inability of the kidney to lower urinary pH below 5.5 after an acid load confirms the impression of a distal renal tubular acidosis. The patient has not received medications that could cause this disorder (such as amphotericin), and there is no clinical or biochemical evidence of primary biliary cirrhosis, chronic active hepatitis, or hypercalciuria with nephrocalcinosis, which may be associated with renal tubular acidosis. Acquired renal tubular acidosis has also been described with rheumatologic diseases, including systemic lupus erythematosus, rheumatoid arthritis, and Sjögren's syndrome. The patient's dysphagia, combined with her poor oral hygiene, suggest the possibility of a lack of saliva, a hallmark of Sjögren's syndrome.

On further questioning, the patient said that she had had symptoms of dry mouth and a recurrent sensation of sand in her eyes for the past year. Serologic testing for antinuclear antibodies was positive (+3 out of 4), anti-Ro (SS-A) was 482 U per milliliter, anti-La (SS-B) was 608 U per milliliter (normal for both is less than 25), and C3 was 63 mg per deciliter (normal range, 50 to 120). Schirmer's test for tear production was positive, meaning no tears, and a salivary-gland biopsy showed lymphocytic infiltration (Fig. 3).

The patient meets five of the six so-called European criteria for the diagnosis of Sjögren's syndrome: the presence of autoantibodies, subjective ocular and oral symptoms, a positive Schirmer's test, and histopathological features. In light of the association between Sjögren's syndrome and lymphoma, the 30-kg weight loss, and the mild lymphadenopathy, lymphoma should be ruled out.

A CT scan of the chest and abdomen revealed no lymphadenopathy or splenomegaly. Treatment with daily doses of sodium bicarbonate (2 g), potassium citrate (2 g), sodium phosphate (1.5 g), potassium chloride (4 g), calcium carbonate (1 g), calcitriol (0.25 μg), and an artificial tear solution was initiated. The patient's condition responded with gradual, progressive improvement. One year later, at this writing, the patient is well. She is now able to walk, has gained weight, and reports a reduction in her general weakness and limb pain.

COMMENTARY

Fractures, bone pain, and reduced bone density in a relatively young patient require a thorough investigation. It is important to differentiate between early-onset osteoporosis and osteomalacia. The distinction between the two can be made most accurately by bone biopsy with double tetracycline labeling.[1] However, this procedure is rarely performed, because the diagnosis can be elicited from the history, physical findings, radiographic evaluation, and laboratory results. In osteoporosis, normal levels of serum calcium, phosphate, alkaline phosphatase, and parathyroid hormone are present; in contrast, abnormalities in at least one of these measurements are common in osteomalacia.

In order for adequate mineralization of the bone to occur, normal concentrations of calcium and phosphate in the extracellular fluid, adequate bio-

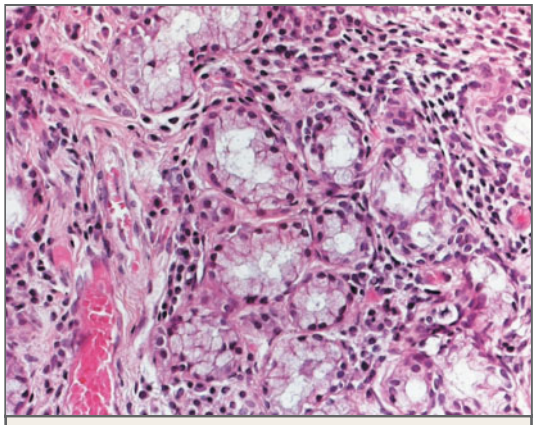


Figure 3. Biopsy Specimen of the Salivary Gland Showing Lymphocytic Infiltrates (Hematoxylin and Eosin).

activity of alkaline phosphatase, a normal pH value at the site of calcification, and the absence of calcification inhibitors are needed.[2] When any of these factors is altered, osteomalacia may result. Vitamin D deficiency or vitamin D resistance accounts for the majority of cases of osteomalacia. The most common cause is dietary deficiency and lack of exposure to sunlight[3]; malabsorptive disorders, such as celiac sprue,[4] underlie many other cases.

Chronic metabolic acidosis induces calcium loss from bone associated with hypercalciuria.[5] It may also induce renal phosphate depletion by way of its direct effect on the sodium phosphate cotransporter.[6] However, there is no agreement whether distal renal tubular acidosis causes osteomalacia.[7] The increased urinary phosphate excretion and hypophosphatemia in our patient may suggest the presence of a proximal tubular defect. In the absence of glycosuria, aminoaciduria, and hypouricemia, we were unable to establish the diagnosis of Fanconi's syndrome. The possibility of multiple tubular dysfunctions in this patient is also supported by the hyposthenuria and mild polyuria, suggesting a defective urine-concentrating capacity, although hypokalemia may also underlie these findings.

In adults, secondary distal renal tubular acidosis may develop as a consequence of calcium disorders (idiopathic hypercalciuria with nephrocalcinosis, primary hyperparathyroidism), use of drugs or toxins (amphotericin, trimethoprim, pentamidine, toluene), or autoimmune diseases (Sjögren's syndrome, rheumatoid arthritis, systemic lupus erythematosus, chronic liver diseases).[8] In this patient, the dysphagia, poor oral hygiene, and eye dryness gave the clues to the diagnosis of Sjögren's syndrome. The most common clinical manifestations of Sjögren's syndrome are exocrine gland involvement: keratoconjunctivitis sicca (dry eyes) and xerostomia (dry mouth and salivary-gland enlargement). Xerostomia can be made manifest in many ways, including dysphagia, dental caries, oral candidiasis, difficulty in speaking, weight loss, or even chronic esophagitis. Abnormal esophageal peristalsis causes severe dysphagia in Sjögren's syndrome.[9]

The patient under discussion met the revised European criteria for the diagnosis of Sjögren's syndrome.[10] The clinical picture suggests primary Sjögren's syndrome because of the lack of clinical features associated with rheumatoid arthritis or other autoimmune disorders.

Tubulointerstitial nephritis is the most common renal complication in Sjögren's syndrome. It is usually recognized in association with distal renal tubular acidosis, nephrogenic diabetes insipidus, and more rarely, proximal tubular abnormalities and Fanconi's syndrome. Up to 25 percent of patients with Sjögren's syndrome have some defect in distal acidification that may be associated with other tubular dysfunctions.[11,12]

There is no consensus regarding treatment. When interstitial nephritis occurs, a course of corticosteroids is usually beneficial; less is known about the benefits of such treatment in the case of distal renal tubular acidosis. There are a few case reports in the literature describing osteomalacia as the presenting manifestation of Sjögren's syndrome.[13-15] In two cases, treatment with alkali, vitamin D, and calcium resulted in improvement; in one case, steroid treatment resulted in similar improvement. In light of the prominent tubular acidification abnormality, this patient was treated with bicarbonate, potassium, and phosphate supplementation along with calcitriol, and she had a good clinical response.

REFERENCES

1. Bingham CT, Fitzpatrick LA. Noninvasive testing in the diagnosis of osteomalacia. Am J Med 1993;95:519-23.
2. Russell JA. Osteomalacic myopathy. Muscle Nerve 1994;17:578-80.
3. Lamberg-Allardt CJ, Outila TA, Karkkainen MU, Rita HJ, Valsta LM. Vitamin D deficiency and bone health in healthy adults in Finland: could this be a concern in other parts of Europe? J Bone Miner Res 2001;16:2066-73.
4. Tovey FI, Hall ML, Ell PJ, Hobsley M. A review of postgastrectomy bone disease. J Gastroenterol Hepatol 1992;7:639-45.
5. Krieger NS, Frick KK, Bushinsky DA. Mechanism of acid-induced bone resorption. Curr Opin Nephrol Hypertens 2004;13:423-36.

6. Wiederkehr M, Krapf R. Metabolic and endocrine effects of metabolic acidosis in humans. Swiss Med Wkly 2001;131:127-32.
7. Fulop M, Mackay M. Renal tubular acidosis, Sjogren syndrome, and bone disease. Arch Intern Med 2004;164:905-9.
8. Rodriguez Soriano J. Renal tubular acidosis: the clinical entity. J Am Soc Nephrol 2002;13:2160-70.
9. Anselmino M, Zaninotto G, Costantini M, et al. Esophageal motor function in primary Sjogren's syndrome: correlation with dysphagia and xerostomia. Dig Dis Sci 1997; 42:113-8.
10. Vitali C, Bombardieri S, Moutsopoulos HM, et al. Preliminary criteria for the classification of Sjogren's syndrome: results of a prospective concerted action supported by the European Community. Arthritis Rheum 1993;36:340-7.
11. Bossini N, Savoldi S, Franceschini F, et al. Clinical and morphological features of kidney involvement in primary Sjogren's syndrome. Nephrol Dial Transplant 2001;16: 2328-36.
12. Vitali C, Tavoni A, Sciuto M, Maccheroni M, Moriconi L, Bombardieri S. Renal involvement in primary Sjogren's syndrome: a retrospective-prospective study. Scand J Rheumatol 1991;20:132-6.
13. Okada M, Suzuki K, Hidaka T, et al. Rapid improvement of osteomalacia by treatment in a case of Sjogren's syndrome, rheumatoid arthritis and renal tubular acidosis type 1. Intern Med 2001;40:829-32.
14. Hajjaj-Hassouni N, Guedira N, Lazrak N, et al. Osteomalacia as a presenting manifestation of Sjogren's syndrome. Rev Rhum Engl Ed 1995;62:529-32.
15. Monto Neto JT, Sesso R, Kirsztajn GM, Da Silva LC, De Carvalho AB, Pereira AB. Osteomalacia secondary to renal tubular acidosis in a patient with primary Sjogren's syndrome. Clin Exp Rheumatol 1991;9:625-7.

EDITORIALS

Problem Solved? West Nile Virus and Transfusion Safety

Lyle R. Petersen, M.D., M.P.H., and Jay S. Epstein, M.D.

In 2002, just three years after its appearance in the Western Hemisphere, West Nile virus caused the largest outbreak of arboviral encephalitis ever recorded in the United States.[1] Epidemiologic investigations that year revealed that West Nile virus could be transmitted by blood transfusion,[2] and mathematical models suggested that hundreds of transmissions had occurred.[3] By July 2003, shortly before a second seasonal outbreak of similar magnitude began, collaborations among blood-collection organizations, test-kit manufacturers, and government agencies culminated in near-universal screening of U.S. blood donations for West Nile virus with the use of newly developed nucleic acid amplification tests.

In this issue of the *Journal*, two reports, one by Stramer et al.[4] and one by Busch et al.,[5] analyze the results of screening 7.1 million donations for West Nile virus, providing the first large-scale evaluation of the program. Initially, screening for West Nile virus was performed only on pooled blood samples from 16 donors ("minipools"). As expected, screening yields were characterized by extreme geographic and temporal variations. Remarkably, viremia was documented in as many as 1 in 150 donors in some areas during epidemics. From among approximately 27.2 million donations screened nationwide during the first two years of screening, from July 2003 through June 2005, 1039 blood donors with viremia were reported nationwide to the Centers for Disease Control and Prevention — a yield of 1 in 26,200. This yield compares with yields ranging from approximately 1 in 100 to 1 in 1000 donations for human immunodeficiency virus (HIV), hepatitis B virus (HBV), and hepatitis C virus (HCV) when screening for those viruses began.

Both Busch et al. and Stramer et al. confirm that the viremia associated with West Nile virus infection is of low titer and borders on the lower limit of sensitivity of the screening tests that use minipools. Of concern, approximately one third to one half of the donations with demonstrable viremia on nucleic acid amplification testing when evaluated individually were not identified by minipool screening, establishing that testing of individual donations had a substantially greater yield than minipool testing in areas with outbreaks. On this basis, individual donations were tested in various regions according to the regional rate of detection of West Nile virus by minipool screening.

However, the true benefits of testing individual donations may be more limited than it first appears. Both groups found that 8 to 15 percent of all donations with demonstrable viremia only on testing of individual donations lacked IgM antibody. Notably, all 30 cases of transfusion-transmitted West Nile virus that were documented from 2002 to 2004 resulted from IgM antibody–negative donations.[2,6,7] Conversely, limited retrospective studies showed that transfused viremic donations missed by minipool screening did not transmit West Nile virus if IgM antibody was present. Further studies are needed to resolve whether IgM-positive donations are ever infectious. However, if we assume that the presence of IgM antibody prevents virus transmission and that nearly all viremic donations reported from July 2003 to June 2005 were based on detection by minipool screening, we estimate that minipool screening alone identified 93 percent of infectious donations, for a residual risk of less than 1 in 350,000, calculated according to the following formula: $[(1039 \div 0.93) - 1039] \div 27.2$ million. This risk can be compared with contemporary risks of transfusion-transmitted HBV of 1 in 220,000 donations, transfusion-transmitted HCV of 1 in 1,600,000 donations, and transfusion-transmitted HIV of 1 in 1,800,000 donations.[8]

The screening program for West Nile virus dif-

fers substantially in concept from those for HBV, HCV, and HIV. Although the risk of transfusion-transmitted HBV, HCV, and HIV derives mostly from a prolonged carrier state in donor populations with a relatively low and stable incidence of infection, the risk associated with West Nile virus derives mostly from a short, asymptomatic period of viremia in populations with an extremely variable and seasonal incidence of infection.[3] Because the future incidence of West Nile virus infections is unknown, the long-term public health benefit of screening blood donations for this virus cannot yet be defined. A second difference is that serologic testing forms the cornerstone of screening for HBV, HCV, and HIV, whereas nucleic acid amplification testing for HCV and HIV eliminates a small, residual risk of transfusion-transmitted infection from incident infections.[9] The high level of viremia typical of acute HCV and HIV infections translates into a very high level of sensitivity of nucleic acid amplification testing with minipools of up to 24 samples. In contrast, screening for West Nile virus relies solely on nucleic acid amplification testing of samples that often have very low viral titers.

The strong working relationships developed and expanded between public health officials and blood-collection agencies since the HIV epidemic in the early 1980s facilitated investigations of the first cases of possible transfusion-transmitted West Nile virus infection. Moreover, recent technological developments, such as screening for HIV and HCV with the use of nucleic acid amplification tests, allowed the rapid development and implementation of screening tests for West Nile virus. Subsequently, the screening program for West Nile virus forged a new rapid-response relationship between transfusion medicine and public health. Because the identification of donors with West Nile virus viremia may provide the earliest indication of mosquito-borne transmission to humans in a community, blood-collection agencies now notify health departments of these donors so that other control efforts can be heightened. For example, in 2004, viremic blood donors were identified first in 7 of the 57 counties in which both viremic blood donors and other human illnesses from West Nile virus were identified.

Our experience with West Nile virus tells us that the next emerging infectious-disease threat to the U.S. blood supply is probably one not yet imagined. Nevertheless, with the right collaboration of scientific disciplines and the adaptation of newer forms of technology, our ability to respond has never been better.

From the National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colo. (L.R.P.); and the Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville, Md. (J.S.E.).

1. O'Leary DR, Marfin AA, Montgomery SP, et al. The epidemic of West Nile virus in the United States, 2002. Vector Borne Zoonotic Dis 2004;4:61-70.
2. Pealer LN, Marfin AA, Petersen LR, et al. Transmission of West Nile virus through blood transfusion in the United States in 2002. N Engl J Med 2003;349:1236-45.
3. Biggerstaff BJ, Petersen LR. Estimated risk of transmission of the West Nile virus through blood transfusion in the US, 2002. Transfusion 2003;43:1007-17.
4. Stramer SL, Fang CT, Foster GA, Wagner AG, Brodsky JP, Dodd RY. West Nile virus among blood donors in the United States, 2003 and 2004. N Engl J Med 2005;353:451-9.
5. Busch MP, Caglioti S, Robertson EF, et al. Screening the blood supply for West Nile virus RNA by nucleic acid amplification testing. N Engl J Med 2005;353:460-7.
6. Update: West Nile virus screening of blood donations and transfusion-associated transmission — United States, 2003. MMWR Morb Mortal Wkly Rep 2004;53:281-4.
7. Transfusion-associated transmission of West Nile virus — Arizona, 2004. MMWR Morb Mortal Wkly Rep 2004;53:842-4.
8. Busch MP, Kleinman SH, Nemo GJ. Current and emerging infectious risks of blood transfusions. JAMA 2003;289:959-62.
9. Stramer SL, Glynn SA, Kleinman SH, et al. Detection of HIV-1 and HCV infections among antibody-negative blood donors by nucleic acid–amplification testing. N Engl J Med 2004;351:760-8.

Where Does Fitness Fit In?

William E. Kraus, M.D., and Pamela S. Douglas, M.D.

Cardiorespiratory fitness, as measured by a number of relatively simple and inexpensive clinical maneuvers, provides strong and independent prognostic information about the overall risk of illness and death, especially that from cardiovascular causes. This relationship extends to men, women, and adolescents. It is valid in apparently healthy persons; in patients with a broad range of maladies, including several types of cancer and cardiovascular disease; and in at-risk patients with diabetes mellitus, the metabolic syndrome, and hypertension.[1-3] However, despite the profoundly important prognostic information provided by simple clinical assessments of fitness, they are rarely used in the clin-

ical setting and often ignored in the exercise-testing laboratory. The article by Gulati et al. in this issue of the *Journal* encourages us to rediscover the power of fitness in both apparently healthy populations and symptomatic referral populations.[4]

For fitness determinations to have broad clinical usefulness, normative values must be developed for commonly assessed demographic groups. Although normative values for men are well established, Gulati et al. address the historical lack of data for women. They present a nomogram for the assessment of normative fitness values for women across a broad age range, derived from data obtained as part of the community-based St. James Women Take Heart Project and validated in a multicenter, symptomatic referral population. This nomogram, similar to that previously obtained for men, now provides clinicians with a means of assessing a woman's relative fitness level using a readily obtainable, inexpensive clinical test. These data inform the care provider and the patient about the relative risk of disease and progression as well as the overall prognosis.

The clinical significance of the findings of Gulati et al. is greatly enhanced by the ease with which fitness can be measured. The exercise treadmill test is accessible and inexpensive, and most patients find it acceptable. The duration of exercise or estimated energy expenditure (expressed as metabolic equivalents, or MET, as used in this study, based on the speed and grade of the treadmill and the body mass of the subject) can readily be assessed. The authors' use of MET to construct their nomogram, as opposed to the time to exhaustion in a single standardized protocol, allows testing to be customized to accommodate the individual subject's ability to exercise and the clinician's preference as to the protocol used in the clinical setting.

A fitness determination puts powerful prognostic information into the hands of clinicians. Such information is of little clinical value unless it can be used to guide an effective intervention that substantially alters risk. Fortunately, in the case of cardiorespiratory fitness, such an intervention is readily available. Men who improve their fitness (as assessed by the duration of exercise) also decrease their cardiovascular risk. In one report, men had an 8 percent reduction in long-term cardiovascular risk for every minute of improvement in exercise capacity.[5] The subjects in this observational study effectively halved their relative risk of death by abiding by the physical-activity recommendations of the American College of Sports Medicine and the Centers for Disease Control and Prevention: 30 minutes per day of at least moderate exercise on most or all days of the week.[6] Recent observational data from the Nurses' Health Study and the Women's Health Study imply that similar relationships among physical activity, exercise, fitness, and outcome also hold for women and that even a small amount of exercise is protective.[7] Since there are many determinants of cardiorespiratory fitness besides habitual exercise (e.g., genetics accounts for up to 50 percent of baseline fitness),[8] it will be important to prove prospectively that regular exercise among relatively unfit persons actually prevents events in proportion to changes in fitness.

Although the currently available data do not demonstrate that fitness should be measured at each clinic visit, a patient's exercise history certainly can and should be obtained routinely. In a paradigm shift, many experts are advocating expanding the vital signs obtained during virtually all visits to include a risk behavior: smoking status. Given the strong effect of fitness on the risk of death from any cause, the risk of disability, the risk of death from cardiovascular causes, and longevity, perhaps assessment of this variable should also be made a routine component of all clinical assessments, by including a simple question regarding exercise habits during each office visit. Counseling sedentary persons to participate in an exercise program is acknowledged to be an important component of cardiovascular-risk management. As has been learned from efforts to reduce smoking, however, unless clinicians make it a regular practice to ask the relevant question and gather the data, they will fail to provide the consistent guidance and encouragement essential to changing patients' unhealthy lifestyle choices.

Exercise testing is routinely used to evaluate patients with symptoms. However, many exercise testing laboratories and clinicians focus on electrocardiographic findings and minimize the prognostic importance of the duration of exercise or total energy expenditure (as expressed in MET). Mark and Lauer decried this "widespread tendency to ignore exercise capacity in clinical management" in favor of focusing on ischemia and the prediction of angiographic anatomy.[9] Extremely poor exercise capacity (such as observed in the cohorts evaluated by Gulati et al., in which some women had maximal exercise capacities of 1.2 to 1.4 MET) portends an extremely poor outcome. The predictive accuracy

of ST-segment deviation is highly dependent on the duration of exercise. More broadly, it is imperative to concentrate on the overall patient in addition to the patency of the patient's coronary vasculature. Just as we now know that predicting coronary events depends not only on the presence of a single vulnerable plaque that is prone to rupture, but also on an expanded concept of the "vulnerable patient," so should we pay very close attention to the systemwide information reflected by a poor level of fitness and not focus solely on regional myocardial ischemia.

Gulati et al. observe that cardiorespiratory fitness deteriorates by approximately 1 percent per year of age. The knowledge of this decline is not new, but previous information was derived almost entirely from studies in men. By providing, for the first time, age-dependent nomograms for women, Gulati et al. identify differences in aging between the sexes (although whether these are significantly different is not addressed). This finding is important, since men and women of similar age and body weight differ in their exercise capacity or cardiorespiratory fitness, irrespective of whatever measure is used, as well as in other measurements of exercise physiology. Further research is needed to understand why the age-related deterioration in cardiorespiratory fitness is greater in women than in men.

There are a few limitations to the analysis provided by Gulati et al. As noted above, the authors did not provide a statistical comparison of the normative values for men and women to clarify the importance of this difference. In addition, they chose a cutoff point of 85 percent of the age-predicted exercise capacity to define populations for assessment of the hazard ratios for death from any cause and death from cardiovascular causes. Other than the obvious parallel to the 85 percent of predicted maximal heart rate often used to judge the adequacy of exercise stress, it is unclear why this value was chosen as a critical threshold. No sensitivity analysis is provided as justification.

Despite these concerns, the report by Gulati et al. provides important new normative values for women regarding exercise capacity and cardiorespiratory fitness and documents a strong relationship between these variables and the risk of death from any cause and from cardiovascular causes. By confirming the critical importance of fitness to health and longevity, the authors place a responsibility on clinicians to use this information for their patients' benefit. Given that such data are readily obtainable in the clinical setting (and often ignored when available on routine exercise testing), compliance should be fairly easy and inexpensive and represents an effective improvement in practice. We hope that this report will provide a stimulus to reintroduce fitness assessments into the routine clinical environment for both women and men.

From the Division of Cardiovascular Medicine, Duke University Medical Center, Durham, N.C.

1. Department of Health and Human Services. Physical activity and health: a report of the Surgeon General. Atlanta: President's Council on Physical Fitness & Sports, National Center for Chronic Disease Prevention and Health Promotion, 1996.
2. Lee S, Kuk JL, Katzmarzyk PT, Blair SN, Church TS, Ross R. Cardiorespiratory fitness attenuates metabolic risk independent of abdominal subcutaneous and visceral fat in men. Diabetes Care 2005; 28:895-901.
3. Blair SN, Kohl HW III, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: a prospective study of healthy men and women. JAMA 1989;262:2395-401.
4. Gulati M, Black HR, Shaw LJ, et al. The prognostic value of a nomogram for exercise capacity in women. N Engl J Med 2005;353: 468-75.
5. Blair SN, Kampert JB, Kohl HW III, et al. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. JAMA 1996;276:205-10.
6. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995;273:402-7.
7. Hu FB, Willett WC, Li T, Stampfer MJ, Colditz GA, Manson JE. Adiposity as compared with physical activity in predicting mortality among women. N Engl J Med 2004;351:2694-703.
8. Bouchard C, Rankinen T, Chagnon YC, et al. Genomic scan for the maximal oxygen uptake and its response to training in the HERITAGE Family Study. J Appl Physiol 2000;88:551-9. [Erratum, J Appl Physiol 2004;96:829.]
9. Mark DB, Lauer MS. Exercise capacity: the prognostic variable that doesn't get enough respect. Circulation 2003;108:1534-6.

Shift-Work Sleep Disorder — The Glass Is More Than Half Empty

Robert C. Basner, M.D.

As many as 20 percent of workers in industrialized nations are shift workers — in other words, people who work either at night or on rotating shifts.[1] Shift-work sleep disorder, defined as a primary complaint of insomnia or excessive sleepiness temporally associated with a work period that occurs during the habitual sleep phase,[2] has been diagnosed in as many as 10 percent of shift workers.[3] Shift

workers overall appear to be at increased risk for peptic ulcer disease, coronary heart disease, insulin resistance, and the metabolic syndrome,[4,5] as well as for sleep deprivation, depression, sleepiness-related accidents, and curtailed family and social activities.[3,4] People who receive the diagnosis of shift-work sleep disorder appear to have a higher morbidity rate than do those without such a diagnosis.[3] A high proportion of shift workers are employed in the health care and transportation industries.

Therefore, a large-scale, well-designed investigation of treatment for sleepiness related to shift work commands considerable attention. Such a study, by Czeisler and colleagues, who are members of the U.S. Modafinil in Shift Work Sleep Disorder Study Group, appears in this issue of the *Journal*.[6] The investigators present a double-blind, placebo-controlled study involving 204 subjects with shift-work sleep disorder who were randomly assigned to either a three-month trial of the wakefulness-promoting drug modafinil at 200 mg daily or placebo; 153 subjects completed the study. Subjects received the study medication before each night shift. It is important to note that the study design allowed for the exclusion of other primary sleep disorders, such as obstructive sleep apnea. The major findings were that modafinil was associated with significant improvements in laboratory measurements of nighttime sleepiness and vigilance, as compared with placebo, and that there was an increased likelihood that subjects taking modafinil would have clinical improvement.

However, enthusiasm for these results should be tempered by putting into perspective what the current research is specifically designed to assess and what it actually finds. It is, basically, a narrowly focused investigation that concerns itself with whether a specific wakefulness-promoting agent, modafinil, is able to improve nocturnal alertness and vigilance in objectively sleepy shift workers during laboratory testing. Modafinil, recently approved by the Food and Drug Administration for the treatment of excessive sleepiness associated with shift-work sleep disorder, probably enhances wakefulness by interactions of adrenergic and dopaminergic systems.[7] At a 200-mg daily dose during the course of four consecutive simulated night shifts among non–shift-worker volunteers, the drug has attenuated sleepiness and neurobehavioral deficits occurring during the hours of the simulated night shifts.[8]

However, promotion of wakefulness and vigilance, as well as inhibition of sleep, is subserved by multiple neuromodulator and neurotransmitter systems,[7] and it is reasonable to postulate that other wakefulness-promoting agents would show similar efficacy in a comparison with placebo, modafinil, or both in shift workers. Improvements in performance and vigilance during the morning hours following more than 40 continuous hours of being awake among normal adults has been found to be similar for modafinil and 600 mg of caffeine,[9] and low-dose caffeine has improved cognitive performance and the ability to remain awake when compared with placebo among healthy adults during extended enforced time awake and circadian desynchrony.[10] In the current study, modafinil did not induce measurable effects on circadian rhythm. Taken together, neither these data nor any other published studies provide evidence to indicate that modafinil is uniquely suited to be used as an enhancer of wakefulness and vigilance in humans subjected to nighttime shift work.

The current study was not designed to investigate the effects of modafinil among subjects who would be defined as having shift-work sleep disorder on the basis of symptoms[2,3]; rather, the investigators selected a small group of subjects with shift-work sleep disorder who had both objectively measured severe nighttime sleepiness and decreased daytime sleep efficiency. In a similar vein, by design, this study does not address many of the major morbidities associated with shift work, nor does it allow an understanding of the potential effects of wakefulness-promoting agents as part of a larger strategy for treating disorders related to shift work. Such a strategy could rationally include behavioral and pharmacologic treatment of the characteristic shortened diurnal sleep times among shift workers, healthy-lifestyle training, provision of breaks,[10] a more fluid scheduling of shifts,[11,12] shift-work naps, manipulation of circadian rhythms, and assessment for and treatment of concomitant sleep disorders that may be contributing to morbidity among some shift workers.[13]

Furthermore, safety risks associated with shift-work sleep disorder have been shown to accrue over the increasing length of shifts, as well as length of time working these shifts, whereas such risks can decrease almost linearly over the course of a single night shift.[14] These clinically significant aspects of night-shift work in relation to potential interactions with a wakefulness-promoting pharmacolog-

ic agent were not addressed by the current study design. It is simplistic to consider that a pill alone could sufficiently modify the effects of this disorder.

How may the efficacy and safety data offered by this study be interpreted? Shift workers receiving modafinil improved from falling asleep within an average of 2.1 minutes during four nocturnal nap opportunities to falling asleep within an average of 3.8 minutes, whereas there was no statistically significant change among the placebo users. The subjects receiving modafinil also showed significant improvements during nighttime vigilance testing, including reduced attention lapses, as compared with their baseline scores. Subjects receiving placebo had worsened performance on this testing. However, limited reliability and validity data exist for sleepiness and vigilance testing in these circumstances, and similar concerns exist regarding the Clinical Global Impression of Change metric used for this investigation; it remains unclear to what extent this format is a valid, disease-specific scale for use in clinical trials among shift workers.[15] Thus, there is no reliable way to know how the improvements found with modafinil during laboratory testing are potentially significant clinical improvements in health, safety, or productivity among shift workers. The results documented here with the modafinil regimen, however, suggest that these workers would still be considered greatly impaired regarding vigilance and wakefulness during the nights they would need to remain awake and alert, as is pointed out by the investigators. Furthermore, although modafinil was associated with self-reports of fewer accidents or near accidents than was placebo, these data were not clearly corroborated, and no productivity or safety data were collected on the job during the time the study was in progress.

Regarding the safety of the drug itself, no adequately powered, randomized, controlled study has documented long-term efficacy or safety of this medication among shift workers, who may have substantial cardiovascular disease.[8] It is of concern, however, that modafinil in this study was associated with increased insomnia, as compared with placebo, thus apparently worsening one of the defining criteria of shift-work sleep disorder even as it improved nighttime sleepiness and vigilance.

What, then, can one conclude from the current study, in the context of other studies regarding health disorders associated with shift work? Reduction of sleepiness and improvement of vigilance are clearly a major therapeutic imperative for shift workers, but the investigators' own most robust conclusion from this study appears to be that "modafinil is of some value in the clinical management of sleepiness associated with shift-work sleep disorder."[6] Modafinil has a reasonable safety profile to date, and it may well be that it will be shown to be an effective and safe adjunct to comprehensive treatment strategies for shift-work sleep disorder. But the current study does not adequately assess the clinical value of this particular drug in shift-work sleep disorder, nor does it justify writing more prescriptions for modafinil. Rather, it serves as a wake-up call for the design and implementation of further scientific studies to address in a cohesive manner the serious health and safety issues that surround us by virtue of our having become, to a large extent, a shift-working society.

From the Columbia University College of Physicians and Surgeons, New York.

1. Eastman CI, Boulos Z, Terman M, Campbell SS, Dijk DJ, Lewy AJ. Light treatment for sleep disorders: consensus report. VI. Shift work. J Biol Rhythms 1995;10:157-64.
2. The international classification of sleep disorders: diagnostic and coding manual. Rev. ed. Rochester, Minn.: American Sleep Disorders Association, 1997:121-4.
3. Drake CL, Roehrs T, Richardson G, Walsh JK, Roth T. Shift work sleep disorder: prevalence and consequences beyond that of symptomatic day workers. Sleep 2004;27:1453-62.
4. Knutsson A. Health disorders of shift workers. Occup Med (Lond) 2003;53:103-8.
5. Karlsson B, Knutsson A, Lindahl B. Is there an association between shift work and having a metabolic syndrome? Results from a population based study of 27,485 people. Occup Environ Med 2001; 58:747-52.
6. Czeisler CA, Walsh JK, Roth T, et al. Modafinil for excessive sleepiness associated with shift-work sleep disorder. N Engl J Med 2005;353:476-86.
7. Boutrel B, Koob GF. What keeps us awake: the neuropharmacology of stimulants and wakefulness-promoting medications. Sleep 2004;15:1181-94.
8. Walsh JK, Randazzo AC, Stone KL, Schweitzer PK. Modafinil improves alertness, vigilance, and executive function during simulated night shifts. Sleep 2004;27:434-9.
9. Wesensten NJ, Belenky G, Kautz MA, Thorne DR, Reichardt RM, Balkin TJ. Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. Psychopharmacology (Berl) 2002;159:238-47.
10. Wyatt JK, Cajochen C, Ritz-De Cecco A, Czeisler CA, Dijk DJ. Low-dose repeated caffeine administration for circadian-phase-dependent performance degradation during extended wakefulness. Sleep 2004;27:374-81.
11. Kerin A, Aguirre A. Improving health, safety, and profits in extended hours operations (shiftwork). Ind Health 2005;43:201-8.
12. Knauth P, Hornberger S. Preventive and compensatory measures for shift workers. Occup Med (Lond) 2003;53:109-16.
13. Sherman BW, Strohl KP. Management of shift work disorder: Alice in Wonderland redux? J Occup Environ Med 2004;46:1010-2.
14. Folkard S, Tucker P. Shift work, safety and productivity. Occup Med (Lond) 2003;53:95-101.
15. Schneider LS, Olin JT, Doody RS, et al. Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change: the Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11:Suppl 2:S22-S32.

CLINICAL IMPLICATIONS OF BASIC RESEARCH

An Aggrecanase and Osteoarthritis

Gerard Karsenty, M.D., Ph.D.

The study of degenerative diseases is often neglected by basic scientists, probably because these complicated pathophysiological conditions are considered to be mere consequences of aging. But there is an urgent need to understand the fundamental mechanism of these diseases, thus allowing the rational development of therapies. Mouse models provide a way to identify otherwise elusive mechanisms of disease because they allow us to determine the in vivo function of a given protein in a given type of cell at a particular time. A pair of recent studies by Stanton and colleagues[1] and Glasson and colleagues[2] realizes this goal in the context of osteoarthritis.

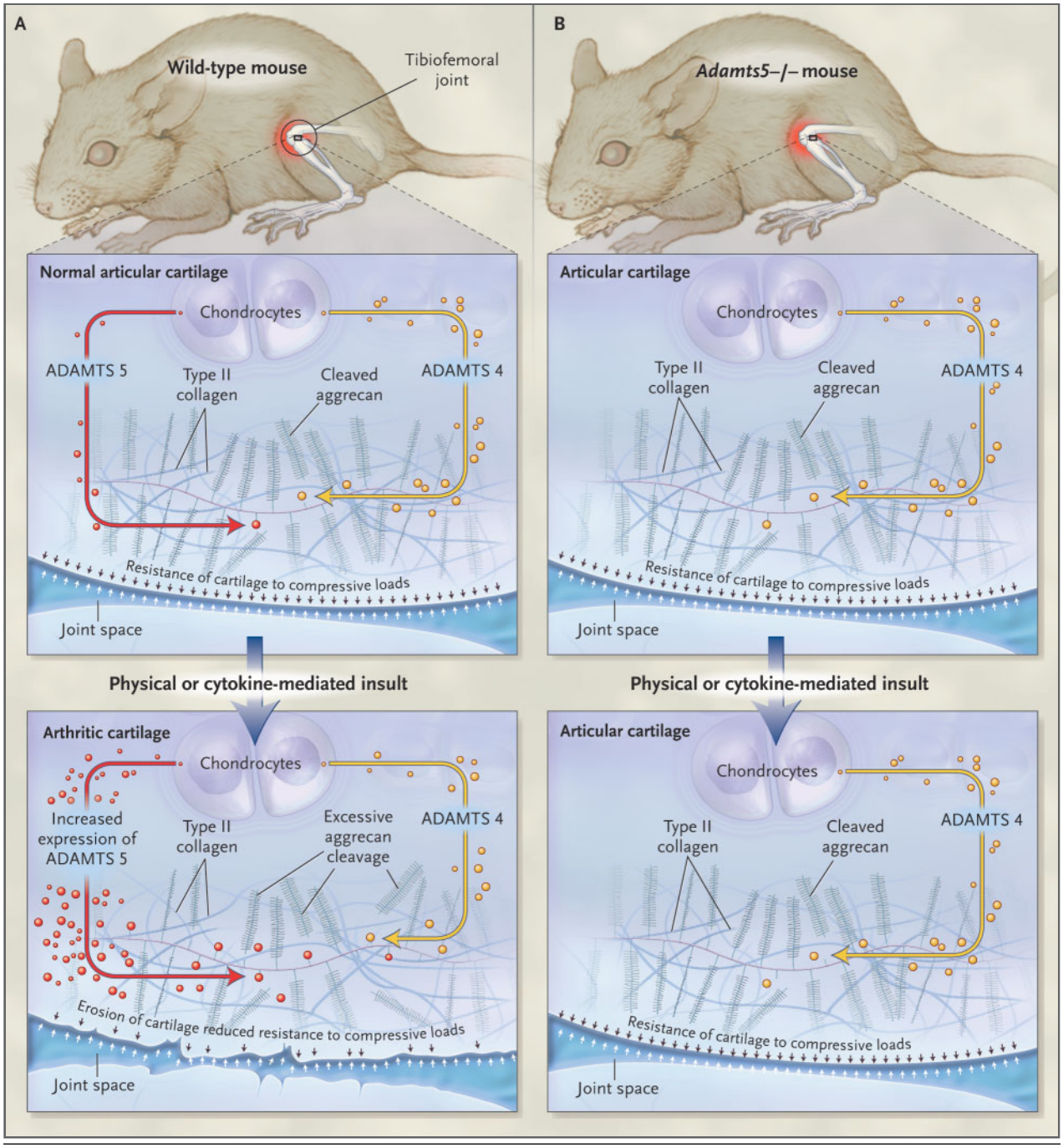
The two main constituents of the cartilaginous extracellular matrix are a type II collagen–rich collagenous network, which provides tensile strength, and a cartilage-specific proteoglycan called aggrecan, which is highly hydrated and thereby allows cartilage to resist a compressive load (Fig. 1). Under physiologic conditions, this cartilaginous extracellular matrix is constantly remodeled through degradation followed by the synthesis of collagen and aggrecan to maintain the integrity of cartilage. In osteoarthritis and inflammatory arthritis, the degeneration of the extracellular matrix far exceeds its synthesis. The extracellular matrix of cartilage wears away, exposing articular cartilage and, eventually, bone. The loss of aggrecan is the primary event leading to the destruction of cartilage, and so an understanding of how aggrecan is normally degraded and whether this process can be forestalled is critical to preventing osteoarthritis.

The articles by Stanton et al.[1] and Glasson et al.[2] provide insight into this process in a simple and convincing way. To date, three enzymes capable of degrading aggrecan have been identified: ADAMTS 1, ADAMTS 4, and ADAMTS 5. The inactivation of ADAMTS 1 had been shown not to protect mice from experimental arthritis, but the importance of ADAMTS 4 and ADAMTS 5 was not known. To address this issue, both groups created mice lacking either ADAMTS 4 or ADAMTS 5. They found that the inactivation of either enzyme did not affect skeletal development or the integrity of articular cartilage in unchallenged mice. Unfazed by this apparently negative result, both groups challenged the mutant mice. One group surgically induced unilateral joint instability,[1] and the other used a model of inflammatory arthritis.[2] The results were essentially identical: ablation of ADAMTS 5 was sufficient to protect against the erosion of cartilage and the occurrence of arthritis (Fig. 1), whereas deletion of ADAMTS 4 had no protective effect. The latter result is consistent with the observation that in vitro, ADAMTS 4 is not the main aggrecanase in mouse articular cartilage. Thus, the experiments in toto point to a single molecule — ADAMTS 5 — as a precise pharmacologic target to prevent osteoarthritis.

These experiments illustrate the elucidation of

Figure 1 (facing page). Control of the Pathophysiological Process Affecting Articular Cartilage in Osteoarthritis.

A pair of recent studies[1,2] used mouse models of osteoarthritis to show that the enzymatic activity of ADAMTS 5, but not that of ADAMTS 4, mediates aggrecan-induced turnover of articular cartilage after physical or cytokine-mediated injury. Under normal conditions (Panel A), cartilaginous extracellular matrix in wild-type mice is constantly remodeled through the degradation and synthesis of collagen and aggrecan. Aggrecan helps cartilage to resist compressive loads. Mice lacking either ADAMTS 4 or ADAMTS 5 (Panel B) develop normally. Physical injury to the meniscotibial ligament or the injection of interleukin-1α into the knee joint[1] (which induces a model of inflammatory osteoarthritis)[2] causes the degradation of cartilage in wild-type mice (Panel A) and mice lacking ADAMTS 4 (*Adamts4*–/–) as a result of a dramatic increase in the expression of ADAMTS 5, leading to excessive cleavage of aggrecan and the erosion of cartilage. In contrast, mice lacking ADAMTS 5 (*Adamts5*–/–) are protected against the degradation and erosion of cartilage (Panel B).



the molecular basis of a complex degenerative disease. The knowledge derived from these experiments can now be used to design novel, adapted therapies for osteoarthritis. If ADAMTS 5 has an identical function in humans — which seems likely — treatment with an inhibitor may prevent the disease or quickly stop its progression.

From the Department of Molecular and Human Genetics, Baylor College of Medicine, Houston.

1. Stanton H, Rogerson FM, East CJ, et al. ADAMTS5 is the major aggrecanase in mouse cartilage in vivo and in vitro. Nature 2005; 434:648-52.
2. Glasson SS, Askew R, Sheppard B, et al. Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. Nature 2005;434:644-8.

CORRESPONDENCE

Screening for the Lynch Syndrome

TO THE EDITOR: Hampel and colleagues (May 5 issue)[1] suggest that a universal screening program for the detection of microsatellite instability in patients with colorectal cancer is feasible and probably desirable. However, it will not be inexpensive. We have estimated that such a screening program would cost in excess of $57,000 per germ-line–mutation carrier detected.[2] Furthermore, the benefits to the proband are modest, since neither treatment nor outcome is likely to change on the basis of test results. The cost-effectiveness of such a program thus depends heavily on the ability to locate and test relatives, since they have the most to gain from early initiation of colorectal-cancer screening. Hampel and colleagues do not report the total number of living relatives of the 23 probands with a deleterious mutation; they report only the number who received counseling and testing. Finally, screening strategies that start with evaluation of probands' family history, such as the Bethesda guidelines,[3] appear to be much more cost-effective than universal screening for microsatellite instability.[2]

Scott D. Ramsey, M.D., Ph.D.
Fred Hutchinson Cancer Research Center
Seattle, WA 98109
sramsey@fhcrc.org

1. Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). N Engl J Med 2005;352:1851-60.
2. Ramsey SD, Burke W, Clarke L. An economic viewpoint on alternative strategies for identifying persons with hereditary nonpolyposis colorectal cancer. Genet Med 2003;5:353-63.
3. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261-8.

THE AUTHORS AND A COLLEAGUE REPLY: Dr. Ramsey questions the validity of our strategy of universal screening mainly because of the perceived high cost relative to the benefit. The claim that benefits to the proband are modest is incorrect: the identification of mutation-positive Lynch syndrome dramatically alters management and outcome.[1] We agree that the number of relatives tested influences the cost-effectiveness of screening. Our 23 probands had 148 first- and second-degree relatives over the age of 18 years who were at high risk. Of 117 relatives tested, 52 were mutation-positive and 65 were mutation-negative. Intensified clinical surveillance of mutation-positive persons significantly reduces cancer-related morbidity and mortality.[2] Relaxation of clinical surveillance in those who are a priori at high risk and yet are found to be mutation-negative is highly beneficial. Unfortunately, too many carriers of the Lynch syndrome mutation are missed even by good evaluation of the family history, and those with strong family histories are often not referred for genetic evaluation.[3] We calculate that our simplified method of immunohistochemical analysis as the first screen will reduce costs in terms of life-years saved. Finally, Dr. Ramsey's methodology (decision analysis and simulation), which resulted in his estimate of $57,000 per germ-line mutation de-

THIS WEEK'S LETTERS

524 Screening for the Lynch Syndrome
525 *Chlamydia pneumoniae* and Acute Coronary Syndrome
528 Histone Deacetylase Activity and COPD
529 Nephrogenic Syndrome of Inappropriate Antidiuresis
530 MRSA in the Community
532 Transatlantic Spread of the USA300 Clone of MRSA

tected, is not validated and is too unstable to predict economic outcomes. In his article, Dr. Ramsey acknowledges the instability of the research approach and execution.[4]

Heather Hampel, M.S.
Albert de la Chapelle, M.D., Ph.D.
W.C. Benton, M.B.A., Ph.D.

Ohio State University
Columbus, OH 43210
delachapelle-1@medctr.osu.edu

1. Lanspa SJ, Jenkins JX, Cavalieri RJ, et al. Surveillance in Lynch syndrome: how aggressive? Am J Gastroenterol 1994;89:1978-80.
2. Järvinen HJ, Aarnio M, Mustonen H, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology 2000;118:829-34.
3. Sweet KM, Bradley TL, Westman JA. Identification and referral of families at high risk for cancer susceptibility. J Clin Oncol 2002; 20:528-37.
4. Ramsey SD, Burke W, Clarke L. An economic viewpoint on alternative strategies for identifying persons with hereditary nonpolyposis colorectal cancer. Genet Med 2003;5:353-63.

Chlamydia pneumoniae and Acute Coronary Syndrome

TO THE EDITOR: Long-term antichlamydial antibiotic therapy did not alter the risk of cardiac events among patients with stable coronary disease, as reported by Grayston et al.,[1] or after an acute coronary syndrome, as reported by Cannon et al.[2] (April 21 issue). Unfortunately, neither of these studies used baseline serologic tests during patient selection. Consequently, only one third of patients had IgA antibody,[1] and only 3.5 percent had chlamydial DNA in peripheral mononuclear blood cells.[2] In the absence of persistent chlamydial infection among the majority of treated patients, there is little reason to expect a clinical benefit with antichlamydial treatment.

Serum IgA has been correlated with the presence of chlamydial DNA in tissue.[3] Furthermore, we found that patients with an acute coronary syndrome who had a baseline IgA titer of 1:32 or higher were at significant risk for the development of myocardial necrosis associated with their index event.[4] This finding highlights the biologic role and prognostic value of IgA, which may serve as a useful tool for patient selection in future clinical trials of antichlamydial treatment in coronary heart disease.

Brian Y.L. Wong, M.D.

Sudbury Regional Hospital
Sudbury, ON P3E 3Y9, Canada

Judy Gnarpe, Ph.D.

University of Alberta
Edmonton, AB T6G 2H7, Canada

1. Grayston JT, Kronmal RA, Jackson LA, et al. Azithromycin for the secondary prevention of coronary events. N Engl J Med 2005; 352:1637-45.
2. Cannon CP, Braunwald E, McCabe CH, et al. Antibiotic treatment of *Chlamydia pneumoniae* after acute coronary syndrome. N Engl J Med 2005;352:1646-54.
3. Gnarpe J, Sparr A, Naas J, Lundback A. Serological analysis of specific IgA to Chlamydia pneumoniae: increased sensitivity of IgA antibody detection using prolonged incubation and high antigen concentration. APMIS 2000;108:357-62.
4. Wong BYL, Gnarpe J, Teo KK, et al. Does chronic Chlamydia pneumoniae infection increase the risk of myocardial injury? Insights from patients with non-ST-elevation acute coronary syndromes. Am Heart J 2002;144:987-94.

TO THE EDITOR: There is a substantial body of evidence that chlamydia contributes to the pathogenesis of atherosclerosis. The two large placebo-controlled clinical trials by Grayston et al. and Cannon et al., which involved treatment with either azithromycin or gatifloxacin, demonstrated that neither of these antibiotics modulates advanced atherosclerotic disease. *Chlamydia pneumoniae* is an intracellular microorganism with a complex life cycle in which the replicating agent can convert to a persistent, nonreplicating form (cryptic bodies) on exposure to a variety of stimuli, including antibiotics.[1] The current mode of action of most antibiotics used as single-agent therapy will have minimal, if any, effect on the persistent, nonreplicating form of *C. pneumoniae*.[2] The importance of combination antimicrobial therapy that inhibits various bacterial targets in human chlamydial-related disease has been demonstrated recently in a prospective trial involving patients with spondyloarthropathy ($P<0.003$ for the comparison between patients who had a response to therapy and those who did not have a response).[3] Ideally, future trials should be based on an efficacious animal model at the same stage of pathogenesis as that of the proposed patient population and that is based on in vitro evidence of cryptic-body elimination as well as evidence

of drug penetration into the key cellular elements of atheromatous plaques.

William M. Mitchell, M.D., Ph.D.
Charles W. Stratton, M.D.

Vanderbilt University School of Medicine
Nashville, TN 37232
bill.mitchell@vanderbilt.edu

1. Gieffers J, Rupp J, Gebert A, Solbach W, Klinger M. First-choice antibiotics at subinhibitory concentrations induce persistence of Chlamydia pneumoniae. Antimicrob Agents Chemother 2004;48:1402-5.
2. Suchland RJ, Geisler WM, Stamm WE. Methodologies and cell lines used for antimicrobial susceptibility testing of Chlamydia spp. Antimicrob Agents Chemother 2003;47:636-42.
3. Carter JD, Valeriano J, Vasey FB. Doxycycline versus doxycycline and rifampin in undifferentiated spondyloarthropathy, with special reference to chlamydia-induced arthritis: a prospective, randomized 9-month comparison. J Rheumatol 2004;31:1973-80.

TO THE EDITOR: Cannon and colleagues compare gatifloxacin with placebo for the secondary prevention of cardiovascular events. This trial provides a unique opportunity to evaluate the safety of gatifloxacin in a high-risk population. The authors discuss associations between gatifloxacin therapy and diarrhea, nausea, vomiting, new-onset diabetes, hyperglycemia, and hypoglycemia. Gatifloxacin has also been associated with torsades de pointes and other ventricular arrhythmias.[1-3] Did the rates of torsades de pointes, ventricular arrhythmias, or sudden death differ between the two groups?

Gatifloxacin provided no significant benefit in terms of any of the predefined primary or secondary end points. However, in the gatifloxacin group there were trends toward increases in the rate of death from all causes (hazard ratio, 1.23; 95 percent confidence interval, 0.85 to 1.82) and of death from coronary artery disease (hazard ratio, 1.55; 95 percent confidence interval, 0.91 to 2.7) (Fig. 3 of the article by Cannon et al.). How many deaths in each group were attributable to potential adverse drug effects, including hyperglycemia, hypoglycemia, torsades de pointes, ventricular arrhythmias, sudden death, hypersensitivity reactions, *Clostridium difficile* colitis, or seizures?[3] The study drug was administered for 10 days each month. How many deaths in each group occurred on a day that the study drug was administered?

Richard Frothingham, M.D.

Veterans Affairs Medical Center
Durham, NC 27705
richard.frothingham@duke.edu

Dr. Frothingham reports having received honoraria from Bayer, Bristol-Myers Squibb, Ortho-McNeil, and Pfizer and having served as a consultant for Bayer, Ortho-McNeil, Otsuka, and Schering-Plough.

1. Frothingham R. Rates of torsades de pointes associated with ciprofloxacin, ofloxacin, levofloxacin, gatifloxacin, and moxifloxacin. Pharmacotherapy 2001;21:1468-72.
2. Bertino JS Jr, Owens RC Jr, Carnes TD, Iannini PB. Gatifloxacin-associated corrected QT interval prolongation, torsades de pointes, and ventricular fibrillation in patients with known risk factors. Clin Infect Dis 2002;34:861-3.
3. Tequin (gatifloxacin). Princeton, N.J.: Bristol-Myers Squibb, November 2004 (package insert).

TO THE EDITOR: Cannon et al. and Grayston et al. appraised the effect of long-term antibiotic therapy for protection against cardiovascular disease and concluded that there was no benefit. However, these studies do not justly reflect the overall effect of the intervention. A major adverse event of prolonged antibiotic treatment — namely, the development of antibiotic resistance — has been neglected.

Assessment of efficacy entails the balance of benefit versus harm. Antibiotics are unique in that the major cost of their use is the development of resistance, which affects both the person treated and the environment. Thus, the intermittent administration of gatifloxacin for 18 months may be harmful. The claim that the use of azithromycin for the treatment of cardiovascular disease is prevalent, thereby justifying the (unassessed) risk of the trial, is poorly based.[1]

Pharmaceutical companies drive these studies, but it is clinicians' duty to impose a complete assessment of adverse events. These two trials are in line with previous meta-analyses[2,3] that showed, with narrow confidence intervals, that there was no benefit associated with the use of antibiotics for cardiovascular protection. Together, these studies amount to 6270 patient-years of antibiotic exposure, with unknown ecologic impact. In the absence of such data, further trials are unjustified.

Mical Paul, M.D.
Abigail Fraser, M.P.H.
Leonard Leibovici, M.D.

Rabin Medical Center
Petah-Tikva 49100, Israel
pil1pel@zahav.net.il

1. Gimenez-Sanchez F, Butler JC, Jernigan DB, et al. Treating cardiovascular disease with antimicrobial agents: a survey of knowledge, attitudes, and practices among physicians in the United States. Clin Infect Dis 2001;33:171-6.
2. Wells BJ, Mainous AG III, Dickerson LM. Antibiotics for the sec-

ondary prevention of ischemic heart disease: a meta-analysis of randomized controlled trials. Arch Intern Med 2004;164:2156-61.
3. Illoh KO, Illoh OC, Feseha HB, Hallenbeck JM. Antibiotics for vascular diseases: a meta-analysis of randomized controlled trials. Atherosclerosis 2005;179:403-12.

DRS. GRAYSTON AND CANNON REPLY: Drs. Mitchell and Stratton point out the potential difficulty in attempting to treat chronic chlamydial infections that may be due in part to transiently nonreplicating antibiotic-resistant forms of *C. pneumoniae*. Because our experience and the literature have shown the need for prolonged courses of antibiotics for the successful treatment of chronic chlamydial infections, such as trachoma, in both the ACES (Azithromycin for the Secondary Prevention of Coronary Events) and PROVE IT–TIMI 22 (Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis in Myocardial Infarction 22) trials we elected to use long courses (one to two years) of antibiotic therapy. The failure of these courses of antibiotics suggests that at this time there is no antibiotic intervention that can be practically administered and that is also effective in reducing the risk of secondary coronary events. The results of these studies do not preclude the possibility that a more effective antibiotic regimen will become available in the future.

Drs. Wong and Gnarpe suggest that baseline serologic tests for chlamydia should have been used to identify participants for our trials. Baseline serologic data were collected in both studies but were not used as admission criteria because we believed that there is not sufficient evidence that either the presence or the absence of *C. pneumoniae* antibody clearly identifies adults at increased risk of coronary events. In addition, these trials offered a unique opportunity to study serologic findings as an indicator of such susceptibility. We found in both trials that the presence of *C. pneumoniae* IgG antibody did not identify participants who had an increased susceptibility to events. The presence of *C. pneumoniae* IgA antibody was assessed in the ACES trial. A total of 1301 participants had IgA antibody, and 2620 did not. IgA antibody was not associated with the risk of the primary end point in either the azithromycin group or the placebo group. Persons with IgA antibody were not shown to be more susceptible to cardiac events than were those without the antibody.

Dr. Frothingham asks about reactions to gatifloxacin. Of the 119 deaths in the PROVE IT–TIMI 22 trial, 19 occurred before the administration of the antibiotic study drug, which began on day 15. Of the remaining deaths, 50 had cardiovascular causes, as classified by the clinical events committee; of these, sudden death occurred in 9 patients in the gatifloxacin group and in 11 in the placebo group. Nonfatal ventricular arrhythmias reported as serious adverse events occurred in 16 patients in each group. Torsades de pointes was not reported as a serious adverse event in either group. Hyperglycemia and hypoglycemia were more common in the gatifloxacin group than in the placebo group, as reported in the article, but no episodes were fatal. Colitis developed in one patient receiving gatifloxacin, in whom it was necessary to stop the study drug, and in none of the patients in the placebo group. (The case of colitis was not reported to be caused by *C. difficile*.) One patient in the gatifloxacin group and none of the patients in the placebo group had seizures that led to discontinuation of the study drug. Thus, despite monthly courses of gatifloxacin for an average of two years, there did not appear to be major adverse events attributable to the known side effects of the drug. We hope this information is helpful to physicians.

Dr. Paul and colleagues raise the issue of antibiotic resistance. Prolonged courses of antibiotics are currently indicated for a variety of conditions, including acne. This safety concern was considered in the design of both trials. If a benefit of prolonged antibiotic therapy in the prevention of cardiac events had been demonstrated, further research on the possible risks associated with this treatment would be needed to assess the balance of risks and benefits. In their letter, Dr. Paul and colleagues dismiss the importance of the ongoing practice of antibiotic treatment of coronary heart disease. The review committee of the National Heart, Lung, and Blood Institute that recommended funding of the ACES trial stated as one of the reasons for their recommendation that there was "current antibiotic treatment of coronary heart disease without evidence of its effectiveness."

Both trials were initiated and planned by the investigators. All data analysis, interpretation, and reporting were performed by the investigators, independent of influence from pharmaceutical companies.

We understood when we planned these antibi-

otic trials that additional information concerning the role of *C. pneumoniae* in atherosclerosis and the treatment of chronic atherosclerotic lesions would be desirable. We believed that the possibility of an additional effective treatment for coronary heart disease was of such importance to the public health that we undertook the trials with the information then available.

J. Thomas Grayston, M.D.
University of Washington
Seattle, WA 98195
for the ACES Investigators
Christopher P. Cannon, M.D.
Brigham and Women's Hospital
Boston, MA 02115
cpcannon@partners.org
for the PROVE IT–TIMI 22 Investigators

Histone Deacetylase Activity and COPD

TO THE EDITOR: In their study, Ito and colleagues (May 12 issue)[1] observed reductions in both the activity and expression of histone deacetylases (HDACs), especially HDAC2, in patients with chronic obstructive pulmonary disease (COPD). In addition, the activity inversely correlated with the severity of COPD.

Histone-modifying enzymes play essential roles in gene regulation.[2] The balance between histone acetylation and deacetylation appears to be crucial to normal cell growth. Disruption of either of these molecular mechanisms has been associated with the development of cancer. Several molecules and genes have been identified or developed or both to inhibit HDACs.[3] Valproic acid, an antiepileptic drug that has been commercially available for decades, has been found to inhibit HDACs, including HDAC2.[4] However, there is no evidence that valproic acid worsens pulmonary function in patients taking the medicine.[5] The "chicken-and-egg" conundrum remains unresolved: Does the reduction of HDAC activity cause severe COPD, or is it a secondary event?

Albert Y. Lin, M.D., M.P.H.
Santa Clara Valley Medical Center
San Jose, CA 95128
albert.lin@hhs.co.santa-clara.ca.us

1. Ito K, Ito M, Elliott WM, et al. Decreased histone deacetylase activity in chronic obstructive pulmonary disease. N Engl J Med 2005;352:1967-76.
2. Kornberg RD, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 1999;98:285-94.
3. Marks PA, Richon VM, Miller T, Kelly WK. Histone deacetylase inhibitors. Adv Cancer Res 2004;91:137-68.
4. Kramer OH, Zhu P, Ostendorff HP, et al. The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2. EMBO J 2003;22:3411-20.
5. Physicians' desk reference. 58th ed. Montvale, N.J.: Thomson PDR, 2004.

TO THE EDITOR: The following statement in the Discussion section of the article by Ito et al. is rather confusing: "In the present study, there was a positive correlation between HDAC activity and disease severity, as measured by the percent of predicted FEV_1 [forced expiratory volume in one second]. . . ." The Results section, Figure 1D, and the abstract clearly indicate that there was apparently a negative correlation between HDAC activity and disease severity. I assume that the authors meant to say that there was a positive correlation between HDAC activity and FEV_1.

Angshu Bhowmik, M.B., B.S.
Homerton University Hospital
London E9 6SR, United Kingdom

THE AUTHORS REPLY: Histone modification regulates many genes, including those involved in normal cell growth. Eleven classic HDACs have been identified.[1] In patients with COPD, we found a marked reduction in HDAC2, with lesser reductions in HDAC5 and HDAC8. Different HDACs appear to be involved in different cellular functions and presumably regulate different sets of genes. Indeed, the targeted reduction of HDAC2 through RNA interference results in reduced responsiveness to corticosteroids in a human epithelial cell line (A549), whereas this reduction is not observed when other classic HDACs are inhibited.[2] Valproate, a nonselective inhibitor of classic HDACs, is associated with 50 percent inhibition of HDAC activity at approximately 200 μg per milliliter (1.4 mmol per liter) in A549 cells. Steady-state plasma concentrations of valproate in patients with epilepsy are 50 to 100 μg per milliliter (0.3 to 0.7 mmol per liter), so it is possible that clinical doses may have

some HDAC inhibitory effect, enhancing inflammation or reducing responsiveness to corticosteroids in patients with inflammatory diseases. There is one report of increased circulating proinflammatory cytokines in children with epilepsy treated with valproic acid.[3] However, we are not aware that a worsening of inflammatory diseases has been investigated or reported with valproate. We agree with Dr. Lin that it is not certain whether the reduction in HDAC activity is a consequence or a cause of severe COPD, but we would like to suggest that it is both and that it provides a molecular basis for the increasing pulmonary inflammation as COPD progresses.

We agree with Dr. Bhowmik that the sentence in the Discussion was incorrectly written. We should have stated that there was a negative correlation between HDAC activity and disease severity.

Peter J. Barnes, D.M., D.Sc.
Ian M. Adcock, Ph.D.
Kazuhiro Ito, Ph.D.

Imperial College
London SW3 6LY, United Kingdom
p.j.barnes@imperial.ac.uk

1. de Ruijter AJ, van Gennip AH, Caron HN, Kemp S, van Kuilenburg AB. Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J 2003;370:737-49.
2. Ito K, Adcock IM, Barnes PJ. Knock out of histone deacetylase-2 by RNA interference enhances inflammatory gene expression and reduces glucocorticoid sensitivity in human epithelial cells. Am J Respir Crit Care Med 2004;169:A847. abstract.
3. Verrotti A, Basciani F, Trotta D, Greco R, Morgese G, Chiarelli F. Effect of anticonvulsant drugs on interleukins-1, -2 and -6 and monocyte chemoattractant protein-1. Clin Exp Med 2001;1:133-6.

Nephrogenic Syndrome of Inappropriate Antidiuresis

TO THE EDITOR: The elegant work presented by Feldman et al. (May 5 issue)[1] shows that point mutations in codon 137 of the V2 vasopressin receptor (V2R) can result in either a loss-of-function mutation (R137H, which is associated with congenital nephrogenic diabetes insipidus) or a gain-of-function mutation (R137C or R137L, which is associated with the congenital nephrogenic syndrome of inappropriate antidiuresis). In the Discussion section of the article, the authors mention the possibility of an activating mutation in aquaporin-2. Because the clinical data were also consistent with this differential diagnosis, how was this possibility ruled out so as to arrive at the initial hypothesis that these infants had hyperactive V2Rs?

The authors also state that arginine vasopressin (AVP) antagonists "would probably be ineffective, given the ligand-independent nature of the lesion." However, subtle conformational changes can take place when a ligand binds to a receptor. Therefore, is it not plausible that some of these conformations could effectively down-regulate activity?

Finally, the acronym "NSIAD" (nephrogenic syndrome of inappropriate antidiuresis) might easily be mistaken for the much more familiar "NSAID" (nonsteroidal antiinflammatory drug). Perhaps renaming the syndrome "congenital pseudo-SIADH" (where SIADH denotes the syndrome of inappropriate antidiuretic hormone secretion) could be considered.

Alan Segal, M.D.

University of Vermont College of Medicine
Burlington, VT 05401
alan.segal@uvm.edu

1. Feldman BJ, Rosenthal SM, Vargas GA, et al. Nephrogenic syndrome of inappropriate antidiuresis. N Engl J Med 2005;352:1884-90.

TO THE EDITOR: Feldman et al. state that since the AVP assay is not optimized to identify low values, sequencing the V2R gene (*AVPR2*) before NSIAD is diagnosed is recommended. They also suggest that there may be additional defects in the V2R signaling cascade in patients who have an NSIAD phenotype but not a V2R mutation. However, the plasma AVP level is low in many acquired diseases in the absence of a V2R gene mutation or defects in the signaling cascade. Thus, such conclusions seem unwarranted.

AVP secretion is almost totally suppressed when the plasma osmolality falls below 275 to 280 mOsm per kilogram, so the plasma osmolality should be measured simultaneously with the plasma AVP level.[1] Moreover, the plasma AVP level is low and the plasma osmolality is normal in patients with the type D secretion pattern of SIADH.[2] Increased sensitivity to AVP or other antidiuretic substances, such as chlorpropamide or the antidiuretic substance produced in prolactinoma, may be present in such patients.[3,4] The authors might recommend se-

quencing the V2R gene after ruling out these clinical conditions.

Chao-Fu Chang, M.D.

Taipei Veterans General Hospital
Taipei 112, Taiwan

Wu-Chang Yang, M.D.

National Yang-Ming University
Taipei 112, Taiwan

Chih-Ching Lin, M.D.

Taipei Veterans General Hospital
Taipei 112, Taiwan
lincc2@vghtpe.gov.tw

1. Robertson GL, Aycinena P, Zerbe RL. Neurogenic disorders of osmoregulation. Am J Med 1982;72:339-53.
2. Hypoosmolal states — hyponatremia. In: Rose BD, Post TW. Clinical physiology of acid-base and electrolyte disorders. 5th ed. New York: McGraw-Hill, 2001:696-745.
3. Moses AM, Fenner R, Schroeder ET, Coulson R. Further studies on the mechanism by which chlorpropamide alters the action of vasopressin. Endocrinology 1982;111:2025-30.
4. Kern PA, Robbins RJ, Bichet D, Berl T, Verbalis JG. Syndrome of inappropriate antidiuresis in the absence of arginine vasopressin. J Clin Endocrinol Metab 1986;62:148-52.

THE AUTHORS REPLY: In response to Dr. Segal: evaluation of the first patient with NSIAD (Patient 2 in our article) included concurrent sequencing of the genes encoding V2R (*AVPR2*) and aquaporin-2. This patient was found to have a gain-of-function mutation in V2R (R137L) and no mutation in aquaporin-2. Subsequent *AVPR2* sequence analysis in the second patient (Patient 1) also demonstrated a missense mutation in V2R (R137C). The gene encoding aquaporin-2 has not yet been sequenced in this patient. Regarding AVP antagonists, Dr. Segal notes that "subtle conformational changes can take place when a ligand binds to a receptor" and might "effectively down-regulate activity." As discussed in a review by Seifert and Wenzel-Seifert,[1] a ligand that stabilizes a constitutively active receptor in an inactive state would be considered an "inverse agonist" rather than an antagonist. It has been reported, however, that a nonpeptide V2R antagonist also has inverse-agonist properties,[2] which may lead to the down-regulation of V2R activity in the naturally occurring constitutively active mutants. Finally, with respect to the naming of this syndrome, we had considered the term "pseudo-SIADH" but agreed with the editors and reviewers of the article, who encouraged the use of the term that was ultimately chosen for the reasons we outlined therein. To avoid confusion with "NSAID," one might want to refer to this condition as "nephrogenic SIAD."

As emphasized in our article and as suggested by Dr. Chang and colleagues, proper evaluation of patients with hyponatremia requires careful consideration of a broad differential diagnosis. The algorithm for laboratory testing during such an evaluation needs to be tailored to the specific patient and to the frequencies of the possible causes of hyponatremia. The causes discussed by Chang et al. are rarely encountered in children. Currently, it is not clear how frequently NSIAD occurs in the general population, and therefore physicians must rely on their index of suspicion for this newly identified disease. Although many conditions may be associated with hyponatremia and low plasma levels of AVP, such a presentation in a male infant with a persistent state of inappropriate antidiuresis strongly suggests the presence of NSIAD. As more epidemiologic information is acquired, the importance of sequencing the V2R gene will become increasingly apparent.

Stephen E. Gitelman, M.D.
Brian J. Feldman, M.D., Ph.D.
Stephen M. Rosenthal, M.D.

University of California at San Francisco
San Francisco, CA 94143
sgitelma@peds.ucsf.edu

1. Seifert R, Wenzel-Seifert K. Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors. Naunyn Schmiedebergs Arch Pharmacol 2002;366:381-416.
2. Morin D, Cotte N, Balestre MN, et al. The D136A mutation of the V2 vasopressin receptor induces a constitutive activity which permits discrimination between antagonists with partial agonist and inverse agonist activities. FEBS Lett 1998;441:470-5.

MRSA in the Community

TO THE EDITOR: Two articles in the April 7 issue (Fridkin et al.[1] and Miller et al.[2]) deal with the subject of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) infection. Fridkin et al. found that 6 percent of cases of MRSA infection were invasive, but they reported no cases of the toxic shock syndrome. Although MRSA strains are toxigenic,[3] MRSA-associated toxic shock syndrome appears rare[4,5] and has not been described with community-acquired strains. We report a case of

the toxic shock syndrome associated with community-acquired MRSA in an 18-year-old woman who presented with malaise, diarrhea, vomiting, and subsequently collapse. She had no risk factors for nosocomial MRSA acquisition. She was febrile, had hypotension, and was drowsy, with mucosal suffusion, widespread erythematous rash, and renal failure. Blood cultures grew MRSA that was sensitive to vancomycin, ciprofloxacin, clindamycin, gentamicin, and tetracycline. The bacterium carried the *mecA* gene and produced enterotoxins G and I. The patient received cefotaxime, flucloxacillin, and gentamicin and was switched to vancomycin after the microbiologic analysis became available; she recovered fully.

The unexpected microbiologic findings in this case raise concern about the empirical choice of antibiotics for patients without risk factors for nosocomial MRSA acquisition. However, true community-acquired MRSA infection remains uncommon,[6] and a move from beta-lactams as first-line therapy is unnecessary and may encourage resistance to vancomycin.

Ann L.N. Chapman, M.D., Ph.D.
Julia M. Greig, M.D.

Royal Hallamshire Hospital
Sheffield S10 2JF, United Kingdom
ann.chapman@sth.nhs.uk

John A. Innes, M.D.

Heartlands Hospital
Birmingham B9 5SS, United Kingdom

1. Fridkin SK, Hageman JC, Morrison M, et al. Methicillin-resistant *Staphylococcus aureus* disease in three communities. N Engl J Med 2005;352:1436-44. [Erratum, N Engl J Med 2005;352:2362.]
2. Miller LG, Perdreau-Remington F, Rieg G, et al. Necrotizing fasciitis caused by community-associated methicillin-resistant *Staphylococcus aureus* in Los Angeles. N Engl J Med 2005;352:1445-53.
3. Schmitz FJ, MacKenzie CR, Geisel R, et al. Enterotoxin and toxic shock syndrome toxin-1 production of methicillin resistant and methicillin sensitive Staphylococcus aureus strains. Eur J Epidemiol 1997;13:699-708.
4. Acland KM, Darvay A, Griffin C, Aali SAA, Russell-Jones R. Staphylococcal scalded skin syndrome in an adult associated with methicillin-resistant Staphylococcus aureus. Br J Dermatol 1999; 140:518-20.
5. Fergie JE, Purcell K. Community-acquired methicillin-resistant Staphylococcus aureus infections in south Texas children. Pediatr Infect Dis J 2001;20:860-3.
6. Salgado CD, Farr BM, Calfee DP. Community-acquired methicillin-resistant Staphylococcus aureus: a meta-analysis of prevalence and risk factors. Clin Infect Dis 2003;36:131-9.

DR. FRIDKIN AND COLLEAGUES REPLY: The case reported by Chapman et al. adds to other reports of severe community-associated MRSA infection with features that resemble the toxic shock syndrome[1,2] and emphasizes that clinicians should consider MRSA a potential pathogen in suspected *S. aureus* infections. Our data clearly suggest that MRSA infections are common, and most involve only skin and soft tissue. However, the prevalence of these infections varies geographically. There are reports from some areas of the United States in which MRSA has become a predominant cause of community-associated staphylococcal disease.[3,4] Although it may be premature to change empirical regimens for suspected *S. aureus* disease, our data support increased vigilance in the collection of appropriate diagnostic specimens and the selection of empirical agents on the basis of an understanding of local disease patterns. We agree that unnecessary use of vancomycin should be avoided. Nevertheless, in the case of a seriously ill patient who may have MRSA infection, empirical therapy with agents active against MRSA, including vancomycin and other drugs, may be necessary.

Jeffrey C. Hageman, M.H.S.

Centers for Disease Control and Prevention
Atlanta, GA 30333

Ruth Lynfield, M.D.

Minnesota Department of Health
Minneapolis, MN 55414

Scott K. Fridkin, M.D.

Centers for Disease Control and Prevention
Atlanta, GA 30333
skf0@cdc.gov

1. Mongkolrattanothai K, Boyle S, Kahana MD, Daum RS. Severe Staphylococcus aureus infections caused by clonally related community-acquired methicillin-susceptible and methicillin-resistant isolates. Clin Infect Dis 2003;37:1050-8.
2. Fergie JE, Purcell K. Community-acquired methicillin-resistant Staphylococcus aureus infections in south Texas children. Pediatr Infect Dis J 2001;20:860-3.
3. Moran GJ, Amii RN, Abrahamian FM, Talan DA. Methicillin-resistant Staphylococcus aureus in community-acquired skin infections. Emerg Infect Dis 2005;11:928-30.
4. Mishaan AM, Mason EO Jr, Martinez-Aguilar G, et al. Emergence of a predominant clone of community-acquired Staphylococcus aureus among children in Houston, Texas. Pediatr Infect Dis J 2005;24:201-6.

DR. MILLER AND COLLEAGUES REPLY: We appreciate the report by Chapman et al. regarding a case of the toxic shock syndrome caused by community-associated MRSA. The gene that encodes toxic shock syndrome toxin 1 (TSST-1) appears to be found uncommonly in community-acquired MRSA strains.[1,2] Nonetheless, this case adds to growing reports of strains of MRSA causing previously unusual clini-

cal syndromes, such as purpura fulminans,[3] necrotizing community-acquired pneumonia and pulmonary septic emboli in healthy persons who do not use injection drugs,[4] and empyema.[5]

Although Chapman et al. state that MRSA infections are "uncommon," these infections are unfortunately becoming quite prevalent globally. In many locales, including in our institutions, community-acquired *S. aureus* infections are more likely to be caused by MRSA than by methicillin-susceptible strains. Therefore, in the treatment of severe or invasive syndromes caused by *S. aureus*, it may be prudent to use vancomycin or another antibiotic that is reliably active against locally circulating MRSA strains in addition to a beta-lactam until results of cultures and susceptibilities are known.

Loren G. Miller, M.D., M.P.H.

Harbor–UCLA Medical Center
Torrance, CA 90509

Françoise Perdreau-Remington, Ph.D.

San Francisco General Hospital
San Francisco, CA 94110

Brad Spellberg, M.D.

Harbor–UCLA Medical Center
Torrance, CA 90509

1. Fey PD, Said-Salim B, Rupp ME, et al. Comparative molecular analysis of community- or hospital-acquired methicillin-resistant Staphylococcus aureus. Antimicrob Agents Chemother 2003;47:196-203.
2. Jamart S, Denis O, Deplano A, et al. Methicillin-resistant Staphylococcus aureus toxic shock syndrome. Emerg Infect Dis 2005;11:636-7.
3. Kravitz GR, Dries DJ, Peterson ML, Schlievert PM. Purpura fulminans due to Staphylococcus aureus. Clin Infect Dis 2005;40:941-7.
4. Gonzalez BE, Martinez-Aguilar G, Hulten KG, et al. Severe staphylococcal sepsis in adolescents in the era of community-acquired methicillin-resistant Staphylococcus aureus. Pediatrics 2005;115:642-8.
5. Schultz KD, Fan LL, Pinsky J, et al. The changing face of pleural empyemas in children: epidemiology and management. Pediatrics 2004;113:1735-40.

Transatlantic Spread of the USA300 Clone of MRSA

TO THE EDITOR: The emergence of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) is of great concern.[1,2] USA300, the predominant epidemic clone in numerous outbreaks in closed communities in the United States,[2] is also increasingly seen in Europe.[3] International travel and the increasing trend of training or working abroad among health care workers probably contribute to its global spread.

We describe the transfer of community-acquired MRSA from the United States to Europe and its successful eradication. A 41-year-old healthy Swiss physician performed a clinical fellowship in the United States from July 2001 to June 2003. Three months before his departure, nasal and pharyngeal swabs from the physician were negative in cultures obtained during the investigation of a case of MRSA infection.

At our institution, in Basel, Switzerland, there is routine screening of health care workers from countries with a high prevalence of MRSA. The nasal and pharyngeal swabs from the physician were found to be positive for MRSA after his return from the United States. The isolate was resistant only to beta-lactam agents and clarithromycin. Molecular typing by pulsed-field gel electrophoresis showed a banding pattern consistent with that of the USA300 strain (Fig. 1). The genotype of the staphylococcal chromosomal cassette was determined to be SCC*mec* type IV, and the Panton–Valentine leukocidin genes (*lukS-PV* and *lukF-PV*) were detected, both by polymerase-chain-reaction analysis.

After extended screening cultures (of axillary, inguinal, perirectal, and rectal swabs) were found to be negative, decolonization was performed for five days with 2 percent mupirocin nasal ointment, 2 percent chlorhexidine oropharyngeal rinses, and 4 percent chlorhexidine body scrub. The physician was furloughed from the care of patients until decolonization was completed. Cultures of the colo-

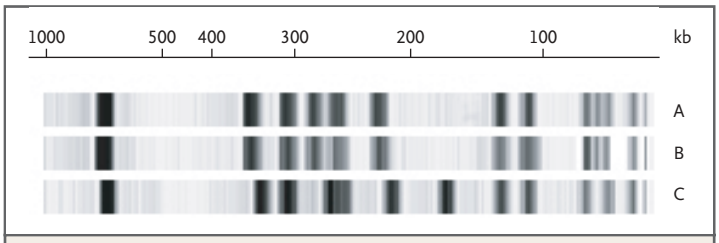


Figure 1. Pulsed-Field Gel Electrophoresis of Three Strains of Methicillin-Resistant *Staphylococcus aureus*.
Lane A shows the strain from the colonized physician. Lane B shows the USA300 strain and Lane C the reference strain, NCTC 8325, as a molecular-weight marker. Banding patterns were processed with GelCompar II software (Applied Maths).

nized sites were negative 2 days after decolonization and remained so for 3, 6, 12, and 18 months thereafter.

This example illustrates a plausible route of global transfer of community-acquired clones of MRSA. It is not clear whether the strain was acquired in a health care institution or in the community, but evidently it was acquired in the United States. Without screening and decolonization, there is the potential for spread into the hospital environment, with dangers of increased endemicity, morbidity, and even mortality.

Institutions with a low prevalence of MRSA should consider screening strategies for health care workers and patients who are native to high-prevalence countries or have recently traveled to such locales. The role of health care workers in the cause and transmission of community-acquired MRSA infection needs to be defined.

Andreas Tietz, M.D.
Reno Frei, M.D.
Andreas F. Widmer, M.D.

University Hospital Basel
CH-4031 Basel, Switzerland
atietz@uhbs.ch

1. Vandenesch F, Naimi T, Enright MC, et al. Community-acquired methicillin-resistant Staphylococcus aureus carrying Panton-Valentine leukocidin genes: worldwide emergence. Emerg Infect Dis 2003; 9:978-84.
2. Chambers HF. Community-associated MRSA — resistance and virulence converge. N Engl J Med 2005;352:1485-7.
3. Wannet W, Heck M, Pluister G, et al. Panton-Valentine leukocidin positive MRSA in 2003: the Dutch situation. Euro Surveill 2004; 9:28. (Also available at http://www.eurosurveillance.org/eq/2004/04-04/eq-04-2004.pdf.)

INSTRUCTIONS FOR LETTERS TO THE EDITOR

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

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

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

BOOK REVIEWS

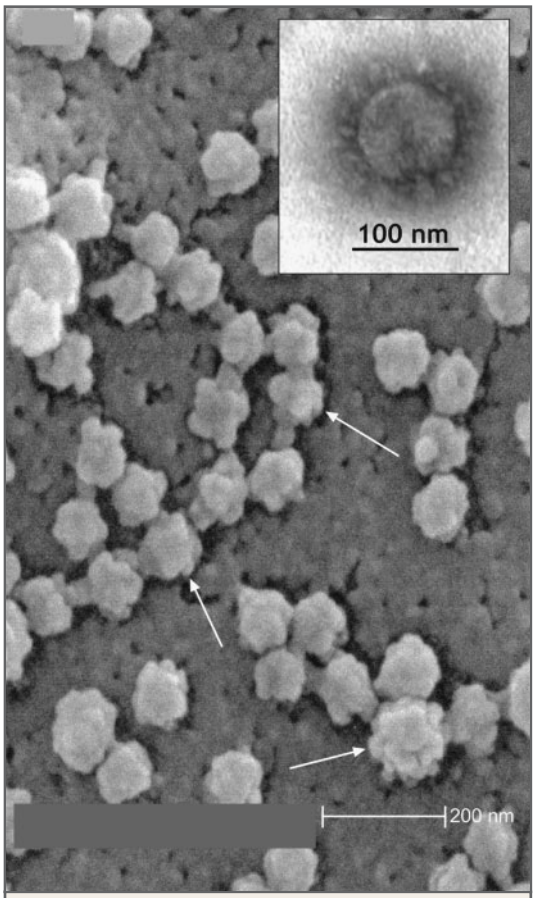
TWENTY-FIRST CENTURY PLAGUE: THE STORY OF SARS

By Thomas Abraham. 165 pp., illustrated. Baltimore, Johns Hopkins University Press, 2005. $24.95. ISBN 0-8018-8124-2.

In war, truth is the first casualty.
— Aeschylus (525–456 B.C.)

SEVERE ACUTE RESPIRATORY SYNDROME (SARS) is the plague of the new millennium. The disease has caused tremendous social, economic, and political disruptions worldwide. In his book, Thomas Abraham engages the reader as he discusses the trials and tribulations that were involved in reining in this infectious-disease epidemic in 2003. He does this skillfully by focusing on "ground zero," where the disease evolved, with two key chapters on China and Hong Kong. The book is richly filled with facts, and they are conveyed in a captivating manner. It is as impressive as the film *Outbreak*. More important, Abraham offers insights into the "dos and don'ts" of managing a public health crisis and provides key learning points that will be useful for dealing with emerging infections in the future. For readers who have had experience with the SARS epidemic firsthand, reading this book will cause déjà vu and bring back many memories, some of which will be heartrending.

As Abraham sees it, the World Health Organization (WHO) played a key leadership role in combating the spread of SARS and forged an exceptional global response to this first new global epidemic of the century. He cites the many quick and decisive actions taken by the WHO, including the unprecedented March 15, 2003, emergency travel advisory and the mobilizing of laboratories to work together to identify quickly the organism that caused SARS. That the previously unknown coronavirus was identified within a month of the outbreak was indeed an extraordinary achievement. Abraham is also candid in pointing out that the head of the WHO had to criticize one of its leading member states for tardiness in its response, a rare action on the part of this intergovernmental health organization.



Scanning Electron Micrograph Showing the Rosette-like Appearance of SARS Coronavirus Particles (Arrows).
In the inset, transmission electron microscopy and negative staining show the form and structure of the virus particle.

Courtesy of the Centers for Disease Control and Prevention/Dr. Mary Ng Mah Lee, National University of Singapore.

Nobody, however, had prior knowledge of the disease or the virus when SARS first struck. As Abraham points out, by the time the disease had spread to other countries, the "doctors in Guangdong knew more about SARS than anyone else." He adds that none of this wisdom was imparted to the rest of the world. This was disastrous. Socioeconomic and political forces were at play at the time and could have contributed largely to this silence. Perhaps, if there had been prompt reporting of the full facts so that others were forewarned

and could have taken preventive measures, history might have taken a different course. The SARS epidemic taught us that communication, collaboration, coordination, and transparency are crucial factors in containing a disease outbreak effectively and efficiently.

Abraham tells us that although tremendous progress has been made in medical science and technology, it was back-to-basics methods such as the old-fashioned techniques of virus isolation and quarantine that helped to contain SARS. The epidemic had also exposed many weaknesses in health care systems around the world in terms of their ability to contain the spread of infectious diseases. It is inevitable that SARS will not be the last epidemic the world has seen. As Abraham aptly puts it, "SARS was a dress rehearsal for the more serious threat posed by a new influenza pandemic." He sounds the warning with detailed accounts of the emergence and reemergence of outbreaks of avian influenza virus in Asia and elsewhere, and he concludes that SARS has offered us many valuable lessons on how to fight new global diseases. Indeed, we can find these inspirations by reading his book.

Suok Kai Chew, M.B., B.S.

Ministry of Health
169854 Singapore

AMERICAN PLAGUES: LESSONS FROM OUR BATTLES WITH DISEASE

By Stephen H. Gehlbach. 273 pp., illustrated.
New York, McGraw-Hill, 2005. $19.95. ISBN 0-07-143790-8.

In 1793, a massive epidemic of yellow fever terrorized Philadelphia. Frustrated by fruitless trials of sundry remedies, Dr. Benjamin Rush finally hit on a putative cure: plentiful bloodletting and purging with the use of the potent laxatives calomel and jalap. As Stephen Gehlbach dryly notes in his new book, the first patient "lost 3 pints of blood, the entire contents of his gastrointestinal tract, and any shreds of dignity. But he survived." Not everyone believed in Rush's gospel of phlebotomy and catharsis. The gloriously abusive journalist William Cobbett, writing under the pseudonym Peter Porcupine, pointed out that obstinate Philadelphians persisted in dying in record numbers, despite the widespread application of Rush's regimen. Rush failed to appreciate this early use of medical statistics to debunk his claims. The politically well-connected Rush, a signatory of the Declaration of Independence, successfully sued the rebarbative "Porcupine" for libel, driving him into bankruptcy and out of the city.

In *American Plagues*, Gehlbach selects 10 such historical tales to illustrate the birth pangs and development of public health in America. Each tale also highlights some principle of epidemiology or evidence-based medicine. Gehlbach describes the disputes in colonial Massachusetts regarding smallpox inoculation and vaccination; Daniel Drake's attempts to unravel the riddle of malaria in the Midwest; and the more recent heroic struggles against tuberculosis, poliomyelitis, and AIDS. *American Plagues* does not deal exclusively with infectious diseases, as the title might suggest. There are also chapters on the initially contentious link between cigarette smoking and lung cancer and the enormous importance of the Framingham Heart Study to our understanding of cardiovascular disease.

Gehlbach weaves some unifying themes through this diverse subject matter. We see how disease so frequently arises from social ills and how the dogmas of today painfully evolved out of the dogged debates of yesteryear. Despite great achievements, the history of public health has seldom been a linear, rational, triumphal progress; more often, it has been a fitful forging ahead amid great obstacles and uncertainties and on a murky, muddy battleground.

We also see how public health, like politics, sometimes makes strange bedfellows. In 18th-century Massachusetts, the euphoniously named physician Zabdiel Boylston, a proponent of smallpox inoculation, found an unlikely ally in the clergyman Cotton Mather. A Puritan intellectual who was a curious mixture of empiricism and credulity, Mather was an enthusiastic naturalist and a member of the Royal Society, although he is best known today as a vigorous defender of the Salem witch trials.

Gehlbach attempts to straddle the line between writing for lay and medical audiences, and for the most part he succeeds. His style is engaging and erudite, and his tone is at times sardonic. *American Plagues* is both a lively history of public health in the United States for medical and general readers and a splendid, painless introduction to the fields of biostatistics and epidemiology suitable for students.

John J. Ross, M.D., C.M.

Caritas St. Elizabeth's Medical Center
Boston, MA 02135
jrossmd@cchcs.org

POSITIVE PREVENTION: REDUCING HIV TRANSMISSION AMONG PEOPLE LIVING WITH HIV/AIDS

Edited by Seth C. Kalichman. 281 pp. New York, Kluwer Academic/Plenum, 2005. $59.95. ISBN 0-306-48699-7.

Traditionally, efforts to reduce the spread of human immunodeficiency virus (HIV) in the United States and elsewhere have focused on people who are at risk for becoming infected owing to their drug use or sexual behavior. The conceptual question has been, What are the circumstances under which people place themselves at risk for infection, and what can be done to reduce their risk? Through community outreach, messages about prevention, and access to health care services, the goal has been to empower people to protect themselves from contracting the virus.

A complementary approach to the prevention of the spread of HIV focuses directly on the potential sources of infection. HIV is spread by people who are infected, so it makes sense to consider how the epidemic can be understood and fought from that angle. Indeed, changing the behavior of one infected person may help prevent many of that person's sexual partners from becoming infected. From this perspective, the conceptual question is, What are the circumstances under which people who are HIV-positive place others at risk, and what are the best ways to address the issue? This is a sensitive question, because it could be perceived as blaming the victim or discriminating against or stigmatizing a group of people. Fortunately, a body of empirical studies has emerged in recent years that addresses this topic in an unbiased and straightforward manner.

Positive Prevention articulates this complementary approach and describes its empirical basis. As the title implies, the book focuses on the population of persons who are HIV-positive and are aware of their status. The nine chapters are authored by many of the top behavioral researchers who conduct studies of this population. The introductory chapter lays out the perspective, and subsequent chapters cover such topics as changes in behavior after diagnosis, self-disclosure of seropositive status to sex partners, and issues for HIV-positive gay and bisexual men, injection-drug users, and adolescents. Other chapters describe behavioral interventions in community and clinical settings. The final chapter, on international perspectives, provides epidemiologic data from the United Kingdom, Switzerland, Australia, India, and South Africa.

The message of the book is clear: the HIV epidemic can and should be addressed with a concerted effort at prevention directed toward those who are infected. It is encouraging that after becoming aware that they are infected, many people substantially reduce their risky sexual behavior and drug use. This outcome underscores the importance of HIV testing so that people can learn their status. The new rapid oral HIV tests, with which results can be obtained within 30 minutes, may increase testing rates. Equally important, those in whom HIV has been diagnosed must be successfully linked to medical care and supportive services. Living with HIV is challenging, and some people continue to engage in risky behavior or relapse periodically. The need for ongoing prevention programs is clear, and the book describes several of the behavioral interventions conducted recently in community and clinical settings.

This book is a "must read" for those interested in the prevention of HIV transmission. It has enough breadth to provide a general overview of "prevention with positives" and enough detail to satisfy and educate those already working in this area. Clinicians who want to develop or refine their approaches to counseling patients with HIV will find this book valuable.

Gary Marks, Ph.D.

Centers for Disease Control and Prevention
Atlanta, GA 30333