The American Journal of Medicine

Volume 120, Issue 3, Pages.191-285(March 2007)

Editorial

1. Pulling the Plug
   Page 191
   Joseph S. Alpert

Commentary

2. The Trialist, Meta-analyst, and Journal Editor: Lessons from ADAPT
   Pages 192-193
   Barbara K. Martin, John C.S. Breitner, Denis Evans, Constantine G. Lyketsos and Curtis L. Meinert

Reviews

3. Hepatocellular Cancer: A Guide for the Internist
   Pages 194-202
   Sameer Parikh and David Hyman

4. Meta-analysis Comparing CT Colonography, Air Contrast Barium Enema, and Colonoscopy
   Pages 203-210.e4
   Alan S. Rosman and Mark A. Korsten

Updates in office management

5. Office Management of Geriatric Urinary Incontinence
   Pages 211-220
   Cassandra F. Gibbs, Theodore M. Johnson II and Joseph G. Ouslander
6. Updates in the Management of Gout
   Pages 221-224
   Michael P. Keith and William R. Gilliland

Diagnosis dilemma: Infectious disease

7. A Cautionary Tale
   Pages 225-228
   Hesham M. Elgouhari and Susan M. Kellie

Diagnosis dilemma

8. An Absorbing Problem
   Pages 229-230
   Naveen Pemmaraju, Themistocles Dassopoulos and Alison R. Moliterno

Images in dermatology

9. Pressured to Appear
   Pages 231-232
   Rashid M. Rashid, Rajesh Yalamanchili and Amor Khachemoune

10. It’s Elemental
    Pages 233-235
    Daniel Petek, Karim Burkhardt and Laurence Legout

ECG image of the month

11. Looks Like Snow
    Pages 236-238
    Vamsee Yaganti, Chetan Naik and Diwakar Jain

Images in radiology

12. The Impersonator
    Pages 239-241
    Hartmut Lode, Jutta Wagner and Hans Scherübl

Clinical research studies

13. Increasing Trends in Incidence of Overweight and Obesity over 5 Decades
    Pages 242-250.e2
    Nisha I. Parikh, Michael J. Pencina, Thomas J. Wang, Katherine J. Lanier, Caroline S. Fox, Ralph B. D’Agostino and Ramachandran S. Vasan
*Pages 251-256*  
M. Alan Brookhart, Jerry Avorn, Jeffrey N. Katz, Joel S. Finkelstein, Marilyn Arnold, Jennifer M. Polinski, Amanda R. Patrick, Helen Mogun and Daniel H. Solmon

15. Development of a Risk Score for Colorectal Cancer in Men  
*Pages 257-263*  
Jane A. Driver, J. Michael Gaziano, Rebecca P. Gelber, I.-Min Lee, Julie E. Buring and Tobias Kurth

16. Smoking, Alcohol Consumption, and Raynaud’s Phenomenon in Middle Age  
*Pages 264-271*  
Lisa G. Suter, Joanne M. Murabito, David T. Felson and Liana Fraenkel

17. Cost-effectiveness of Treatment for Hepatitis C in an Urban Cohort Co-infected with HIV  
*Pages 272-279*  
Nicole G. Campos, Joshua A. Salomon, Julie C. Servoss, David P. Nunes, Jeffrey H. Samet, Kenneth A. Freedberg and Sue J. Goldie

AJM online

Clinical research studies

18. NSAID Use and Progression of Chronic Kidney Disease  
*Pages 280.e1-280.e7*  
Katherine Gooch, Bruce F. Culleton, Braden J. Manns, Jianguo Zhang, Helman Alfonso, Marcello Tonelli, Cy Frank, Scott Klarenbach and Brenda R. Hemmelgarn

19. Assessment of FIBROspec II to Detect Hepatic Fibrosis in Chronic Hepatitis C Patients  
*Pages 280.e9-280.e14*  
Atif Zaman, Hugo R. Rosen, Ken Ingram, Christopher L. Corless, Esther Oh and Katie Smith

Clinical communication to the editor

20. Risedronate-induced Hepatitis  
*Pages e1-e2*  
Michael B. Phillips
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>ADAPT Trial Data</td>
<td>e3</td>
<td>John Breitner, Denis Evans, Constantine Lyketsos, Barbara Martin and Curtis Meinert</td>
</tr>
<tr>
<td>22</td>
<td>The Reply</td>
<td>e5</td>
<td>Shelley R. Salpeter and Eric J. Topol</td>
</tr>
<tr>
<td>23</td>
<td>The Reply</td>
<td>e7</td>
<td>Joseph S. Alpert</td>
</tr>
<tr>
<td>24</td>
<td>Treatment of Staphylococcus aureus Prosthetic Valve Endocarditis</td>
<td>e9</td>
<td>Ali Hassoun</td>
</tr>
<tr>
<td>26</td>
<td>Orthostatic Intolerance in Chronic Fatigue Syndrome</td>
<td>e13</td>
<td>Peter C. Rowe and Katherine E. Lucas</td>
</tr>
<tr>
<td>27</td>
<td>Asian Indians and Coronary Artery Disease Risk</td>
<td>e15</td>
<td>Sastry Prayaga</td>
</tr>
<tr>
<td>28</td>
<td>Obstructive Sleep Apnea in Minorities</td>
<td>e17</td>
<td>Miguel A. Arias, Alberto Alonso-Fernández and Francisco García-Río</td>
</tr>
<tr>
<td>29</td>
<td>The Reply</td>
<td>e19</td>
<td>John C. LaRosa and Clinton D. Brown</td>
</tr>
<tr>
<td>30</td>
<td>Outcomes and Physician Specialty Among Patients with Heart Failure:</td>
<td>e21</td>
<td>Carlos Campillo and Jordi Forteza-Rey</td>
</tr>
</tbody>
</table>
31. The Reply
   *Page e23*
   JoAnne Micale Foody

32. Reduce the Risk of Reinfarction and Related Complications after Myocardial Infarction
   *Pages e25-e26*
   Charles Shang

33. The Reply
   *Page e27*
   Stein Ørn and Kenneth Dickstein

34. Progression of Chronic Kidney Disease: Can It Be Prevented or Arrested?
   *Page e29*
   Ashok K. Malani, Hussam Ammar and Mohamed H. Yassin

35. The Reply
   *Page e31*
   Bertrand L. Jaber and Nicolaos E. Madias

APM perspectives

36. Physician Preferences and Attitudes Regarding Industry Support of CME Programs
   *Pages 281-285*
   Paul S. Mueller, C. Christopher Hook and Scott C. Litin
EDITORIAL

Pulling the Plug

PULLING THE PLUG

In the vernacular of the house officer, pulling the plug means discontinuing life support in a badly damaged patient whose survival is highly unlikely. This act reflects acceptance by the family and the healthcare team that aggressive measures aimed at keeping the patient alive are futile and not in the patient’s best interest. While discontinuing life support will always be a painful and difficult act—both for the family of the dying patient and for the medical staff—it was even more complicated when I was in training 35 years ago. At that time, physicians had to repeatedly explain to family members that brain death meant the death of the individual, regardless of what other organs were still functioning.

As a result of extensive media coverage, the general public now understands that brain function is the central tenet of human existence and that without substantial cerebral activity, functioning internal organs do not imply that the patient is alive. People have been so well-schooled on the concept of brain death, that the overwhelming majority of Americans had a strongly negative response when Florida Governor Jeb Bush, the Florida Legislature, and the US Congress stepped into Terri Schiavo’s case, supporting her parents’ efforts to maintain her in a persistent vegetative state—against the wishes of Schiavo’s husband.1

Today, patients commonly express an aversion to a life without consciousness, a life maintained by artificial ventilation and nutrition. Some will detail their sentiments in an advance health directive. Family members, stating that their loved one “never wanted to live on a machine,” often request discontinuation of a comatose patient’s life support early in the hospitalization and at times, before the extent of the patient’s brain injury has been defined. Consequently, the conversations I have with families forced into this clinical situation currently focus on how quickly and how securely the prognosis for brain recovery can be determined so that an informed and compassionate decision can be made.

Although The American Journal of Medicine does not normally review books, I want to call attention to a recently published book by attorney W.H. Colby.1 Mr. Colby recounts his experience as the principal attorney for the Cruzan family, one of the most influential right-to-die cases in US legal history. He also discusses similar cases, including that of Paul Brophy, a firefighter and emergency medical technician who suffered a cerebral hemorrhage and then lapsed into a persistent vegetative state.2,3,4 After several years, his family asked that artificial nutrition and hydration be discontinued, a request opposed by the hospital caring for Mr. Brophy. The case went to the Massachusetts Supreme Court, and in a split decision, the court found for the Brophy family, and nutrition and hydration were discontinued.

I testified on behalf of the Brophys and have since remained particularly sensitive to these matters. For that reason, I appreciated Colby’s clear, detailed reporting of the arguments and outcomes in the most important right-to-die cases in the US during the last 30 years. Every clinician who might care for patients on life support or shepherd family members through the experience should consider reading it. The nonmedical population would do well to ponder these issues as well. At one point or another, we will all face the quandaries described. It is my hope that physicians find Colby’s book to be provocative and a source of support. I welcome responses to this editorial at jalpert@email.arizona.edu.

Joseph S. Alpert, MD
Robert S. and Irene P. Flinn Professor of Medicine
Special Assistant to the Dean
University of Arizona College of Medicine
Tucson
jalpert@email.arizona.edu

References
COMMENTARY

The Trialist, Meta-analyst, and Journal Editor: Lessons from ADAPT

The Alzheimer’s Disease Anti-inflammatory Prevention Trial (ADAPT) was designed to investigate whether the nonselective nonsteroidal anti-inflammatory drug naproxen sodium (Aleve, Bayer, Consumer Care, Morristown, NJ) or the selective cyclooxygenase-2 inhibitor celecoxib (Celebrex, Pfizer, New York, NY) can prevent Alzheimer’s dementia or delay cognitive decline. Recruitment for the trial began in mid-2001. Some 3 ½ years later, more than 2500 participants had been enrolled. On December 17, 2004, the National Cancer Institute announced the finding of increased cardiovascular risks with the use of celecoxib in its Adenoma Prevention with Celecoxib (APC) trial. That same day, both the Pfizer-sponsored Prevention of Sporadic Adenomatous Polyps trial and ADAPT halted study treatments, even though these trials did not show the same degree of risk as the APC trial. The rationale for suspending treatments in ADAPT was discussed in a statement read at the February 18, 2005, joint meeting of the Food and Drug Administration Arthritis Advisory Committee and the Drug Safety and Risk Management Advisory Committee.1

This commentary recounts the ADAPT experience after the termination of study treatments, first with respect to our repeated unsuccessful attempts to publish the ADAPT cardiovascular and cerebrovascular safety data, and second relating to the unauthorized use of our unpublished data in a meta-analysis. We write here to describe our perception of growing tension among the researchers who conduct clinical trials, meta-analysts who use and summarize trial data, and journal editors who publish both individual trial results and meta-analyses. We suggest some lessons to be learned.

It is ADAPT policy to seek publication in a peer-reviewed journal as soon as possible when results lead to a major protocol change and to forego public presentation of those results before publication. Therefore, with strong encouragement from our sponsor, the National Institute on Aging, we began before the close of 2004 to write our rationale for suspending treatment and enrollment in ADAPT and the trial’s cardiovascular and cerebrovascular results. Also in keeping with policy, we declined the request of the Food and Drug Administration to present the trial’s safety data in February 2005. We expected, mistakenly, that our safety results would be published within a few months of the meeting.

Our initial submission for publication described the ADAPT Steering Committee’s rationale for suspending treatment and also provided limited preliminary results. In response to urgings of reviewers, later revisions of the article presented complete (extending through the entire treatment period) and more extensive cardiovascular and cerebrovascular event data. To our surprise, however, the article was rejected by 5 journals in succession. The overriding reason for rejection was that the safety results were not definitive, owing to the moderate effect sizes and wide confidence intervals. Although the study treatments were terminated early and the number of events was small, the absolute number of events (10 cardiovascular deaths, 39 nonfatal myocardial infarctions, and 24 strokes) was actually larger than the numbers reported in the New England Journal of Medicine by the APC trial (10 cardiovascular deaths, 21 nonfatal myocardial infarctions, and 11 strokes)—only distributed such that the point estimates of the treatment effects were weaker.

A further difficulty was that ADAPT was not designed as a safety trial, and thus, the method of capturing safety-related events was less than perfect. Of course, with hindsight, one might have done a better job of capturing detailed data on cardiovascular events, but it is difficult to design for the unexpected. To deal with this problem, we could have resorted to a process of post hoc review and adjudication of cardiovascular events, but such a process has its own limitations. Because one cannot identify missed events, such a method can only decrease the absolute risks observed, and one cannot be certain that its results would not be biased, for example, by controversies such as that now surrounding the cardiovascular risks of nonsteroidal anti-inflammatory drugs. We thought it preferable to rely on the fact that the trial was randomized and double-masked, and thus, that any differences observed were not likely attributable to bias in ascertainment of events. This perspective seemed to carry little weight with reviewers or editors.

Thus, the first point of tension illustrated by the ADAPT experience relates to the desire of reviewers and journal editors for results that are definitive (i.e., likely to be reproduced in a repeat experiment). However, trials...
are not typically conducted to prove harm. Instead, trialists have a duty to cease treatment if risks seem to outweigh the prospect for benefit, and this is particularly the case with safety outcomes in prevention trials. Furthermore, we question the wisdom of dismissing trial results as “nondefinitive,” not because of the quantity of data available but because of the magnitude of the $P$ values attached. Instead, we would argue that if, as has often been stated, investigators have an ethical obligation to publish efficacy and safety results regardless of their nature or direction, then editors also must recognize their duty to publish such results even when they are not “definitive.” We commend the new online journal *PLoS Clinical Trials* for making it their mission “to maximize the number of trials whose results are available in the public domain,” regardless of “the direction of results, size, or perceived importance of the trial.”

Our difficulty and delay in publishing the ADAPT cardiovascular and cerebrovascular results set the stage for another event—the unauthorized use of our unpublished data (likely from a confidential data monitoring report generated just before suspension of study treatments) in a meta-analysis by Salpeter et al. published in the July 2006 issue of *The American Journal of Medicine*. Our letter to the editor about the use of ADAPT data and Dr. Salpeter’s response have been included in AJM Online section of this issue.

This latter event exemplifies the potential for tension between trialists and meta-analysts. Good methodology for meta-analyses posits that data should be used from all relevant sources, whether published or not. Indeed the drive to register trials, endorsed by both meta-analysts and journal editors, stems in large part from a desire to avoid publication bias in the systematic review of trial data. Therefore, when relevant data are not published or otherwise in the public domain, meta-analysts regularly seek unpublished data from study investigators or sponsors. An unfortunate consequence is that the completeness of meta-analyses depends on the decision of investigators or sponsors to comply with the request for data. They should certainly comply if they are not likely to publish the data in question, but will likely decline if they are planning their own publication. They also may be obliged to withhold data while their article is pending review, as was actually the case for ADAPT when Salpeter et al. conducted and published their meta-analysis. Trial investigators should not be seen as hindering the progress of evidenced-based medicine if they decline to provide data that they themselves are actively attempting to publish.

More emphatically, the desire to increase the scope of a given meta-analysis provides no justification for the use of unpublished data without proper approval. Meta-analysts and producers of systematic reviews are obliged to ensure that any unpublished data in their articles are properly obtained and vetted, and to report in their publication the procedures used to acquire such data. Likewise, journal editors receiving meta-analysis articles have a duty to require that authors attest to the possession of permission for use of unpublished data. Many journals currently require authors to obtain permission to cite personal communications or to list an individual in an acknowledgment. The use of unpublished data should surely be held to a similar standard. We are pleased that, as the result of the ADAPT experience, the Editor-in-Chief of this Journal has revised policy to require proper assurances and disclosures for future publications. We hope other journals will follow suit.

**Barbara K. Martin**

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Md

John C. S. Breitner

Department of Psychiatry and Behavioral Sciences, University of Washington, and VA Puget Sound Health Care System, Seattle, Wash

Denis Evans

Department of Internal Medicine, Rush University Medical Center, Chicago, Ill

Constantine G. Lyketsos

Department of Psychiatry, Johns Hopkins School of Medicine, Baltimore, Md

Curtis L. Meinert

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Md

**References**


Hepatocellular cancer is the fifth most common cause of cancer and the third leading cause of cancer-related deaths worldwide.1 Its incidence has increased dramatically in the United States because of the spread of hepatitis C virus infection and is expected to increase for the next 2 decades. Hepatitis B virus, hepatitis C virus, and chronic heavy alcohol use leading to cirrhosis of the liver remain the most important causes. The diagnosis of hepatocellular cancer rests on a combination of radiologic, serologic, and histopathologic criteria. Liver transplantation is the only definitive treatment. Resection of the tumor and other percutaneous therapies are more commonly used in practice, because most hepatocellular cancers are detected at an advanced stage. Patients who are at high risk for the development of hepatocellular cancer should be screened with an ultrasound of the liver every 6 months. The prognosis is dependent on both the underlying liver function and the stage at which the tumor is diagnosed. The aim of this review is to familiarize internists in screening, diagnosis, and referral of patients with hepatocellular cancer in an appropriate and timely fashion. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Hepatocellular cancer; Hepatitis C; Chronic alcoholism; Cirrhosis; Screening

Hepatocellular cancer is the fifth most common cause of cancer and the third leading cause of cancer-related deaths worldwide. A large, retrospective cohort study confirmed an almost 2-fold increase in the incidence of hepatocellular cancer from 1975 to 1998 in the United States.2,3 This increase is primarily related to the spread of hepatitis C virus (HCV) infection, which peaked in the United States in the late 1980s.4 Given the time lag of 2 to 4 decades between the onset of infection and the development of cirrhosis, it has been predicted that the incidence of hepatocellular cancer will continue to increase over the next 2 decades.4 According to the American Cancer Society, there will be 19,160 new cases diagnosed and 16,780 deaths due to this disease in the United States in 2007.5

Males are more commonly affected than females in the ratio of 3:1 to 9:1.6 The mean age of presentation of hepatocellular cancer in Europe and the United States is approximately 60 years.7 This is in contrast with patients in Asia and Africa, where it is between 20 and 50 years.

CAUSE

The major clinical risk factor for the development of hepatocellular cancer is cirrhosis of the liver. Chronic infections with hepatitis B virus (HBV) and HCV and chronic heavy alcohol use are the most important risk factors for the
development of cirrhosis. HBV accounts for the majority of hepatocellular cancer in China and Africa, where most of the infection is acquired early in life either from mother to the offspring or by horizontal transmission. In contrast, HCV accounts for most of the cases in the Western hemisphere. Chronic alcohol use of greater than 80 g per day for more than 10 years increases the risk of hepatocellular cancer 5-fold. Furthermore, chronic alcohol use in HBV or HCV infection doubles the risk of hepatocellular cancer over either infection alone.

The magnitude of risk of hepatocellular cancer from cirrhosis due to other causes is not well known. There have been reports in patients with hereditary hemochromatosis, alpha-1 antitrypsin deficiency, and autoimmune hepatitis. An increased incidence has been associated with smoking and exposure to aflatoxin, a mycotoxin that contaminates peanuts and soybeans, and causes mutations in the p53 tumor suppressor gene. Approximately one quarter of all cases diagnosed in the United States do not have any of these risk factors. There is growing interest in the role of insulin resistance syndrome as a risk factor for these cryptogenic cases. Insulin resistance syndrome is present in virtually all cases of nonalcoholic steatohepatitis and cirrhosis in 10% to 20% of cases. Diabetes and obesity, 2 major manifestations of insulin resistance syndrome, have been shown to double the risk of hepatocellular cancer (Table 1).

### CLINICAL FEATURES

The typical clinical manifestations of hepatocellular cancer are right upper quadrant abdominal pain, early satiety, and weight loss. However, more and more hepatocellular cancers are now detected at an asymptomatic stage because of the growing awareness of these tumors in patients with chronic liver disease and cirrhosis. Other clinical presentations, such as spontaneous rupture of the tumor into the peritoneal cavity, obstructive jaundice, and bony pain from metastasis, are extremely uncommon. Various paraneoplastic syndromes have been associated with hepatocellular cancer. These include erythrocytosis (erythropoietin), hypoglycemia (insulin-like growth factor), and hypercalcemia (parathyroid-related protein). Physical findings in patients with hepatocellular cancer generally reflect the severity of the underlying chronic liver disease and cirrhosis. The liver may be enlarged and a vascular bruit is sometimes heard, consistent with hypervascularity of the tumor.

### CLINICAL SIGNIFICANCE

- In the US, the incidence of hepatocellular cancer doubled between 1975 and 1998 and is expected to continue to increase for the next 2 decades.
- The American Association for the Study of Liver Diseases recommends an ultrasound of the liver every 6 months in high-risk patients to screen for hepatocellular cancer.
- Liver transplantation remains the only definitive treatment of hepatocellular cancer, although surgical resection and percutaneous therapies are more commonly applied in routine clinical practice.

### DIAGNOSIS

A consensus statement from the European Association for the Study of Liver Diseases (EASL) has been formulated to help clinicians standardize diagnostic approaches (Table 2).

### Lesions Greater Than 2 Centimeters in Diameter

In nodules greater than 2 cm diameter in size, a diagnosis of hepatocellular cancer can be made if any 2 imaging studies (including ultrasonography, computed tomography, magnetic resonance imaging, or hepatic arteriography) show increased vascularity. Alternatively, only 1 imaging study with an alpha-fetoprotein level greater than 400 ng/mL is diagnostic. Our ability to diagnose these tumors noninvasively rests on the premise that they are seen on a background of cirrhosis and enhance with contrast on rapid-sequence imaging secondary to their neovascularity. These radiologic criteria for diagnosis have excellent diagnostic accuracy with reported sensitivity of 100% and specificity of 98.8%. In cases of indeterminate radiologic findings, fine-needle aspiration biopsy is recommended.

### Lesions Less Than 2 Centimeters in Diameter

Imaging techniques for lesions less than 2 cm do not have sufficient accuracy in distinguishing hepatocellular cancer from other conditions. Alpha-fetoprotein levels may be normal or only slightly elevated and thus provide no diagnostic utility. Hepatic lesions less than 1 cm in size have a less than 50% chance of being malignant, and serial ultrasound
(every 3 months) is recommended. On the other hand, fine-needle aspiration biopsy should be performed in lesions between 1 and 2 cm in size.

Role of Liver Biopsy

The role of liver biopsy has been the subject of great controversy. For almost all other types of cancer, histopathologic confirmation is necessary to make a diagnosis. However, as elucidated in Table 2, both the EASL and United Network for Organ Sharing criteria do not require a biopsy for the diagnosis of hepatocellular cancer in lesions greater than 2 cm. For lesions less than 2 cm in size where high-quality imaging or expertise in reading these images is not available, a biopsy is recommended.

There is a small but definite risk of tumor seeding from an invasive biopsy. The prevalence rates have been reported to be anywhere between 0.003% and 5%. However, it is unclear whether it leads to metastatic disease or worse survival because a majority of patients are treated by excision of the subcutaneous tumor deposit. Also, the false-negative rate from biopsy of lesions less than 2 cm is approximately 30% to 40%. Thus, a negative biopsy does not conclusively rule out the diagnosis of hepatocellular cancer.

Role of Serum Markers

The 3 most commonly used serum markers are alpha-fetoprotein, Lens culinaris agglutinin-reactive alpha-fetoprotein (alpha-fetoprotein-L3), and protein induced by vitamin K antagonist-II. The sensitivity and specificity of these markers to diagnose hepatocellular cancer vary according to the threshold level used. Total alpha-fetoprotein has a sensitivity of 60% and specificity of 90% at cutoff values between 10 and 20 ng/mL. A systematic review confirmed the poor diagnostic ability of alpha-fetoprotein alone in detecting hepatocellular cancer at any level of pretest risk. It is a much better diagnostic test in the presence of a hepatic mass where a cutoff value of greater than 400 ng/mL is used in combination with imaging criteria. An increase in the percentage of alpha-fetoprotein-L3 over the total alpha-fetoprotein (>10%) is specific for small hepatocellular cancer. Protein induced by vitamin K antagonist-II is also more specific than total alpha-fetoprotein in detecting hepatocellular cancer. However, these are not available in most nonresearch laboratories in the United States at this time.

SCREENING

Although there is no definite evidence that screening in hepatocellular cancer improves survival, many physicians screen patients in high-risk groups with either serum alpha-fetoprotein or ultrasound of the liver or both. Two recent randomized controlled trials completed in China demonstrated a significant reduction in hepatocellular cancer-related mortality in patients who underwent screening. Ultrasound of the liver is the preferred screening test because it has a sensitivity of 84% and specificity of more than 90%. A combination of alpha-fetoprotein and ultrasound has been reported to increase the sensitivity by 5% to 10% over ultrasound alone, but it also increases costs and false-positive rates.

The United States Preventive Services Task Force, National Comprehensive Cancer Network, and American Cancer Society do not have any specific guidelines for screening patients for hepatocellular cancer. The National Cancer Institute recommends against routine screening for lack of a survival benefit. The American Association for the Study of Liver Diseases and EASL recommend ultrasound of the
liver every 6 months for high-risk patients.\textsuperscript{42} (Table 3 and Figure 1).

**NATURAL HISTORY AND PROGNOSIS**

Prospective studies have shown that most hepatocellular cancers develop through a progressive pathway from pre-malignant nodular lesions to cancerous lesions in the cirrhotic liver.\textsuperscript{43} Progression takes an average of approximately 2 to 4 decades from the initial time of infection with HBV or HCV to the development of cirrhosis. Thereafter, the annual risk of hepatocellular cancer is 2\% to 3\% for HBV, 1\% to 7\% for HCV, and 1\% for alcohol-induced cirrhosis.\textsuperscript{10,44} Hepatocellular cancer can develop in the absence of cirrhosis in patients with HBV infection at a rate of 0.26\% to 0.6\% per year.\textsuperscript{44} Recent studies have shown that treatment of chronic HCV infection with interferon monotherapy in patients with cirrhosis decreases the risk of hepatocellular cancer, and it is expected that combination therapy with pegylated interferon and ribavirin may reduce this risk even further.\textsuperscript{45,46}

Predicting survival in hepatocellular cancer is complicated by the fact that 2 disease processes, namely, the tumor and underlying cirrhosis, are present simultaneously. Numerous studies have shown that prognosis is directly proportional to the degree of hepatic function, suggesting that cirrhosis rather than mass size of the tumor is the main determinant of outcome. The median survival of untreated patients with newly diagnosed hepatocellular cancer is weeks to months.\textsuperscript{47} A number of factors are associated with worse outcome: male sex, advanced age, etiologic agent

**Table 3** High-Risk Groups for Screening of Hepatocellular Cancer

<table>
<thead>
<tr>
<th>Cirrhosis:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatitis B</td>
</tr>
<tr>
<td>Hepatitis C</td>
</tr>
<tr>
<td>Alcoholic cirrhosis</td>
</tr>
<tr>
<td>Hereditary hemochromatosis</td>
</tr>
<tr>
<td>Primary biliary cirrhosis</td>
</tr>
<tr>
<td>Nonalcoholic steatohepatitis</td>
</tr>
<tr>
<td>Patients waiting on the liver transplant list</td>
</tr>
</tbody>
</table>

No cirrhosis: Chronic hepatitis B carriers: males aged >40 y and females aged >50 y, family history of hepatocellular cancer in a patient with chronic hepatitis B.

Screening for patients with cirrhosis secondary to alpha-1 antitrypsin deficiency, autoimmune hepatitis, and Wilson’s disease is considered low-moderate risk, and there are no recommendations for screening at this time.

---

Figure 1 Surveillance of hepatocellular cancer in patients with cirrhosis of the liver. HCC = hepatocellular cancer; AFP = alpha-fetoprotein.
evolve. In the 1980s, poor patient selection led to dismal
tation as the primary method of treatment continue to
Liver Transplantation.
but need further validation.
before using resection, it is
Resection of the tumor is the treatment of choice
Surgical Therapy
ing cancer-prone cirrhotic liver (Figure 2).
The definitive treatment of hepatocellular cancer is liver
MANAGEMENT
The definitive treatment of hepatocellular cancer is liver
transplantation; this cures both the cancer and the underly-
ing cancer-prone cirrhotic liver (Figure 2).
Surgical Therapy
Resection. Resection of the tumor is the treatment of choice
for hepatocellular cancer.49,50 Before using resection, it is
necessary to demonstrate sufficient liver reserve by calcu-
lating the Child-Pugh Score (Table 5). Generally, patients
with Child-Pugh class A can safely undergo resection. How-
ever, not all patients with class A have homogenous liver
function, and therefore, presence of portal hypertension is
assessed to determine feasibility of resection. With these
criteria, a 5-year survival of approximately 60% to 70%
can be achieved. However, tumor recurrence complicates ap-
proximately 70% of patients at 5 years.51 Adjuvant chemo-
therapy and chemoembolization have not been shown to be
of any added benefit.52,53 Various other adjuvant treatment
approaches, including internal radiation with I-131–labeled
lipiodol,54 adoptive immunotherapy with activated lympho-
cytes,55 and interferon,56,57 have shown promising results
but need further validation.
Liver Transplantation. The indications for liver transplan-
tation as the primary method of treatment continue to
evolve. In the 1980s, poor patient selection led to dismal
outcomes with 5-year survival of less than 40%.58 A land-
mark clinical trial in 1996 established the so-called “Milan
Criteria” for selecting ideal candidates for liver transplan-
tation. The investigators included patients with a single
lesion less than 5 cm in size or 3 lesions each less than 3 cm
in size. Their 5-year survival exceeded 70%, and the recur-
rence rate was less than 15%.59 Indeed, these results were
duplicated in other studies in which the 5-year survival
exceeded 75%, far greater than survival after resection or
ablation.60 These criteria are still used today for listing
patients on the transplant registry and are listed on the
United Network for Organ Shearing website28 (Table 6).
The assignment of priority scores for liver transplanta-
tion is based on the Model for End-Stage Liver Disease
score, which uses laboratory values of serum creatinine,
total bilirubin, and international normalized ratio. Patients
with hepatocellular cancer are assigned a higher Model for
End-Stage Liver Disease score and thus, get a priority for
transplant over those with similar degrees of liver dysfunc-
tion, and without hepatocellular cancer, who are waiting for
a transplant. However, a shortage of donors has led to
unacceptably high dropout rates because of deaths or ap-
ppearance of contraindications, with survival decreasing to
less than 50% using an intention-to-treat principle.61
Given the difficulties in obtaining a cadaveric liver in a
timely fashion, “bridging” therapies such as surgical resec-
tion, neoadjuvant local ablation, and chemoembolization
have been tried with promising results.62,63 Living-donor
liver transplantation is being increasingly performed in the
United States with results comparable to those undergoing
cadaveric transplantation.64
Nonsurgical Therapy
The majority of hepatocellular cancers identified at initial
presentation are unresectable and do not qualify to be on the
transplant list. A number of other options are available.
Local Ablation. Local ablation uses image-guided chemi-
cal (ethanol, acetic acid) and thermal (radiofrequency,
cryoablation) techniques. Ablation is commonly used with
curative intent in patients with unresectable tumors, with
survival similar to resection. Percutaneous ethanol injection

dimensional: 585.0x783.0

Table 4 Barcelona Clinic Liver Cancer Staging Classification47

<table>
<thead>
<tr>
<th>Stage</th>
<th>Performance Status Test*</th>
<th>Tumor Stage</th>
<th>Okuda Stage †</th>
<th>Liver Function Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>Single</td>
<td>I-II</td>
<td>Child Pugh ‡ A-B</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>Large multinodular</td>
<td>I-II</td>
<td>Child Pugh ‡ A-B</td>
</tr>
<tr>
<td>C</td>
<td>1-2</td>
<td>Vascular invasion, extrahepatic spread</td>
<td>I-II</td>
<td>Child Pugh ‡ A-B</td>
</tr>
<tr>
<td>D</td>
<td>3-4</td>
<td>Any</td>
<td>III</td>
<td>Child Pugh ‡ C</td>
</tr>
</tbody>
</table>

*Performance status test is based on the Eastern Co-Operative Oncology Group performance scale: 0: asymptomatic, 1: symptomatic and fully
ambulatory, 2: symptomatic and in bed <50% of the day, 3: symptomatic and in bed >50% of the day, 4: bedridden.
†Okuda staging system (I-III) is another staging system that takes into account the size of the tumor, presence of ascites, and albumin and bilirubin concentrations.
‡For Child-Pugh Classification, see Table 5.

(HCV worse than HBV), presence of more than 1 risk
factor, size, number and doubling time of nodules, vascular
invasion, and distant metastasis.
Because of the heterogeneous nature of hepatocellular
cancer with respect to its cause, epidemiologic background,
and severity of hepatic dysfunction, a worldwide staging
system is not in place. The most commonly used staging
system for solid tumors, TNM classification, has severe
limitations because it does not include the severity of un-
derlying cirrhosis. Therefore, other staging systems such as
the Barcelona Clinic Liver Cancer staging classification48
(Table 4) and Cancer of Liver Italian Score have been
developed.

MANAGEMENT
The definitive treatment of hepatocellular cancer is liver
transplantation; this cures both the cancer and the underly-
ing cancer-prone cirrhotic liver (Figure 2).

MANAGEMENT
The definitive treatment of hepatocellular cancer is liver
transplantation; this cures both the cancer and the underly-
ing cancer-prone cirrhotic liver (Figure 2).

MANAGEMENT
The definitive treatment of hepatocellular cancer is liver
transplantation; this cures both the cancer and the underly-
ing cancer-prone cirrhotic liver (Figure 2).

MANAGEMENT
The definitive treatment of hepatocellular cancer is liver
transplantation; this cures both the cancer and the underly-
ing cancer-prone cirrhotic liver (Figure 2).
was the most commonly applied technique, with a response rate of 70% to 100%. However, radiofrequency thermal ablation is more commonly used now because it can achieve better control of disease and improve survival compared with percutaneous ethanol injection. The major limitation of local ablation is its inability to achieve meaningful response rates in infiltrative lesions and in tumors larger than 4 to 5 cm in size.

**Transarterial Therapy**, Transarterial interventions are available for treatment of large unresectable hepatocellular cancers that are not amenable to resection or percutaneous therapies. These are generally used with palliative intent to reduce tumor burden. The most commonly used techniques include transcatheter arterial chemoembolization and transarterial radioactive iodine with lipiodol. A systematic review of randomized trials for unresectable hepatocellular carcinoma showed that in patients with compensated cirrhosis and good functional status, arterial embolization improved 2-year survivals. Postembolization syndrome, associated with abdominal pain and fever, can sometimes occur and precipitate ascites and hepatic encephalopathy.
Patients with advanced liver disease (Child-Pugh C) and portal vein thrombosis should not undergo these therapies because of the risk of precipitating acute liver failure.

**Combination Therapy.** Combined therapy with transcatheter arterial chemoembolization followed by radiofrequency thermal ablation has been shown to produce good local response, especially in tumors less than 5 cm. However, the overall usefulness of this procedure needs to be established in a larger number of patients.

**Systemic Treatment.** Numerous systemic therapies, including doxorubicin, tamoxifen, megestrol, interferon alfa, and anti-androgens, have been tried and compared in randomized trials. The use of most of these agents is associated with significant toxicity without any discernible benefit with regard to survival or complete response. However, systemic treatment has shown promising results and may replace the Milan criteria in a larger number of patients.

**FUTURE TRENDS**

Proteomics has led to the discovery of new molecular markers, such as des-gamma carboxyprothrombin and human hepatocyte growth factor, for screening hepatocellular cancer, and these are being validated for clinical use. Antiangiogenesis agents such as vascular endothelial growth factor antibodies and thalidomide, nonspecific inhibitors of carcinogenesis such as Sandostatin and arsenic, and better means of delivering radiation such as yttrium microspheres are all being actively investigated for the treatment of hepatocellular cancer. Expanding the criteria for selecting patients for liver transplantation, such as the University of California San Francisco criteria, which include a single tumor less than 6.5 cm, 3 or less nodules with the largest being less than 4.5 cm, and total tumor diameter less than 8 cm, has shown promising results and may replace the Milan criteria.

**CONCLUSION**

The incidence of hepatocellular cancer is increasing in the Western world, including the United States. Although HBV, HCV, and alcohol use constitute the most important risk factors for the development of hepatocellular cancer, diabetes and obesity may contribute to increased carcinogenicity. Primary care physicians taking care of patients with chronic viral hepatitis and cirrhosis will need to have a heightened awareness of hepatocellular cancer, because the translation of bench research into clinical practice will lead to newer diagnostic tests, better therapeutic options, and improved survival of these patients. Finally, an important frontier in the battle against hepatocellular cancer will be the application of effective preventive strategies aimed at decreasing the risk of transmission of HBV and HCV, and the development of safe and effective medications for the treatment of chronic HBV and HCV.

**ACKNOWLEDGMENTS**

We are indebted to Drs Richard Goodgame, Hashem El-Serag, and Prashant Kapoor for their critical review of this article.

**References**


---

**Table 5 Child-Pugh Score**

<table>
<thead>
<tr>
<th>Abnormality</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encephalopathy (grade)</td>
<td>None</td>
<td>1-2</td>
<td>3-4</td>
</tr>
<tr>
<td>Ascites</td>
<td>None</td>
<td>Slight</td>
<td>Moderate</td>
</tr>
<tr>
<td>Albumin (g/dL)</td>
<td>&gt;3.5</td>
<td>2.8-3.5</td>
<td>&lt;3.5</td>
</tr>
<tr>
<td>Prothrombin time prolonged (sec)</td>
<td>1-4</td>
<td>4-6</td>
<td>&gt;6</td>
</tr>
<tr>
<td>Bilirubin (mg/dL)</td>
<td>1-2</td>
<td>2-3</td>
<td>&gt;3</td>
</tr>
</tbody>
</table>

Class A: 5-6 points (good operative risk); Class B: 7-9 points (moderate operative risk); Class C: 10-15 points (poor operative risk).


Meta-analysis Comparing CT Colonography, Air Contrast Barium Enema, and Colonoscopy

Alan S. Rosman, MD, Mark A. Korsten, MD

Section of Gastroenterology and Medicine Program, James J. Peters VA Medical Center and Mount Sinai School of Medicine, New York, NY.

ABSTRACT

INTRODUCTION: Published studies have reported a wide range of sensitivities and specificities for computed tomographic (CT) colonography for polyp detection, generating controversy regarding its diagnostic accuracy.

METHODS: A meta-analysis of published studies comparing the accuracies of CT colonography and colonoscopy for polyp detection was performed. The pooled per-patient sensitivities and specificities were calculated at various thresholds for polyp size. Summary receiver operating characteristic (sROC) curves were also constructed.

RESULTS: Thirty studies were included in the meta-analysis of CT colonography. The pooled per-patient sensitivity of CT colonography was higher for polyps greater than 10 mm (0.82, 95% confidence interval [CI], 0.76-0.88) compared with polyps 6 to 10 mm (0.63, 95% CI, 0.52-0.75) and polyps 0 to 5 mm (0.56, 95% CI, 0.42-0.70). Similarly, the exact area under the sROC curve (area ± standard error) was higher using a threshold greater than 10 mm (0.898 ± 0.063) compared with thresholds of greater than 5 mm and any size (0.884 ± 0.033 and 0.822 ± 0.059, respectively). There were no significant differences in the diagnostic characteristics of 2-dimensional versus 3-dimensional CT colonography. At a threshold greater than 5 mm, the exact area under the sROC curve was significantly higher for endoscopic colonoscopy compared with CT colonography (0.998 ± 0.006 vs 0.884 ± 0.033, P < .005).

CONCLUSIONS: CT colonography has a reasonable sensitivity and specificity for detecting large polyps but was less accurate than endoscopic colonoscopy for smaller polyps. Thus, CT colonography may not be a reasonable alternative in situations in which a small polyp may be clinically relevant. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: CT colonography; Barium enema; Colonoscopy; Summary receiver operating characteristic curve; Meta-analysis
METHODS

Data Identification

All relevant published studies relating to CT colonography were identified by computer-assisted search of the MEDLINE database from January 1996 to November 2005 using SilverPlatter’s MEDLINE (Ovid Technologies, New York, NY). References were retrieved using keywords that included (“virtual” or “CT” or “computed” or “CAT”) near “colon*.” Studies were considered eligible for inclusion in the meta-analysis if all subjects undergoing CT colonography also underwent endoscopic colonoscopy (as a reference standard) and the studies reported the per-patient sensitivities and specificities for polyp detection. Because specificities were generally reported on a per-patient basis, a per-patient sensitivity was then needed for constructing an sROC curve. Thus, the studies that only reported the per-polyp sensitivity but not the per-patient sensitivity were not included in the meta-analysis. The raw data for determining sensitivities and specificities were also needed to perform a continuity correction for all studies because some studies reported sensitivities or specificities equal to 100%. Studies were also excluded if they had a limited number of patients with disease or controls (n < 5) or had an excess number of colorectal cancers rather than benign polyps without subgrouping these 2 categories (ie, polyps vs cancers). In cases of multiple articles from the same institution, the dates for study inclusion were evaluated to ensure that the patients were not overlapping. Eligible reports were reviewed by the 2 authors to determine final eligibility.

STATISTICAL METHODS

Sensitivities and specificities were calculated using standard formulas and pooled using the inverse of the variance of the rates. sROC curves were constructed from the per-patient sensitivities and false positivities (1-specificities) of the studies. In this method, the sensitivity and false positivity are transformed into its logistic transform (also called the logits), defined as the natural log of the positivity rate/(1-positivity rate). Because some studies reported a sensitivity or specificity of 100%, 0.5 was added to each cell of the 2-by-2 tables (number of true positives, false positives, true negatives, and false negatives) for all of the studies. A linear regression was then performed using the difference between the logits of the true positives and false positives as the dependent variable and the sum of the logits of the true positive and false positives as the independent variable. Linear regression was performed using the Statistical Package for the Social Sciences for Windows version 10 (SPSS Inc, Chicago, Ill). The y-intercept is a measure of the diagnostic odds ratio, and the slope is a measure of how the odds ratio is dependent on the threshold. A slope of zero results in a symmetric, well-behaved sROC curve (“homogeneous case”), whereas slopes significantly different from zero result in an asymmetric sROC curve. The y-intercept and slope were then substituted in the sROC function to generate the sROC curve. The sROC curves and their 95% confidence intervals (CIs) were graphically constructed at 3 different sized thresholds for clinically significant polyps (any size, > 5 mm, and > 10 mm) using Sigma Plot for Windows version 8 (SPSS Inc).

The accuracy of the tests was assessed by 3 different statistical methods, the index Q*, the homogeneous area under the curve (AUC), and the exact AUC. The index Q* corresponds to the uppermost point on the sROC curve in which sensitivity equals specificity. The index Q* values, their standard error, and statistical comparisons were calculated using the method of Moses et al. The area under the sROC curve was calculated using the method of Walter. In sROC analysis, an index Q* or AUC close to 1.0 signifies a test with high discrimination, whereas values close to 0.5 imply poor discrimination. The calculation of the area under the sROC curve consists of integrating the sROC function. In many cases, the sROC function cannot be integrated by standard calculus methods, and thus a numeric integration method is required to calculate the exact AUC (Mathcad version 7, MathSoft, Inc, Cambridge, Mass). In cases in which the slope of linear regression is close to zero (known as the homogeneous case), the homogeneous AUC becomes a reasonable estimate of the exact AUC. The homogeneous AUC values can be calculated from a formula derived from integral calculus. The standard errors of the homogeneous and exact AUC values were calculated using the method described by Walter. In cases in which the slope was not significantly different from zero, both the homogeneous and exact AUCs were calculated. Statistical comparisons of the AUCs were performed using the formula of Hanley and McNeil. Finally, the 95% CIs of the sROC curves were constructed using the method described by Moses et al. Examples of 2 hypothetic test curves are shown in Figure 1 to further illustrate sROC analysis.

RESULTS

Included Studies

As shown in Table 1, 30 studies were included in the meta-analysis. Twenty-six studies primarily used 2-dimensional (2D) image reconstruction for
their initial analysis of CT colonography, 3 studies used a 3-dimensional (3D) reconstruction (also known as “fly-through”), and 1 study compared both methods (Table 1). The studies generally used the findings from endoscopic colonoscopy as the gold standard for detecting polyps. However, some of the studies attempted to further verify their endoscopic findings. Eight studies used segmental unblinding during colonoscopy, in which the endoscopist was notified of the results of the CT colonography during withdrawal from each segment of the colon. If a lesion was seen on CT colonography but not on colonoscopy, the segment was then reexamined. Van Gelder et al used a different approach in which a second-look colonoscopy was performed if the CT colonography revealed a discrepant finding from the first colonoscopy that could not be explained by residual stool, haustral folds, or a prominent ileocecal valve. Almost all of the studies used a cathartic preparation for CT colonography. However, 1 study alternatively used “fecal tagging” by adding diatrizoate meglumine and diatrizoate sodium 48 hours before CT colonography. Finally, 1 study also compared the results of air contrast barium enema with CT colonography and endoscopic colonoscopy.

### Diagnostic Characteristics of Computed Tomographic Colonography

The pooled sensitivities and specificities for CT colonography are shown in Table 1 at different thresholds. The sensitivity and specificity of CT colonography increased as the threshold for defining a clinically significant polyp became more stringent (ie, increasing size). Furthermore, the pooled per-patient sensitivities of CT colonography was higher for polyps greater than 10 mm (0.82, 95% CI, 0.76-0.88) compared with polyps 6 to 10 mm (0.63, 95% CI, 0.52-0.75) and polyps 0 to 5 mm (0.56, 95% CI, 0.42-0.70). The sROC curves of CT colonography at the 3 different thresholds (any size, > 5 mm, and > 10 mm) are shown in Figures 2 to 4. The diagnostic parameters of the sROC curves (homogeneous AUC, exact AUC, and index Q*) also improved with increasing size of the threshold for polyps (Figures 2-4). These results suggest that CT colonography had a reasonable diagnostic accuracy for large polyps but not for smaller polyps.

### Comparison of 2-Dimensional Versus 3-Dimensional Computed Tomographic Colonography

We also subdivided the studies by whether the study used 2D versus 3D (“fly-through”) algorithms for the primary image analysis as described by Hara. As shown in Table 2, the sensitivities for 3D CT colonography were not significantly different from 2D CT colonography at either a threshold of greater than 5 mm or a threshold greater than 10 mm. (There were insufficient data to evaluate a threshold of any size polyp.) By using sROC analysis at the polyp threshold greater than 5 mm, all 3 parameters (homogeneous AUC, exact AUC, and index Q*) were higher for 2D, but the comparisons did not reach statistical significance ($P = .28$, $P = .29$, and $P = .29$, respectively). At the threshold greater than 10 mm, all 3 parameters were higher.
for 3D, but the comparisons did not reach statistical significance (P = .26, P = .24, and P = .23, respectively).

**Comparison with Air Contrast Barium Enema**

We identified 2 published studies that prospectively compared the diagnostic accuracy of air contrast barium enema with endoscopic colonoscopy.\(^4\)\(^1\),\(^4\)\(^9\) In the study by Winawer et al,\(^4\)\(^9\) the threshold used was a polyp of any size. After segmental unblinding was performed, the sensitivity and specificity for air contrast barium enema were 0.38 and 0.86, respectively. By using a threshold greater than 5 mm, Rockey et al\(^4\)\(^1\) reported a per-patient sensitivity and specificity for air contrast barium enema of 0.41 and 0.82, respectively. At a higher threshold of greater than 10 mm, the per-patient sensitivity and specificity were 0.48 and 0.90, respectively.\(^4\)\(^1\) Although the limited number of studies prevented a formal sROC analysis for air contrast barium enema, one could compare the diagnostic characteristic values of air contrast barium enema with the sROC curves for CT colonography. As shown in Figures 2 to 4, the sensitivities of barium enema (at their corresponding value for false positivity) were below the lower limit of the 95% CI for the sROC curve.

**Comparison with Colonoscopy**

Nine of these studies attempted to further verify the findings of colonoscopy either by segmental unblinding or second-look colonoscopy in selected cases. Seven of these studies reported the per-patient sensitivities and specificities of colonoscopy (Table 3). The pooled sensitivities for endoscopic colonoscopy were significantly higher than CT colonography at a threshold of greater than 5 mm and greater than 10 mm. (There were an insufficient number of studies to evaluate the diagnostic characteristics of endoscopy colonoscopy using a threshold of any size polyp.) The sROC curves for endoscopic colonoscopies are shown in Figures 5 and 6. Because of the small number of studies, meaningful CIs could not be constructed. At a polyp threshold greater than 5 mm, colonoscopy had better diagnostic accuracy than CT colonography using the following sROC parameters: homogeneous AUC (0.999 ± 0.002 vs 0.888 ± 0.027, P < .001), exact AUC (0.998 ± 0.006 vs 0.884 ± 0.033, P < .005), and index Q* (0.987 ± 0.013 vs 0.819 ± 0.028, P < .001). At a threshold of greater than 10 mm, colonoscopy was significantly more accurate than CT colonography as determined by homogeneous AUC (0.999 ± 0.001 vs 0.928 ± 0.036, P = .05) and index Q* (0.990 ± 0.004 vs 0.863 ± 0.043, P < .005). However, the comparison of exact AUC did not reach statistical significance (0.999 ± 0.001 vs 0.898 ± 0.063, P = .10).
DISCUSSION

Colonoscopy is currently thought to be the most accurate diagnostic test for polyp screening. However, several surveys of American populations have reported a low adherence to endoscopic screening. CT colonography has been proposed as a less-invasive alternative to colonoscopy. Most studies have reported a patient preference for CT colonography compared with colonoscopy. However, the accuracy of CT colonography is currently debated. Two recent meta-analyses have statistically pooled the sensitivities and specificities of CT colonography for polyp detection. Sosna et al reported that the sensitivity varied with polyp size and that the specificity was 95% for polyps greater than 10 mm. Mulhall et al also reported similar findings but noted a consistently high specificity across polyp sizes. However, a statistical pooling of the sensitivity and specificity of a test may not be adequate for several reasons. First, sensitivity and specificity represent a tradeoff in the radiologist’s criteria for distinguishing a polyp from an artifact. By loosening the criteria, a test will become more sensitive but less specific. As explained in the examples provided by Irwig et al, a pooled estimate of sensitivities and specificities may give an inaccurate assessment of the diagnostic accuracy. Furthermore, it may be difficult to compare pooled estimates of the diagnostic characteristics of 2 different tests. For example, if one test has a pooled sensitivity of 90% and a specificity of 75%, and another test has a pooled sensitivity and specificity both

<table>
<thead>
<tr>
<th>Modality</th>
<th>Sensitivity*</th>
<th>Specificity*</th>
<th>Homogeneous AUC†</th>
<th>Exact AUC†</th>
<th>Index Q* †</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5 mm</td>
<td>0.73 (0.62-0.85)</td>
<td>0.88 (0.84-0.92)</td>
<td>0.898 ± 0.0356</td>
<td>0.898 ± 0.038</td>
<td>0.829 ± 0.038</td>
</tr>
<tr>
<td>2D</td>
<td>0.75 (0.52-0.98)</td>
<td>0.76 (0.65-0.88)</td>
<td>0.843 ± 0.037</td>
<td>0.843 ± 0.036</td>
<td>0.775 ± 0.034</td>
</tr>
<tr>
<td>&gt;10 mm</td>
<td>0.81 (0.74-0.87)</td>
<td>0.96 (0.94-0.98)</td>
<td>0.910 ± 0.053</td>
<td>0.866 ± 0.087</td>
<td>0.842 ± 0.059</td>
</tr>
<tr>
<td>3D</td>
<td>0.80 (0.61-0.99)</td>
<td>0.96 (0.94-0.98)</td>
<td>0.975 ± 0.022</td>
<td>0.971 ± 0.023</td>
<td>0.927 ± 0.038</td>
</tr>
</tbody>
</table>

AUC = area under the curve; 3D = 3-dimensional; 2D = 2-dimensional.

Table 2 Diagnostic Characteristics of Two-Dimensional Versus Three-Dimensional Computed Tomographic Colonography

Table 3 Diagnostic Characteristics of Endoscopic Colonoscopy

Pooled value

None of the comparisons between 2D and 3D CT colonography were statistically significant.

*Rate and 95% confidence limits in parentheses.
†Value ± standard error.
equal to 85%, it is difficult to determine which test is more accurate.

We thus used both statistical pooling and an sROC approach to evaluate CT colonography. Our analysis revealed a reasonable sensitivity and specificity for detecting patients with large polyps (ie, > 10 mm) but decreased accuracy with thresholds of smaller polyp sizes. sROC analysis also suggests that CT colonography seems to be more accurate than air contrast barium regardless of polyp size. The use of 3D software algorithms (“fly-through”) for initial analysis of the CT images did not seem to significantly improve the accuracy. Finally, CT colonography is less accurate than endoscopic colonoscopy for polyp detection.

Although some authorities have advocated systems that initially construct axial or 2D images and reserve 3D analysis for problem solving, others have suggested that initial evaluation using 3D analysis may be more accurate. Our meta-analysis was not able to validate the superiority of 3D systems. Furthermore, using 3D algorithms for initial evaluation is more time-consuming than using 2D images. It is possible that differences in the quality of endoluminal image reconstruction among commercially available systems may account for some of the variability in the accuracy of 3D image analysis. It is also possible that combining 3D image reconstruction with fecal tagging may improve the diagnostic accuracy of CT colonography.

Endoscopic colonoscopy had significantly higher sensitivities and specificities than CT colonography at either a threshold of greater than 5 mm or greater than 10 mm. By using sROC analysis, colonoscopy had a higher diagnostic accuracy than CT colonography at a threshold of greater than 5 mm. At a threshold of greater than 10 mm, the homogeneous AUC and index Q* were significantly higher, whereas the comparison of exact AUC did not reach statistical significance. Thus, endoscopic colonoscopy seems to be more accurate than CT colonography, particularly for detecting polyps less than 10 mm. The design of these studies included in our meta-analysis does have a bias in favor of colonoscopy because they used colonoscopy as the “gold standard.” The use of segmental blinding during colonoscopy may not necessarily correct this bias because polyps that are hidden in mucosal folds may be difficult to locate even after a second “unblinded” attempt.

Given that colonoscopy is more accurate for detecting smaller polyps, it is reasonable to evaluate their clinical significance. In a retrospective study, Kulling et al reported that 8.5% of polyps less than 5 mm and 15.5% of polyps less than 10 mm have advanced features (villous component or severe dysplasia). Another concern is the rate of progression of polyps to more advanced lesions. Stryker et al also reported 8% of polyps greater than 10 mm can progress to invasive cancer, although the natural history of small to medium-sized polyps is much less certain. In patients with hereditary nonpolyposis syndrome, the progression of small polyps to cancer may require only 2 to 3 years. Given the limitations of CT colonography for small polyps, we think that this modality should not be considered as a first-line screening test in patients with a strong family history of colorectal cancer. If CT colonography is used for colon cancer screening in situations without a strong family history, we think that it should be repeated more frequently than the recommended surveillance schedules for colonoscopy.

CONCLUSION

CT colonography shows both promise and limitations. It seems to be superior to air contrast barium enema but less accurate than colonoscopy, particularly for detecting small polyps. Thus, it may be an appropriate colon cancer screening test in situations in which small polyps are not clinically relevant. However, it should be performed more frequently than surveillance colonoscopy given the potential progression of small polyps to more advanced neoplasia.
ACKNOWLEDGMENTS

The authors thank Drs. Calvin Eng and Steven Lehrer for their mathematic assistance and Dr. Vahram Haroutunian for his assistance in constructing the graphs.

References


38. Rosman and Korsten sROC Comparison of CT Colonography and Colonoscopy


<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Initial Method of Image Reconstruction for CT Colonography</th>
<th>Verification of Colonoscopy</th>
<th>Comments</th>
<th>True Positives/ Total with Polyps of Any Size</th>
<th>True Negatives/ Total Without Polyps of Any Size</th>
<th>True Positives/ Total with Polyps &gt; 5 mm</th>
<th>True Negatives/ Total Without Polyps &gt; 5 mm</th>
<th>True Positives/ Total with Polyps &gt; 10 mm</th>
<th>True Negatives/ Total Without Polyps &gt; 10 mm</th>
<th>True Positives/ Total with Polyps 0-5 mm</th>
<th>True Positives/ Total with Polyps 6-10 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen et al[19]</td>
<td>Lower GI symptom, GI bleeding, or prior polyps</td>
<td>2D</td>
<td>No</td>
<td>Ultra low-dose technique</td>
<td>45/59</td>
<td>59/78</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Cotton et al[20]</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>Both methods, 3D data</td>
<td>Yes</td>
<td>Not reported</td>
<td>47/104</td>
<td>462/496</td>
<td>25/42</td>
<td>547/558</td>
<td>Not reported</td>
<td>22/62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton et al[20]</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>Both methods, 2D data</td>
<td>Yes</td>
<td>Not reported</td>
<td>41/104</td>
<td>449/496</td>
<td>23/42</td>
<td>535/558</td>
<td>Not reported</td>
<td>18/62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dachman et al[21]</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>2D</td>
<td>No</td>
<td>Two observers; results averaged</td>
<td>7/16</td>
<td>25/28</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Fenlon et al[22]</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>2D</td>
<td>No</td>
<td>Not reported</td>
<td>42/51</td>
<td>41/49</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Fletcher et al[23]</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>2-D</td>
<td>No</td>
<td>Half of the patients received oral contrast; used combined supine and prone view</td>
<td>Not reported</td>
<td>Not reported</td>
<td>114/130</td>
<td>36/50</td>
<td>82/96</td>
<td>78/84</td>
<td>Not reported</td>
<td>32/34</td>
</tr>
<tr>
<td>Ginnerup Pedersen et al[24]</td>
<td>Cancer/polyp surveillance, lower GI symptoms, anemia</td>
<td>2D</td>
<td>Yes</td>
<td>Not reported</td>
<td>40/44</td>
<td>101/104</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Hara et al[25]</td>
<td>Suspected polyp on barium enema or prior polyp</td>
<td>2D</td>
<td>No</td>
<td>Two observers, results averaged</td>
<td>Not reported</td>
<td>Not reported</td>
<td>16.5/25</td>
<td>28/45</td>
<td>9/12</td>
<td>52.5/58</td>
<td>12.5/30</td>
<td>7.5/13</td>
</tr>
<tr>
<td>Study</td>
<td>Study Population</td>
<td>Initial Method of Image Reconstruction for CT Colonography</td>
<td>Verification of Colonoscopy</td>
<td>Comments</td>
<td>True Positives/Total with Polyps of Any Size</td>
<td>True Negatives/Total Without Polyps of Any Size</td>
<td>True Positives/Total with Polyps &gt; 5 mm</td>
<td>True Negatives/Total Without Polyps &gt; 5 mm</td>
<td>True Positives/Total with Polyps &gt; 10 mm</td>
<td>True Negatives/Total Without Polyps &gt; 10 mm</td>
<td>True Positives/Total with Polyps 0-5 mm</td>
<td>True Positives/Total with Polyps 6-10 mm</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Hara et al</td>
<td>Patients at high risk of neoplasia</td>
<td>2D</td>
<td>No</td>
<td>Single and multidetector row helical CT systems, results pooled</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>12/14</td>
<td>205/223</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Hoppe et al</td>
<td>Patients referred for clinically indicated colonoscopy (nonscreening)</td>
<td>2D</td>
<td>Yes</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>26/34</td>
<td>51/58</td>
<td>19/20</td>
<td>65/66</td>
<td>Not reported</td>
<td>7/14</td>
</tr>
<tr>
<td>Iannaccone et al</td>
<td>Screening, GI symptoms, or prior polyps</td>
<td>2D</td>
<td>No</td>
<td>Cancers excluded from analysis</td>
<td>69/72</td>
<td>47/50</td>
<td>83/86</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Iannaccone et al</td>
<td>Screening, prior, or family history of polyps, hemato-positive stools, iron-deficient anemia, lower GI symptoms</td>
<td>2D</td>
<td>Yes</td>
<td>Fecal tagging achieved with oral contrast solutions; 3 readers, results averaged</td>
<td>71/79</td>
<td>114.3/124</td>
<td>44/48</td>
<td>128.3/138</td>
<td>17/17</td>
<td>186/186</td>
<td>27/31</td>
<td>27/31</td>
</tr>
<tr>
<td>Johnson et al</td>
<td>Asymptomatic patients at higher than average risk for colon cancer</td>
<td>2D</td>
<td>No</td>
<td>Scans were double-read</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>30/47</td>
<td>625/656</td>
<td>Not reported</td>
<td>45/69</td>
</tr>
<tr>
<td>Johnson et al</td>
<td>Symptomatic or at increased risk of colorectal neoplasia</td>
<td>2D</td>
<td>No</td>
<td>Multicenter</td>
<td>Not reported</td>
<td>Not reported</td>
<td>42/61</td>
<td>23/32</td>
<td>33/44</td>
<td>36/49</td>
<td>Not reported</td>
<td>9/17</td>
</tr>
<tr>
<td>Laghi et al</td>
<td>Lower GI symptoms Patients undergoing colonoscopy</td>
<td>2D</td>
<td>No</td>
<td>Cancers excluded from analysis</td>
<td>15/17</td>
<td>32/34</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Lui et al</td>
<td>Patients undergoing colonoscopy</td>
<td>2D</td>
<td>No</td>
<td>Two observers; results averaged; used thin sections</td>
<td>8/10</td>
<td>13/15</td>
<td>6/6</td>
<td>17.5/19</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>2/4</td>
</tr>
<tr>
<td>Study</td>
<td>Study Population</td>
<td>Initial Method of Image Reconstruction for CT Colonography</td>
<td>Verification of Colonoscopy</td>
<td>Comments</td>
<td>True Positives/Total with Polyps of Any Size</td>
<td>True Negatives/Total Without Polyps of Any Size</td>
<td>True Positives/Total with Polyps &gt;5 mm</td>
<td>True Negatives/Total Without Polyps &gt;5 mm</td>
<td>True Positives/Total with Polyps &gt;10 mm</td>
<td>True Negatives/Total Without Polyps &gt;10 mm</td>
<td>True Positives/Total with Polyps 0-5 mm</td>
<td>True Positives/Total with Polyps 6-10 mm</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>----------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Mani et al</td>
<td>Patients undergoing colonoscopy</td>
<td>2D</td>
<td>No</td>
<td>Pooled 3 radiologists’ interpretations rather than computer-aided-detection</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>7.3/10</td>
<td>29.8/31</td>
<td>Not reported</td>
</tr>
<tr>
<td>Mendelson et al</td>
<td>Symptomatic or at increased risk of colorectal neoplasia</td>
<td>2D</td>
<td>No</td>
<td>Some patients had supine scan only, whereas others had both prone and supine</td>
<td>19/47</td>
<td>44/53</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Morrin et al</td>
<td>Symptomatic or at increased risk of colorectal neoplasia</td>
<td>2D</td>
<td>No</td>
<td>Most patients also received IV contrast; results of IV contrast and noncontrast groups pooled</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>12/14</td>
<td>52/52</td>
<td>11/20</td>
</tr>
<tr>
<td>Pescatore et al</td>
<td>Symptomatic or at increased risk of colorectal neoplasia</td>
<td>2D</td>
<td>No</td>
<td>Two teams of observers; results averaged</td>
<td>17.5/24</td>
<td>17/26</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Pickhardt et al</td>
<td>Asymptomatic patients for screening or positive family histories</td>
<td>3D</td>
<td>Yes</td>
<td>Oral contrast given</td>
<td>Not reported</td>
<td>Not reported</td>
<td>149/168</td>
<td>848/1065</td>
<td>45/48</td>
<td>1138/1185</td>
<td>Not reported</td>
<td>104/120</td>
</tr>
<tr>
<td>Pineau et al</td>
<td>Patients undergoing colonoscopy (screening and nonscreening)</td>
<td>2D</td>
<td>Yes</td>
<td>Oral contrast</td>
<td>55/89</td>
<td>82/116</td>
<td>38/45</td>
<td>133/160</td>
<td>18/20</td>
<td>175/185</td>
<td>17/44</td>
<td>20/25</td>
</tr>
<tr>
<td>Rex et al</td>
<td>Asymptomatic patients for screening or polyps detected on flexible sigmoidoscopy</td>
<td>2D</td>
<td>No</td>
<td></td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>8/10</td>
<td>32/36</td>
<td>3/16</td>
</tr>
</tbody>
</table>
Table 1  Characteristics of Studies Included in the Meta-Analysis (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Initial Method of Image Reconstruction for CT Colonography</th>
<th>Verification of Colonoscopy</th>
<th>Comments</th>
<th>True Positives/Total with Polyps of Any Size</th>
<th>True Positives/Total Without Polyps of Any Size</th>
<th>True Positives/Total with Polyps &gt;5 mm</th>
<th>True Positives/Total Without Polyps &gt;5 mm</th>
<th>True Positives/Total with Polyps &gt;10 mm</th>
<th>True Positives/Total Without Polyps &gt;10 mm</th>
<th>True Positives/Total with Polyps 0-5 mm</th>
<th>True Positives/Total with Polyps 6-10 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rockey et al⁴¹</td>
<td>Positive fecal occult blood, iron-deficient anemia, prior history or strong family history of neoplasia, or hematochezia</td>
<td>2D Yes</td>
<td>Patients also underwent double contrast barium enema</td>
<td></td>
<td>Not reported</td>
<td>Not reported</td>
<td>85/155</td>
<td>409/459</td>
<td>37/63</td>
<td>529/551</td>
<td>48/92</td>
<td>210.e4</td>
</tr>
<tr>
<td>Taylor et al⁴²</td>
<td>Lower GI symptoms, GI bleeding, or other</td>
<td>2D Yes</td>
<td></td>
<td></td>
<td>15/25</td>
<td>25/29</td>
<td>10/19</td>
<td>34/34</td>
<td>9/10</td>
<td>44/44</td>
<td>5/6</td>
<td>1/9</td>
</tr>
<tr>
<td>Van Gelder et al⁴³</td>
<td>Prior or family history of neoplasia</td>
<td>3D No</td>
<td>Selected the high radiation dose scans</td>
<td></td>
<td>16/27</td>
<td>6/23</td>
<td>9/10</td>
<td>22/40</td>
<td>Not reported</td>
<td>Not reported</td>
<td>7/17</td>
<td>Not reported</td>
</tr>
<tr>
<td>Van Gelder et al⁴⁴</td>
<td>Personal or family history of polyps or increased cancer risk</td>
<td>3D Yes</td>
<td>(second-look colonoscopy in selected cases)</td>
<td></td>
<td>87.5/141</td>
<td>33/108</td>
<td>35/45</td>
<td>142/204</td>
<td>26/31</td>
<td>200.5/218</td>
<td>52.5/96</td>
<td>9/14</td>
</tr>
<tr>
<td>Wessling et al⁴⁵</td>
<td>Cancer screening or lower GI symptoms</td>
<td>2D No</td>
<td></td>
<td></td>
<td>10/15</td>
<td>25/33</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>5/9</td>
<td>2/3</td>
</tr>
<tr>
<td>Wong et al⁴⁶</td>
<td>Lower GI symptoms and polyp surveillance</td>
<td>2D No</td>
<td></td>
<td></td>
<td>16/27</td>
<td>41/44</td>
<td>Not reported</td>
<td>Not reported</td>
<td>7/8</td>
<td>63/63</td>
<td>9/17</td>
<td>0/2</td>
</tr>
<tr>
<td>Yee et al⁴⁷</td>
<td>Symptomatic or at increased risk of colorectal neoplasia</td>
<td>2D No</td>
<td></td>
<td></td>
<td>164/182</td>
<td>85/118</td>
<td>Not reported</td>
<td>Not reported</td>
<td>42/49</td>
<td>241/251</td>
<td>65/79</td>
<td>50/54</td>
</tr>
<tr>
<td>Yee et al⁴⁸</td>
<td>Both symptomatic and screening</td>
<td>2D No</td>
<td>Scanned in both supine and prone positions</td>
<td></td>
<td>103/114</td>
<td>56/68</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Pooled value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.73 (0.66-0.81)</td>
<td>0.77 (0.69-0.86)</td>
<td>0.77 (0.68-0.85)</td>
<td>0.84 (0.79-0.89)</td>
<td>0.82 (0.76-0.88)</td>
<td>0.96 (0.95-0.97)</td>
<td>0.56 (0.42-0.70)</td>
<td>0.63 (0.52-0.75)</td>
</tr>
</tbody>
</table>

CT = computed tomographic; GI = gastrointestinal; 2D = 2-dimensional; 3D = 3-dimensional; IV = intravenous.
ABSTRACT

Urinary incontinence is a common and potentially disabling condition affecting 15% to 30% of those aged 65 years and older. It adversely affects physical health, psychological well-being, and health care costs. Even when it is not curable, proper management of urinary incontinence can lead to improved quality of life for patients and caregivers. Despite its prevalence, many geriatric patients suffering from urinary incontinence are undiagnosed and untreated. Patients often do not report the problem, and health care professionals often do not ask about it. Health care professionals should therefore learn to identify, evaluate, and manage urinary incontinence based upon the available evidence and practice guidelines. Although the evidence base for specific recommendations for the office evaluation and management of geriatric urinary incontinence is limited, a basic evaluation to identify treatable causes of incontinence, referral of appropriate patients for further evaluation, and several noninvasive management strategies can greatly improve these symptoms in many older patients. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Urinary incontinence; Drug therapy; Behavior therapy; Referral and consultation; Office management

PATHOGENESIS OF URINARY INCONTINENCE

Normal urination requires the coordination of several physiological processes. Somatic and autonomic nerves carry information on bladder volume to the spinal cord and motor output is adjusted accordingly. The cerebral cortex exerts a predominantly inhibitory influence while the brain stem facilitates urination by coordinating urethral sphincter relaxation and detrusor contraction. As the bladder fills, sympathetic tone contributes to closure of the bladder neck and relaxation of the dome of the bladder, and inhibits parasympathetic tone. At the same time, somatic innervation maintains tone in the pelvic floor musculature, including the striated muscle around the urethra. When urination occurs, sympathetic and somatic tones diminish, resulting in decreased urethral resistance. Cholinergically mediated parasympathetic tone increases and causes the bladder to contract. Urine flow occurs when bladder pressure exceeds urethral resistance. Normal bladder capacity is 300 to 500 mL, and the first urge to void generally occurs between bladder volumes of 150 and 300 mL.

Continence depends upon intact micturition physiology and functional toileting ability. Age-related changes in the lower urinary tract occur in both continent and incontinent older persons. Compensatory mechanisms outside the lower urinary tract, such as mobility and cognition, play an
Acute Incontinence
Acute urinary incontinence refers to incontinence of sudden onset precipitated by potentially reversible, treatable conditions. Table 2 lists reversible conditions that can cause or contribute to geriatric urinary incontinence. Note that these factors can both precipitate acute incontinence, and contribute to chronic incontinence. While treating these conditions may not resolve the urinary incontinence, they may make other treatments more effective.5,6

Chronic Urinary Incontinence
Table 3 reviews the basic types, causes, and treatments of chronic urinary incontinence.6

Urge Incontinence. Urge incontinence is the most common and bothersome type of urinary incontinence in older ambulatory care patients. It is characterized by abrupt urgency with leakage of urine that can be large or small. Urge incontinence can be associated with frequency and nocturia (a syndrome known as overactive bladder). Among older women, urge incontinence may be accompanied by symptoms of stress incontinence. In older men, it commonly presents with the irritative symptoms of prostatic enlargement, including frequency and nocturia. Urge urinary incontinence is most often associated with detrusor overactivity, which may be age-related, idiopathic, secondary to central inhibitory pathway lesions, related to previous pelvic irradiation, or due to local bladder irritation of bladder sensory or motor innervation. Detrusor overactivity may coexist with impaired detrusor contractility, causing urge incontinence with an elevated postvoid residual in the absence of bladder outlet obstruction.8 The incomplete bladder emptying in these patients may result in urinary frequency and predispose them to urinary retention when they are treated with anticholinergic drugs.

Stress Incontinence. Stress incontinence is most often associated with weakened pelvic floor supporting tissues and consequent hypermobility of the bladder outlet and urethra. Leakage of urine occurs with increases in intra-abdominal pressure such as with cough, position changes, exercise, laughing, or sneezing. Stress incontinence is the second most common type of urinary incontinence in older women, where it is usually associated with lack of estrogen or previous vaginal deliveries or surgery. It is uncommon in men unless they have had urethral surgery or irradiation. Obesity can exacerbate urinary incontinence. Stress urinary incontinence is less commonly due to intrinsic sphincter deficiency. In women, intrinsic sphincter deficiency can occur with trauma from anti-incontinence surgery or with severe urethral atrophy and causes urine leakage with minimal increases in intra-abdominal pressure or continual leakage during physical activity.6

Overflow Incontinence. Overflow incontinence results from detrusor muscle weakness, bladder outlet obstruction, or both. Patients may experience symptoms of dribbling, weak urinary stream, intermittency, hesitancy, frequency, and nocturia. A diagnosis of overflow incontinence may be challenging because of the overlap in symptoms with other types of urinary incontinence.5 Anticholinergics, narcotics, and alpha-adrenergic agonists can contribute to overflow urinary incontinence (Table 2).

Functional Incontinence. Functional urinary incontinence is associated with the inability or lack of motivation to reach a toilet on time. Contributing factors include inaccessible toilets, mobility disorders, cognitive impairment, and psychological disorders. Patients with functional incontinence due to cognitive or mobility impairment require systematic toileting assistance as a component of their management.6

### Table 1. Adverse Effects of Urinary Incontinence

<table>
<thead>
<tr>
<th>Category</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical health</td>
<td>● Skin irritation/breakdown</td>
</tr>
<tr>
<td></td>
<td>● Recurrent urinary tract infections</td>
</tr>
<tr>
<td></td>
<td>● Falls/fractures</td>
</tr>
<tr>
<td></td>
<td>● Sleep disruption</td>
</tr>
<tr>
<td>Psychological health</td>
<td>● Isolation/social withdrawal</td>
</tr>
<tr>
<td></td>
<td>● Depression</td>
</tr>
<tr>
<td></td>
<td>● Anxiety</td>
</tr>
<tr>
<td></td>
<td>● Dependency</td>
</tr>
<tr>
<td>Social consequences</td>
<td>● Increased caregiver burden</td>
</tr>
</tbody>
</table>

Predisposition to institutionalization

Economic costs

- Treatment costs
  - Pads, diapers
  - Drugs
  - Surgery
  - Devices
- Complications of treatment
- Lost productivity

OFFICE EVALUATION OF INCONTINENCE IN OLDER PATIENTS

Many older patients who are bothered by urinary incontinence do not spontaneously complain about it. Thus, screening questions should be used to identify this condition. Simple questions included in a review of systems can be helpful, such as: “Do you have trouble with your bladder?” “Do you lose urine when you do not want to?” “Do you wear pads or adult diapers for protection?”

Once urinary incontinence has been identified, the main goals of office management are to diagnose and treat reversible causes, determine the predominant type of urinary incontinence to guide treatment, identify who needs further evaluation, prevent complications, and improve quality of life for patients and caregivers.

The basic urinary incontinence evaluation should include a focused history, targeted physical examination, urinalysis, and postvoid residual determination in most patients. In performing the history and physical examination, special attention should be given to mobility, mental status, medications, volume status, and accessibility of toilets that may be causing or worsening the incontinence. The history also should focus on the characteristics of the incontinence, the most bothersome symptoms, and the treatment goals and preferences of the patient and caregivers. Bladder records, voiding diaries, and standard symptom scales like the American Urological Association Symptom Inventory for men (Figure 1) can be helpful in characterizing symptoms and following treatment response. Physical examination should include abdominal, rectal, neurological, and genital/pelvic examinations. Selected patients should be considered for further urologic, gynecologic, or urodynamic evaluation (Table 5, Figure 2).

Several simple diagnostic tests can be performed in the office (Table 4). The International Continence Soci-
ety recommends a postvoid residual (PVR) urine measurement by a noninvasive method before institution of pharmacological or surgical treatment of incontinence.9 Patients with storage-specific symptoms (ie, frequency, urgency, incontinence) without complaints of decreased bladder emptying, with normal bladder sensation, and without anatomical, neurological, organ-specific, or co-morbid risk factors for retention may be assessed for bladder emptying by history and physical examination alone.9 In general, a PVR >50 mL is considered adequate bladder emptying, while a PVR >200 mL is considered inadequate emptying and warrants referral to a specialist.9 Although these guidelines can be useful in directing patient management, each patient and situation must be considered individually.

| Table 3 Basic Types, Causes, and Treatments of Persistent Urinary Incontinence |
|---|---|---|---|
| Types | Presentation | Common Causes | Primary Treatment |
| Stress | Involuntary loss of urine (usually small amounts) with increases in intra-abdominal pressure (eg, cough laugh, exercise) | Weakness of pelvic floor musculature and urethral hypermobility. Bladder outlet or urethral sphincter weakness | Regular voiding to avoid a full bladder Pelvic muscle exercises α-adrenergic agonist (controversial) Estrogen (topical) (controversial) Peri-urethral injection Surgical bladder neck suspension or sling | |
| Urge | Leakage of urine (variable but often larger volumes) because of inability to delay voiding after sensation of bladder fullness is perceived | Detrusor overactivity, isolated or associated with one or more of the following: - Local genitourinary condition such as tumors, stones, diverticuli, or outflow obstruction - Central nervous system disorders such as stroke, dementia, parkinsonism, spinal cord injury | Antimuscarinic drugs Topical estrogen (for severe vaginal atrophy or atrophic vaginitis) Bladder training (including pelvic muscle exercises) | |
| Mixed | Combination of urge and stress symptoms | Combination of above causes | One or a combination of above, targeting most bothersome symptom(s) first | |
| Overflow | Leakage of urine (usually small amounts) resulting from mechanical forces on an over-distended bladder, or from other effects of urinary retention on bladder and sphincter function | Anatomic obstruction by prostate, stricture, large cystocele Acontractile bladder associated with diabetes mellitus or spinal cord injury Neurogenic (detrusor-sphincter dyssynergy), associated with multiple sclerosis and other suprasacral spinal cord lesions | Surgical removal of obstruction Bladder retraining Indwelling catheterization | |
| Functional | Urinary accidents associated with inability to toilet because of impairment of cognitive or physical functioning, psychological unwillingness, or environmental barriers. | Severe dementia and other neurological disorders. Psychological factors such as depression and hostility. | Behavioral interventions with toileting assistance Environmental manipulations Incontinence undergarments and pads | |


**MANAGEMENT**

**General Principles**

There are several therapeutic options available for managing urinary incontinence in older patients. Table 3 outlines the primary treatments for the basic types of chronic incontinence. Treatment decisions should be individualized and will depend on the findings of the basic evaluation and preferences of the patient and the health care provider.

Supportive measures can be helpful in managing urinary incontinence and should be used with other, more specific treatment measures. Education about bladder health, environmental manipulation (such as safe lighted path to the bathroom), the appropriate use of toilet sub-
stitutes (urinals, bedside commodes), modifications of diuretic and fluid intake patterns, and good skin care are all important examples of such supportive measures.6

**Behavioral Interventions**

Behavioral interventions are well studied in the geriatric population and recommended by most guidelines as an initial approach to therapy. They are generally noninvasive, effective for the most common types of geriatric incontinence, and conducive to an outpatient primary care setting.10-12 Key aspects of behavioral therapy are outlined in Table 6.

Behavioral interventions can be categorized as “patient-dependent” and “caregiver-dependent” therapies. Patient-dependent interventions necessitate adequate function, learning capability, and motivation of the patient. Their goal is to restore a satisfactory pattern of voiding and continence.6 Pelvic floor muscle (Kegel) exercises are an essential component of patient-dependent behavioral interventions and involve repetitive con-
For each question, circle the answer that best describes your situation.

<table>
<thead>
<tr>
<th>Question</th>
<th>Not at all</th>
<th>Less than one in five times</th>
<th>Less than half of the time</th>
<th>About half of the time</th>
<th>More than half of the time</th>
<th>Almost always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In the past month, how often have you had a sensation of not emptying your bladder completely after you finished voiding?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. In the past month, how often have you had to urinate again less than 2 hours after you finished urinating before?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. In the past month, how often have you found you stopped and started again several times when you urinated?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. In the past month, how often have you found it difficult to postpone urination?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. In the past month, how often have you had a weak urinary stream?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. In the past month, how often have you had to push or strain to begin urination?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. In the past month, how many times did you typically get up to urinate from the time you went to bed until you arose in the morning?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Total: __________________

SCORING KEY: 0 to 7 = mild, 8 to 19 = moderate, 20 or more = severe


Figure 1  The American Urological Association, Seven Symptom Inventory for Benign Prostatic Hyperplasia

tractions and relaxations of the pelvic floor muscles. These exercises can be taught by having women squeeze the examiner’s finger during vaginal examination or by having men or women squeeze the examiner’s finger by contracting the anal sphincter during rectal examination. Many older patients have difficulty identifying and isolating pelvic floor muscles and may benefit from referral to a physical therapist or a nurse specialist. Although biofeedback and electrical stimulation have not been shown to increase the efficacy of behavioral therapy,

Table 5  Examples of Criteria to Refer an Older Patient with Incontinence for Further Urologic, Gynecologic, or Urodynamic Evaluation

- Surgery/irradiation involving the pelvic area or lower urinary tract within past 6 months
- Two or more symptomatic urinary tract infections in a 12-month period
- Marked pelvic prolapse on physical examination
- Post-void residual volume >200 mL measured by straight catheterization or bladder ultrasound
- Difficulty passing a 14-Fr straight urinary catheter
- Marked prostatic enlargement, prominent asymmetry, or induration of the prostatic lobes
- Greater than 5 red cells per high-power field on repeated microscopic examinations in the absence of infection
- Persistent bothersome symptoms after adequate trials of behavioral or drug therapy
these techniques can be helpful in patients who have difficulty isolating the proper muscles.

Caregiver-dependent interventions (regular toileting, prompted voiding) are useful in functionally disabled patients. The success of these interventions is largely dependent on caregiver knowledge and motivation, rather than on the patient’s physical function and mental status.6

**Drug Treatment**

Table 7 lists the various drugs used to treat the different types of incontinence. Drug treatment is often prescribed in conjunction with behavioral therapy, which can enhance its effectiveness.6 For urge or mixed incontinence, antimuscarinic drugs targeting bladder smooth muscle are most commonly used. A variety of drugs are available with similar efficacy in clinical trials.13,14 All can have bothersome systemic anticholinergic side effects, especially dry mouth, constipation, and exacerbation of gastroesophageal reflux. Reductions in urge incontinence in clinical trials are generally in the range of 60% to 80% (compared with 30% to 50% on placebo),10,14,15 with an overall incidence of dry mouth of approximately 20% (>5% severe).14 These medications may precipitate urinary retention, and patients at risk, including diabetics, those with elevated baseline postvoid residual volumes, and men with outflow obstruction should be monitored carefully.6 Elderly patients on anticholinergics should be monitored for worsening cognitive function or drug-induced delirium, as there have been reports of cognitive side effects mimicking dementia.16-18

The role of estrogen in treating geriatric incontinence is controversial. Topical estrogens may be used in postmenopausal women with urge incontinence associated with atrophic vaginitis or severe vaginal atrophy.19,20 However, combination oral hormone replacement therapy with estrogen...
and progestin has been associated with an increased incidence of incontinence.\textsuperscript{21,22} In men with benign prostatic enlargement, irritative voiding symptoms associated with urge incontinence may be effectively treated with an $\alpha$-adrenergic antagonist (average reduction in AUA scores of 4 to 6 points\textsuperscript{23}). Newer agents (alfuzosin and tamsulosin) may have less effect on blood pressure and should be considered for older men who have low blood pressure, orthostasis, or dizziness. If $\alpha$-adrenergic antagonist therapy alone does not control irritative voiding symptoms, the addition of an antimuscarinic drug should be considered. Although $\alpha$-adrenergic antagonists alone are probably not efficacious for long-term treatment of overflow incontinence,\textsuperscript{24} 5 $\alpha$-reductase inhibitors alone and in combination have been shown to reduce the voiding symptoms from benign prostatic hyperplasia, as well as the incidence of urinary retention.\textsuperscript{25}

Estrogens and $\alpha$-adrenergic agonists have been recommended as treatments for stress urinary incontinence, however, use of both of these medications is controversial. Estrogens may worsen urinary incontinence,\textsuperscript{21,22} and there is weak evidence to suggest that the use of alpha adrenergic agonists is better than placebo.\textsuperscript{26} Drug treatment for chronic overflow incontinence using a cholinergic agonist or an $\alpha$-adrenergic antagonist is rarely efficacious.\textsuperscript{6}

Drug treatment is often prescribed in conjunction with behavioral therapy, which can enhance the effectiveness of drug treatment.

**Surgery**

Surgery should be considered in selected geriatric patients who fail nonsurgical treatment and want further intervention, and in those with a major anatomic abnormality. For women with stress incontinence, surgery offers high short-term cure rates but is invasive, potentially morbid, and lacking long-term follow-up data. Stress incontinence in older women also may be effectively treated with periurethral injection of bulking agents.\textsuperscript{1} Surgery also may be indicated in men in whom incontinence is associated with anotomically or urodynamically documented outflow obstruction. For older males with postprostatectomy stress incontinence, periurethral injection of bulking agents may be helpful in milder cases. Artificial urinary sphincters may be considered, but their cure rates are only 50% and morbidity may be as high as 40%.\textsuperscript{27} The decision must be individualized, weighing carefully the degree to which the symptoms bother a patient and the potential risks and benefits of surgical treatment.

**Incontinence Undergarments and Catheters**

Incontinence undergarments can be helpful in patients with urinary incontinence but should be used appropriately. Although extensively marketed and readily available, these products are relatively expensive and are not covered by Medicare or most other insurance. They should generally not be used as the first response to urinary incontinence or before a diagnostic evaluation is performed. Many older patients prefer undergarments to other treatments.\textsuperscript{28} For patients who prefer or need these undergarments, proper fit and absorptive capabilities are important. Information on products and other educational material can be found on internet sites of national organizations such as the National

---

**Table 6: Key Aspects of Behavioral Therapy for Incontinence**

<table>
<thead>
<tr>
<th>Patient/caregiver education and resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Booklets</td>
</tr>
<tr>
<td>• Websites</td>
</tr>
<tr>
<td>○ American Geriatrics Society Foundation for Health in Aging (<a href="http://www.healthinaging.org">www.healthinaging.org</a>)</td>
</tr>
<tr>
<td>○ National Association for Continence (<a href="http://www.nafc.org">www.nafc.org</a>)</td>
</tr>
<tr>
<td>○ The Simon Foundation (simonfoundation.org)</td>
</tr>
<tr>
<td>○ American Urogynecological Society (<a href="http://www.augs.org">www.augs.org</a>)</td>
</tr>
<tr>
<td>○ National Institute on Aging (<a href="http://www.nia.nih.gov">www.nia.nih.gov</a>)</td>
</tr>
<tr>
<td>○ National Institute of Diabetes and Digestive and Kidney Diseases (<a href="http://www.niddk.nih.gov">www.niddk.nih.gov</a>)</td>
</tr>
<tr>
<td>○ American Urological Association (<a href="http://www.UrologyHealth.org">www.UrologyHealth.org</a>)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fluid management</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Avoid caffeine</td>
</tr>
<tr>
<td>• Limit nighttime fluids (for nocturia/nocturnal incontinence)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alleviation of constipation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Dietary fiber</td>
</tr>
<tr>
<td>• Fluid intake</td>
</tr>
<tr>
<td>• Appropriate use of stool softeners, laxatives, suppositories</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pelvic floor muscle exercises</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Self-taught or taught during examination</td>
</tr>
<tr>
<td>• Taught using biofeedback</td>
</tr>
<tr>
<td>• Home practice for strengthening</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use of exercises in situations that precipitate incontinence and for urge suppression</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Regular voiding to avoid a full bladder (stress incontinence), or prevent involuntary bladder contractions (urge incontinence)</td>
</tr>
</tbody>
</table>

| Regular toileting assistance (for mobility impaired patients) |

---

218 The American Journal of Medicine, Vol 120, No 3, March 2007
Antimuscarinic
Darifenacin 7.5-15 mg q day Increase bladder capacity Urge or mixed with urge predominant Dry mouth, constipation blurry vision, elevated intraocular pressure, cognitive impairment, delirium

Oxybutynin
Short-acting 2.5-5 mg tid Diminish involuntary bladder contractions
Long-acting 5-30 mg q day
Transdermal 3.9 mg patch q 4 days

Solifenacin
Tolterodine 5-10 mg q day Increase periurethral blood flow Stress Stress

Trospium
α-Adrenergic agonist* 20 mg bid Stimulate contraction of urethral smooth muscle Stress Headache, tachycardia, elevation of blood pressure

Pseudoephedrine
Topical estrogen* 30-60 mg tid, or 60-120 mg long-acting Strengthen periurethral tissues Stress Local irritation

Topical cream
0.5-1.0 g per day for 2 weeks, then twice weekly Increase periurethral blood flow Stress

Vaginal estradiol ring
One ring every 3 months

Vaginal tablets
One 25-μg tablet per day for 2 weeks, then twice weekly

Cholinergic agonists*
Bethanechol
Stimulate bladder contraction Overflow incontinence with atonic bladder Bradycardia, hypotension, bronchoconstriction, gastric acid secretion, diarrhea Postural hypotension

α-Adrenergic antagonists
Alfuzosin
10 mg qd Relax smooth muscle of urethra and prostatic capsule Urge incontinence and related irritative symptoms associated with benign prostatic enlargement

Doxazosin
1-8 mg qd†

Tamsulosin
0.4 mg qd

Terazosin
1-10 mg qhs†


*Weak evidence supporting use in treatment of urinary incontinence.
†Most studies have used the higher range dosages (doxazosin 4 to 8 mg and terazosin 5 to 10 mg).

Association for Continence (NAFC.org) and the Simon Foundation (Simonfoundation.org).

Urinary catheters should never be used as a means of convenience but may be appropriate management in specific patient situations. Clean intermittent catheterization is a treatment option for those with ongoing bladder emptying problems and high postvoid residual. Frequency of catheterization needs to be based on individual bladder volumes and patient tolerance. Chronic indwelling catheters should be used only after alternative management strategies have been considered. They are considered appropriate in patients with urinary retention causing persistent overflow incontinence, symptomatic infections, or renal dysfunction that cannot be corrected surgically or medically and cannot be managed practically with intermittent catheterization. Chronic indwelling catheters also can be considered for patients with skin wounds, pressure sores, or irritations being contaminated by incontinent urine and for patients who are terminally ill or severely impaired for whom bed and clothing changes are uncomfortable. While both indwelling catheters and diapers should not be considered as a first resort, a small proportion of patients may express a strong preference for this type of management of their urinary incontinence.

References


Updates in the Management of Gout
Michael P. Keith, MD, William R. Gilliland, MD, MHPE
Uniformed Services University of the Health Sciences, Bethesda, Maryland.

ABSTRACT
The majority of patients with gout are cared for by primary care physicians. Although both the physician and patient may easily recognize the acute arthritis of gout, errors in selecting the most appropriate medication and proper dose are common. The clinical stages of gout include asymptomatic hyperuricemia, intermittent gouty arthritis, and chronic tophaceous gout. Treatment of gout is usually considered after the first attack of arthritis, typically podagra. The aims of treatment are to alleviate the pain and inflammation associated with acute attacks, prevent future attacks, and decrease uric acid levels. Confusion frequently arises because certain medications such as colchicine may have dual purposes: to treat an acute attack and to suppress future attacks. The purpose of this management update is to provide practical advice about prescribing the proper medication considering both treatment goals and patient comorbidities. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Gout; Management; Colchicine; Nonsteroidal anti-inflammatory agents; Xanthine oxidase; Probencid; Febuxostat

As with many diseases, it is important to recognize that the course of classic gout passes through various stages: asymptomatic hyperuricemia, acute intermittent gout, and chronic tophaceous gout. Asymptomatic hyperuricemia is a common biochemical abnormality and is strongly associated with an increased incidence of gout. In one longitudinal study, subjects with uric acid levels between 7.0 and 8.0 mg/dL had a 3% cumulative risk of developing gout, whereas subjects with uric acid levels of 9.0 mg/dL or more had a cumulative risk of 22%. However, the majority of patients with hyperuricemia never develop gout.

Acute intermittent gouty arthritis is generally the first clinical manifestation of gout. Classically described as an acute monoarthritis associated with severe pain escalating over a 6- to 12-hour period, swelling, and erythema, gout is easily recognized by physicians and patients. Because systemic features such as fever and chills may accompany acute attack, it is important to make a definitive diagnosis of gout by aspiration and demonstration of characteristic monosodium urate crystals.

Finally, the transition from acute intermittent gout to chronic tophaceous gout develops over 10 years or more. This transition is characterized by the intercritical periods between gouty attacks becoming less defined, diminished intensity of the pain, and persistent joint abnormalities.

ACUTE MANAGEMENT OF GOUT
Acute gouty arthritis most commonly presents as monoarthritis of the lower extremities with the first metatarsophalangeal joint being the most common. Early, appropriate therapy is associated with a decreased duration of the attack and less pain and disability. Currently available treatment options for acute gout include colchicine, nonsteroidal anti-inflammatory drugs (NSAIDs), and corticosteroids. Urate-lowering agents such as allopurinol or probenecid should not be started or discontinued during an acute attack.

Colchicine, a plant derivative, inhibits leukocyte activation and migration and is most effective when given in the first 24 to 48 hours of the attack. It is relatively contraindicated in patients with renal insufficiency. Some authors advocate its use in patients when the diagnosis is suspected, but not confirmed; if the patient does not respond after up to

0002-9343/$ -see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjmed.2006.02.044
Inflammatory arthritis. Intravenous or intramuscular corticosteroids are also be useful in oligo- and polyarticular presentations. In the case of acute monoarthritis resulting from gout, intra-articular corticosteroids may be the treatment of choice. Oral corticosteroids should be considered when patients decline arthrocentesis or have failed treatment with colchicine and/or an NSAID. Systemic corticosteroids may also be useful in oligo- and polyarticular presentations. Intravenous or intramuscular corticosteroids are also options, especially in patients who cannot take medications by mouth. Adrenocorticotropic hormone has also been shown to be effective in acute gout and is thought to be effective even in patients who are adrenally insufficient, possibly through a melanocortin receptor as shown in a rat model of inflammatory arthritis.

A summary of treatment options for acute gout is presented in Table 1.

**CHRONIC MANAGEMENT OF GOUT**

The long-term management of gout consists of antihyperuricemic or urate-lowering therapy. In general, indications for urate-lowering therapy include 2 or more gout attacks per year, tophaceous gout, erosive arthritis on radiographs, and uric acid kidney disease, including urate nephropathy, uric acid nephropathy, and uric acid nephrolithiasis. A serum uric acid level of less than 6.0 mg/dL is generally recommended as an initial target for antihyperuricemic therapy, because a serum uric acid below this level has been associated with a reduced frequency or prevention of gout attacks.

Two classes of drugs are available for use as urate-lowering agents. The xanthine oxidase inhibitors block metabolic steps in the synthetic pathway of uric acid. Drugs likely to be encountered by physicians in the United States are allopurinol and febuxostat, the first new therapeutic agent for gout developed since the advent of allopurinol in the 1960s. Oxypurinol, a metabolite of allopurinol, is not widely available in the United States and therefore will not be discussed. The uricosurics, such as probenecid, act to increase renal uric acid secretion. Other uricosurics include sulfinpyrazone and benz bromarone, which are also not readily available in the United States.

Prophylaxis of gout is an important concern when initiating antihyperuricemic therapy because such therapy may cause a gout flare. In fact, a recent clinical trial showed that 64% of patients treated with allopurinol experienced a flare...
during the first year of therapy. In general, gout should be in remission before starting urate-lowering therapy. However, the recommended duration of remission is open to debate even among rheumatologists. Agents that are available for gout prophylaxis in this setting include colchicine or an NSAID. When given for prophylaxis, colchicine should be given at a dose of no more than 0.6 mg twice daily with lower doses required in patients with chronic renal insufficiency. Similarly, low doses of an NSAID may be used, for example, naproxen 250 mg twice daily. Low-dose prednisone (≤10 mg daily) may also be considered in patients who cannot tolerate colchicine or an NSAID because of comorbidities.

Management of a gout flare in a patient receiving concomitant antihyperuricemic therapy is the same as for acute gout except that urate-lowering agents should not be stopped. The duration of prophylactic therapy varies widely among practicing rheumatologists with recommendations ranging from 3 to 12 months. One suggestion is to continue prophylactic therapy until the serum uric acid level has been below goal for 3 to 6 months and the patient has had no acute attacks during this period.2

**XANTHINE OXIDASE INHIBITORS**

Allopurinol is a purine analogue inhibitor of xanthine oxidase that blocks the conversion of xanthine to uric acid. Allopurinol is probably the agent most commonly used by practicing rheumatologists, despite the fact that approximately 90% of patients with gout are uric acid under-excreters. This is because it is effective in both uric acid under-excreters and overproducers, and a 24-hour measurement of urine uric acid is not required before its use. Typical doses range from 100 mg daily up to 300 mg daily in patients with normal renal function, but occasionally higher doses may be required. The initial dose should be reduced in patients with impaired renal function because of the increased rate of adverse events at higher starting doses in these patients. Patients taking allopurinol should have periodic laboratory monitoring to include complete blood count, hepatic enzymes, and serum uric acid. We recommend testing the uric acid at baseline and subsequently quarterly until the target uric acid level is reached in the patient.

In regard to drug interactions, allopurinol can increase the plasma levels of theophylline and warfarin. Some patients who require allopurinol may be currently taking azathioprine or its metabolite, 6-mercaptopurine. In this case, the dose of azathioprine/6-mercaptopurine should be reduced by 50%, because these drugs are metabolized through the xanthine oxidase pathway and failure to adjust the dose can lead to severe myelosuppression.2

The prescribing physician must be aware that side effects may occur in up to 20% of allopurinol-treated patients. The most common include gastrointestinal intolerance and skin rash. More severe reactions include fever, myelosuppression, toxic epidermal necrolysis, hepatitis, and vasculitis. The most severe reaction is the allopurinol hypersensitivity syndrome. Its manifestations include fever, rash, eosinophilia, progressive kidney failure, and possibly death. Although this may occur in any patient treated with allopurinol, it is most common in patients with chronic kidney disease who are taking diuretics. Patients should be advised to stop the medication and contact their physician should fever or rash develop.

Patients who develop a reaction to allopurinol may be considered for allopurinol desensitization. It has been shown to be effective and can be conducted by oral or parenteral administration. In general, patients requiring allopurinol desensitization should be referred to an allergist.

Febuxostat is a novel, nonpurine inhibitor of xanthine oxidase that is awaiting Food and Drug Administration approval. A recent large clinical trial showed that it is effective in reducing serum uric acid levels and has a favorable safety profile.6 Febuxostat is metabolized in the liver, and thus, dose adjustments for renal impairment are not required.2 Despite the fact that its exact role in the management of gout has not been determined, febuxostat is a welcome addition to the therapeutic armamentarium available for chronic gout. Although cost may be an issue with the use of febuxostat, it will likely be useful in patients who are intolerant to allopurinol or in patients who, because of chronic renal insufficiency, are not candidates for probenecid. The use of febuxostat in lieu of allopurinol will avoid possible drug interactions in patients who require therapy with azathioprine or 6-mercaptopurine. Likewise, it may prove to be useful as add-on therapy in patients in whom a serum uric acid level of less than 6.0 mg/dL cannot be reached because of dose limitations to allopurinol due to chronic kidney disease. The dosing of febuxostat, based on clinical trials, is likely to be 80 mg or 120 mg once daily.6

**URICOSURIC AGENTS**

Probenecid is the ideal antihyperuricemic drug in young patients with gout who have normal renal function, are under-excreters of uric acid, and have no history of nephrolithiasis. Measurement of a 24-hour urine uric acid concentration should be performed in patients for whom probenecid therapy is initiated. Patients on a regular purine diet who excrete less than 800 mg of uric acid in the urine per day are considered under-excreters.

Physicians should counsel their patients that probenecid can cause nephrolithiasis, and therefore they should maintain an adequate urine output. Furthermore, patients should be advised that doses of aspirin greater than 81 mg per day will diminish its efficacy.2 Addition of a xanthine oxidase inhibitor may be necessary if the patient cannot achieve a serum urate level of less than 6 mg/dL on probenecid alone.

**ADJUNCTIVE THERAPIES**

In addition to the standard agents discussed in the chronic management of gout, other drugs have been found to have
salutary effects on serum uric acid. Vitamin C was shown to have a mild urate-lowering effect at a dose of 500 mg daily. The effect of higher doses is not known. Medications used for comorbid conditions may likewise have a beneficial effect in gout. Fenofibrate should be considered in gout patients with hypertriglyceridemia, and losartan or amlodipine should be considered in gout patients with hypertension as all of these agents have been demonstrated to have mild uricosuric effects. The authors do not, however, recommend these agents instead of the standard agents indicated for the management of chronic gout.

Lifestyle and medication changes also play an adjuvant role in the management of gout. With the increasing popularity of potent uric acid-lowering agents, low purine diets have been less emphasized in the management of gout. Although it is difficult to manage gout by diet alone, it is reasonable for patients to limit their consumption of purine-rich foods, such as meats (especially organ meats), seafood (particularly shellfish and anchovies), and vegetables and legumes. Through a variety of mechanisms, excessive alcohol intake also has been associated with increased frequency of gout attacks. Primary care providers should also be aware of a number of medications that cause hyperuricemia because of decreased excretion of urate, including cyclosporine, nicotinic acid, furosemide, thiazide diuretics, ethambutol, pyrazinamide, and aspirin. If possible, alternative medications should be considered. A summary of basic principles for managing gout is outlined in Table 2.

CONCLUSION
Gout is a gratifying disease for physicians to manage because of the dramatic response if appropriate therapy is initiated promptly. Unfortunately, gout remains one of the most frequently mismanaged diseases in both emergency departments and primary care settings. The management approach needs to be comprehensive and individualized, taking into account factors unique to the patients such as comorbidities and disease manifestations and potential side effects of the medications.

References

<table>
<thead>
<tr>
<th>Table 2 Ten Principles for Managing Gout</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Observing intracellular monosodium urate crystals in joint fluid is the only way to make a definitive diagnosis of gout.</td>
</tr>
<tr>
<td>2. Asymptomatic hyperuricemia rarely requires therapy.</td>
</tr>
<tr>
<td>3. Colchicine, NSAIDs, and corticosteroids are equally efficacious in the management of acute gout.</td>
</tr>
<tr>
<td>4. Consider uric acid-lowering therapy in patients with two or more attacks per year.</td>
</tr>
<tr>
<td>5. A serum uric acid level of 6.0 mg/dL is the initial target level when using uric acid-lowering therapy.</td>
</tr>
<tr>
<td>6. Urate-lowering agents should neither be stopped nor started during an acute gout flare.</td>
</tr>
<tr>
<td>7. Use concomitant prophylaxis with oral colchicine or NSAIDs when initiating urate-lowering therapy for a period of 3 to 12 months.</td>
</tr>
<tr>
<td>8. Allopurinol, rather than uricosurics, should be used in patients with renal stones, tophaceous gout, or a glomerular filtration rate of less than 60 mL/min, or who are urate overproducers.</td>
</tr>
<tr>
<td>9. Febuxostat, a novel nonpurine xanthine oxidase inhibitor, may be a useful alternative medication in patients with renal insufficiency.</td>
</tr>
<tr>
<td>10. Lifestyle modification and avoidance of medications that increase uric acid levels may also play an adjunctive management role.</td>
</tr>
</tbody>
</table>

NSAID = nonsteroidal anti-inflammatory drug.
PRESENTATION
A negative test result proved misleading in the case of a 51-year-old woman from Chihuahua, Mexico. During one of her frequent visits to a daughter in the United States, the woman presented with a 2-month history of intermittent hemoptysis, weight loss, and night sweats. She had no history of chest pain, dyspnea on exertion, fever, chills, hematemesis, melena, or epistaxis, and she denied use of tobacco, alcohol, or injectable drugs. However, she had type 2 diabetes mellitus, hypertension, and 12 years earlier, she underwent mastectomy, chemotherapy, and radiotherapy for cancer of the right breast.

Five years before her current admission, the patient was diagnosed with pulmonary tuberculosis based on a positive (30 mm) tuberculin skin test and a radiographic finding of a left apical cavitation (Figure 1). At that time, her husband had confirmed pulmonary tuberculosis, so although her diagnosis was never verified with a culture, she received antitubercular therapy for 4 months.

ASSESSMENT
Physical examination revealed the following: blood pressure, 170/88 mm Hg; heart rate, 88 beats per minute; respiratory rate, 18 breaths per minute; temperature, 99.7°F (37.6°C); and oxygen saturation of 98% on 2 liters oxygen by nasal cannula. Fine rales were audible at both lung bases. A cardiac examination was normal, and she had no hepatosplenomegaly or lymphadenopathy. Laboratory studies disclosed a white blood cell count of 6.9 x 10^3/mm^3, hemoglobin of 10.6 g/dL, platelet count of 288 x 10^3/mm^3, and a creatinine of 1.1 mg/dL. Liver function tests and a coagulation profile were normal. Computed tomography of the chest showed right upper-lobe infiltrate with a thin-walled-cavity (Figure 2), a 2.7 pleural-based scar in the left apex, scattered ground-glass opacities in both lung fields, and mediastinal and right hilar lymphadenopathy.

Mycobacterial and fungal sputum smears were reported negative, as were tests for serum Histoplasma antigen and serum Coccidioides antibody. Chronic granulomatous inflammation was evident with right upper-lobe tissue biopsy guided by computed tomography; again, mycobacterial and fungal smears were negative. Nonetheless, her history suggested recurrent pulmonary tuberculosis, so isoniazid, rifampin, and ethambutol were initiated. Subsequent mycobacterial and fungal cultures of sputum and lung tissue were negative.

DIAGNOSIS
Three months later, the patient developed massive hemoptysis and right upper lobectomy was performed. Fungal stain and culture of the lung tissue showed Coccidioides species (Figure 3). Enzyme immunoassay for Coccidioides IgG was positive, and the complement fixation antibody titer was 1:8. Pulmonary coccidioidomycosis was diagnosed.

Coccidioides species are dimorphic fungi endemic to the southwest regions of the United States, Northern Mexico, and parts of Central and South America. These include C. immitis, the cause of coccidioidomycosis, and C. posadasii, but no clinical laboratory method differentiates between them. In the soil, Coccidioides species exist as mycelia that mature to form the infective arthroconidia. Nearly all infections are the result of inhaling arthroconidia, which become spherules in the lungs. Cutaneous inoculations with extension to regional lymph nodes have been reported but are exceedingly rare. Most patients with coccidioidomycosis live in endemic areas. Clinicians in other locales can encounter the disease in travelers from endemic regions or in former residents who experience reactivation of latent infections.

At least one half to two thirds of all infections are subclinical or sufficiently mild not to prompt medical evaluation. Primary Coccidioides infections most frequently
Figure 1  Five years before the patient’s current admission, a left apical cavity was discovered on a radiograph.

Figure 2  A right upper-lobe infiltrate with a thin-walled-cavity was noted on a computed tomography of the chest.
manifest as community-acquired pneumonia 1 to 3 weeks after exposure. Distinguishing coccidioidomycosis from other etiologies is usually difficult without laboratory confirmation, such as detection of anticoccidioidal antibodies in serum or identification of \textit{Coccidioides} species in sputum or another respiratory specimen. Approximately 5% to 10% of infections trigger pulmonary sequelae, usually nodules or cavities. The latter, which can be present at any stage of the primary infection, are typically solitary, near the pleura, thin-walled, and under 4 cm in diameter. Half regress after 2 years without antifungal therapy. In the chronic phase of infection, patients are generally asymptomatic. Some patients develop a chronic fibrosing pneumonic process characterized by infiltrates and cavitations that commonly involve more than 1 lobe. Lesions might cause local or systemic symptoms, such as night sweats and weight loss. This form of infection is not common among persons with T-cell deficiencies but seems to be associated with diabetes or preexisting pulmonary fibrosis.

Direct examination of sputum and other respiratory specimens may reveal the diagnostic spherules of \textit{Coccidioides}, particularly in patients who produce copious sputum or who have multilobar infiltrates. Bronchoscopy is normally performed in immunosuppressed or severely ill patients, especially if they have diffuse infiltrates. In one study of 30 patients with an abnormal x-ray and a culture or histology indicative of coccidioidomycosis, bronchoscopy yielded the diagnosis in 69% of cases (after patients with solitary pulmonary nodules were excluded from analysis). Only 32% of prebronchoscopy sputum samples produced a positive culture.

Serology is the most common method of diagnosis in primary coccidioidal infections. Even minimally reactive results are often diagnostically important, and a negative serologic test never excludes coccidioidal infection. Repeating serology 1 or more times over 2 months increases the sensitivity of serologic diagnosis, especially for recently acquired infections.

The differential diagnosis for focal or multifocal cavitary lung diseases includes neoplasms such as bronchogenic carcinomas and lymphomas, pulmonary infarct, septic embolism, immunologic disorders such as Wegener’s granulomatosis and rheumatoid nodules, and infection (see Table). In cavitary lung disease caused by \textit{Mycobacterium tuberculosis}, sputum smears are likely to be positive for acid-fast bacilli. In a study of 977 patients with culture-proven pulmonary tuberculosis, sputum smears were positive for acid-fast bacilli in 20% to 40% of patients with minimal disease, 60% to 70% of patients with moderately severe disease, and 90% to 95% of patients with advanced cavitary disease. Among 155 patients with tuberculosis, positive smears were found in 98% of patients with cavitary disease compared with 70% of patients with noncavitary disease.

Patients can have coexisting pulmonary tuberculosis and coccidioidomycosis. A review of 43 such cases noted that simultaneous onset occurred in less than one third of them. Symptoms of each disease usually emerged sequentially.

**MANAGEMENT**

Our patient’s antitubercular therapy was replaced by itraconazole, which produced a good clinical response. While multiple diagnostic tests for presumed pulmonary tuberculosis were negative, other tests also failed to reveal her diagnosis before surgery. We postulate that she initially had culture-negative pulmonary tuberculosis based on her exposure history, tuberculin skin test conversion, and the left apical cavity, which subsequently healed.

This case highlights 3 important clinical points when evaluating cavitary lung disease. First, pulmonary coccidioidomycosis cannot be excluded with negative serological tests or cultures. In addition, tuberculosis is an unlikely cause for cavitary lung disease if sputum smears for acid-fast bacilli are negative. Finally, pulmonary tuberculosis and coccidioidomycosis can coincide in patients at risk.

![Figure 3](image-url) Fungal stain and culture of lung tissue removed during lobectomy revealed the causative organism.
References


An Absorbing Problem

Naveen Pemmaraju, Themistocles Dassopoulos, Alison R. Moliterno
Johns Hopkins University School of Medicine

PRESENTATION

Often people live with chronic, non-life-threatening conditions for many years before the underlying diagnosis is revealed. The current case is a 30-year-old woman who was diagnosed with iron deficiency anemia during her first pregnancy 13 years earlier. Numerous trials of oral iron supplementation failed to correct her anemia. She was referred to a hematologist for further evaluation.

ASSESSMENT

A life-long resident of Baltimore, Maryland, the patient lived there with her husband and children and worked full-time as a human resources manager. She was a non-smoker and reported no alcohol or illicit drug use. Her only medicine was a multivitamin which she took occasionally. She practiced no dietary restrictions, including red meat. The patient’s past medical history was significant for 3 uncomplicated vaginal deliveries. Her mother had a history of atherosclerotic vascular disease, and her father and brother were well. None of her 3 children, ages 7, 8, and 12, had any medical history, including iron deficiency.

Upon review of her symptoms, the patient reported cravings for ice, restless legs, and fatigue. She described episodes of lightheadedness, occasional left-hand numbness and tingling, and 1 pre-syncopal episode. She denied gastrointestinal bleeding, menorrhagia, blood transfusion history, pain in her mouth, dietary restrictions or changes, food allergies, weight changes, diarrhea, constipation, dyspepsia, lactose intolerance, nausea, vomiting, abdominal bloating or pain, rashes or skin lesions.

The physical examination revealed a pale woman in no distress. She was afebrile, blood pressure of 118/63, pulse 76, normal respirations, and a weight of 56.9 kilograms (125.1 pounds). She was not orthostatic. Her sclera and conjunctivae were normal. Cardiac exam revealed a normal S1 and S2 and a faint systolic ejection murmur. The lungs were clear to auscultation bilaterally, abdomen was soft and nontender, and no organomegaly was noted. She was hemoccult negative. There was mild bilateral pretibial edema. Vibratory sense and light touch were both intact. No skin lesions were noted.

Laboratory testing showed a white blood cell count of 4140/cu mm, hemoglobin 8.7 g/dL, hematocrit 31.2%, mean cell volume 69.8 FL, and platelets 358,000 K/cu mm. The peripheral blood smear revealed hypochromic, microcytic red blood cells (Figure A). Absolute reticulocyte count was 54.7 K/cu mm (24-88), ferritin 3 ng/mL (10-300). Thyroid stimulating hormone, triiodothyronine, and vitamin B12 levels were normal. In the recent past, other studies included a normal lactate dehydrogenase, normal red cell CD55 and CD59 expression, and normal serum protein electrophoresis. A hemoglobinopathy screen did not reveal either abnormal hemoglobins or elevated levels of hemoglobin F or A2. A serum erythropoietin level was elevated at 107 mIU/mL (4-19), and a bone marrow exam revealed absent iron stores.

To further investigate the source of the iron deficiency, a malabsorption workup was initiated. Studies drawn at the time of diagnosis revealed an elevated IgG gliadin antibody of 39 (negative <11), elevated IgA gliadin antibody of >100 (negative <11), Endomysial IgA antibody (EMA) titer of 1:160, and anti-tissue transglutaminase (tTG) IgA antibody of 83 μ/mL (>8 positive). The patient was referred for upper endoscopy. Biopsies of the duodenum showed villous blunting, increased intraepithelial lymphocytes, crypt hyperplasia, and increased chronic inflammation consistent with celiac disease (Figure B). The remainder of the malabsorption workup revealed normal levels of vitamins A, D, E, and folate, a normal prothrombin time, and a normal bone density study.

DIAGNOSIS

This patient was diagnosed with celiac disease. Celiac disease is an autoimmune enteropathy due to an inflammatory reaction to gluten-containing grains. Celiac disease is associated with a spectrum of small bowel malabsorption; the proximal duodenum is the primary site of iron absorption. It is now recognized that celiac disease, which was formerly thought to be rare, is common among those of European/North American heritage, with estimates as high as 1 out of every 120 to 300 people.1 Predisposition to celiac disease...
disease is associated with specific major histocompatibility complex haplotypes, although only a minority of individuals with these haplotypes develops celiac disease, indicating that other genes and environmental factors are required. There is a varied and long list of presenting symptoms in patients with celiac disease that result in delay of diagnosis as a patient is referred to a number of subspecialists. The common clinical manifestations include: dermatitis herpetiformis, diarrhea, iron deficiency, osteoporosis, unexplained weight loss, vitamin or folate deficiency, associated autoimmune diseases (eg, type I diabetes), and unexplained peripheral neuropathy. Recent studies indicate that iron deficiency anemia is the most frequent presenting feature of celiac disease. Similar to our patient, up to 50% of adult patients with celiac disease do not report significant diarrhea. Unsworth reported that up to 10% of patients referred to a gastroenterologist with iron deficiency anemia are diagnosed with unsuspected celiac disease. In a prospective screening study, 1 in 44 patients with iron deficiency anemia was diagnosed with histologically proved celiac disease, compared to only 1 out of 498 patients without iron deficiency anemia. Similar to our patient, half of the female celiac patients in this trial were premenopausal. The clinical diagnosis of celiac disease has been augmented by well-performing serologic assays. In patients with celiac disease, anti-endomysial antibody (EMA) testing has a high specificity (as high as 99%) and sensitivity (as high as 100%). More recently the anti-tissue transglutaminase 2 (tTG) IgA enzyme-linked immunosorbent assay has been shown to have less inter-observer variability, to be less expensive than EMA, and to have an overall sensitivity of greater than 90%. In a systematic review, Rostom found that the traditionally used anti-gliadin antibody was not as sensitive or specific as EMA and anti-tTG in the detection of celiac disease. The gold standard diagnostic tool is still the small intestine biopsy. Conducted during upper endoscopy, multiple biopsies are collected from the distal duodenum. These will classically demonstrate hyperplastic crypts, blunting of villi, and increased numbers of lymphocytes and plasma cells in the lamina propria. A combination of better laboratory testing and greater clinical awareness of celiac disease are contributing to the rising incidence of cases. Many patients may not exhibit the full clinical spectrum of classical celiac disease, emphasizing that serologic screening should be initiated in cases where iron deficiency, especially iron deficiency refractory to oral supplementation as in our patient, is present. The National Institutes of Health has recently launched a celiac awareness campaign, with resources for physicians and patients.

**MANAGEMENT**

The patient started a gluten-free diet. The patient was encouraged to and subsequently did contact a celiac support group which was thought to be beneficial as well. Two and a half months after the gluten-free diet was instituted, a repeat anti-tTG study was negative. For her severe iron deficiency anemia, she received intravenous iron supplementation. Four months after her visit to the hematologist, the patient was found to be making a remarkable recovery both in terms of her iron deficiency anemia (hemoglobin recovered to 14.2 gm/dL, restless legs and pagophagia resolved) and also in terms of her nutritional status, as she had gained 7 pounds since starting the gluten-free diet.

**References**


**Figure**

A. Peripheral blood smear is noted for microcytic and hypochromic red cells with anisocytosis and poikilocytosis. B. Duodenal biopsy is noted for blunting of villi, hyperplastic crypts and lymphocytic infiltration of the lamina propria.
Pressured to Appear

Rashid M. Rashid, BS,a Rajesh Yalamanchili, MD,b Amor Khachemoune, MD, CWS,c

aMD-PhD Program, Loyola University Chicago Stritch School of Medicine, Chicago, Ill; bState University of New York Downstate Medical Center, Department of Dermatology, Brooklyn, NY; and cRonald O. Perelman Department of Dermatology, New York University Medical Center, New York, NY

PRESENTATION

Sometimes patients endure a problem for quite a long time before seeking care. Such was the case of a 37-year-old woman who had had skin-colored papules on the medial aspects of her heels for more than 10 years. She reported intermittent foot pain after standing for prolonged periods of time. The patient was not a marathon runner or involved in any other extreme physical activities. Her medical and family histories suggested no contributory factors.

ASSESSMENT

The patient was a healthy, thin woman in no distress. An examination of her feet revealed multiple, minimally-tender, soft, compressible, skin papules. These ranged in color from skin-toned to yellowish and measured 3 to 4 mm on the medial aspect of both feet (Figure). However, when the patient put weight on the plantar surface of her feet, the lesions became larger, measuring 5 to 6 mm in diameter. No other significant physical findings were discovered.

DIAGNOSIS

Physical exertion, whether recreational or not, produces innumerable physical injuries and patient complaints. Heel pain is commonly encountered and seen by all physicians.1 However, the majority of reports and reviews commonly overlook dermatologic aspects of such pain, focusing instead on areas such as mechanical or structural dysfunction.2 Our patient’s discomfort had a dermatologic origin: piezogenic pedal papules.

These lesions were first described in 1968 as dermatoceles in a report on a patient with papules along the medial aspect of his heel and complaints of foot pain.3 The authors postulated that the papules were herniations of fatty tissue through the connective tissue trabeculae of the heel.

The observed prevalence of these lesions is quite variable in the literature, with no established consensus.4,5 However, multiple authors have concurred that this often-overlooked finding is quite common in the general population, occurring with similar frequency in children and adults.6,7 Unless these papules induce discomfort, the diagnosis is not commonly made during a regular examination.

On clinical presentation, the papules are round, skin-colored to yellow, and 2 to 10 mm in diameter. Classically, they develop on the medial aspect of the heel, but other locations are possible.7 For example, piezogenic papules have been reported on the wrist and palm.8-10 It is also important to note that the morphological appearance can be atypical. In one case, a solitary piezogenic papule became ulcerated, thereby resembling an adnexal tumor and requiring histopathology for accurate diagnoses.11 Ordinarily, the papules become grossly visible when the patient is standing, and they tend to recede when weight is removed from the foot.10

Various etiologies have been proposed for piezogenic pedal papules, including vigorous physical activity, hereditary factors, normal but repeated pressure forces in susceptible individuals, and in more recent years, acquired collagen defects such as Ehlers-Danlos syndrome.4,5,12,13 The widely accepted pathogenesis involves degeneration of the septa and trabeculae in the stroma of the connective tissue at the affected site. Subcutaneous fat protrudes through these defects to produce papules. It is believed that small papules are more likely to remain asymptomatic than larger lesions, which form after trabecular degeneration causes the fusion of smaller fat chambers.5,9

Pain is an occasional finding with piezogenic pedal papules. Since 1968, it has been attributed to ischemia produced when fat is forced into the dermis along with associated vasculature and nerves.3 Histologically, the papules are characterized by degeneration or obliteration of the trabeculae and septa within the subcutaneous fat and the destruct-
tion of the elastic fibers normally present within the overlying dermal connective tissue.9

It is interesting to note that the current terminology of piezogenic pedal papules is a minor misnomer. "Piezogenic" is derived from the Greek word "piezein"—to press—implying an origin secondary to pressure. However, though believed to be a type of hernia-like process, the actual trigger for the pathogenesis of piezogenic pedal papule is not known with certainty. Nevertheless, it is accepted that these papules are usually only visible upon application of pressure or when the patient is standing, thus allowing some merit to the use of the term “piezogenic.”

MANAGEMENT

Several treatment modalities have been proposed for the management of painful piezogenic pedal papules. These include simple analgesia; weight loss; the use of supportive external pressure devices; taping of the heel; and even surgical intervention in severe, refractory cases.14 Still, there is a paucity of studies specifically evaluating the efficacies of these therapeutic options. Recently, the use of local electro-acupuncture has garnered attention as preliminary studies offer some promise of relief in symptomatic patients.15 Further studies are clearly warranted to rigorously assess the best means of treatment for these patients.

Our patient was treated with conservative measures and was subsequently lost to follow-up. In summary, piezogenic pedal papules are a benign dermatological finding. However, it is essential to recognize and diagnose these lesions in order to adequately treat those symptomatic patients for whom this condition may become debilitating. Also, when suspecting this diagnosis, remember to examine your patient in a standing position. Finally, increasing awareness of these lesions is of great importance. This will promote further scientific reporting, analysis of additional therapeutic modalities, and exploration of potential prophylactic options.

ACKNOWLEDGMENT

The authors want to thank Dr. Benjamin N. Lockshin for his assistance with this manuscript.

References

PRESENTATION
An injury to the skin seemingly led to a terrible course of events for one patient, but the fundamental source of trouble was far below the surface of her body. A 77-year-old woman with a history of diabetes mellitus and arteriopathy was admitted to the emergency department for cellulitis of her left foot and multiple necrotic lesions on her left limb. Two weeks before, after skin trauma to the left tibia, she developed additional violaceous nodules, which rapidly turned into deep purple skin discoloration and extremely painful necrotic skin ulcers (Figures 1 and 2).

ASSESSMENT
On physical examination, the patient had a fever of 101.3°F (38.5°C), multiple necrotic lesions, and necrosis of the third toe, complicated by cellulitis. Her foot was cold, and distal pulses were not present. No abnormalities were observed on cardiac, abdominal, or pulmonary exams. Laboratory analyses showed normal renal function, coagulation, and serum calcium and phosphate levels. Testing for cryoglobulinemia and the lupus anticoagulant were negative. The patient’s white blood count was 12.3 x 10^3/cells/mm^3 (3% segmented neutrophils); her C-reactive protein level was 170 mg/L; and an arteriography revealed distal stenosis of the arteries with no possibility of revascularization, as well as severe calcifications. Despite appropriate antibiotics and repeated debridement, her condition worsened. Amputation became necessary. Knee disarticulation and pathology were performed (Figure 3).

DIAGNOSIS
Pathology confirmed a diagnosis of calciphylaxis, a rare, necrotizing skin condition. It occurs in less than 5% of patients with end-stage renal failure, and it is extremely rare in patients who have normal renal function, as ours did.1

The disorder, originally described in 1898, was reproduced in nephrectomized rats and given its name by researcher Hans Selye, who suggested that calciphylaxis was the result of a precise sensitization and subsequent challenge.1-3

Patients who develop calciphylaxis generally have deranged calcium-phosphate metabolism. Specifically, most have longstanding, inadequately-controlled renal failure and secondary hyperparathyroidism, which enhances the release of phosphate and calcium from bone. Calciphylaxis is more frequently encountered among Caucasians and 3 times more common among women than men. Additional risk factors in patients who are already metabolically predisposed are obesity; recent, abrupt weight loss; diabetes mellitus; hypoalbuminemia; and use of certain medications, including iron dextran and warfarin.1-10

In vitro models indicate that the combination of hyperphosphatemia and low-normal plasma calcium will result in an increased calcium-phosphate product and the tendency to cause calcification of vessels. It has been suggested that a calcium-phosphate product greater than 70 mg^2/dL^2 is associated with calciphylaxis.4-6 However, recent studies evaluating vascular calcification in nondialysis patients have found that the smooth muscle cells play an active role, including an increased expression of osteopontin, which stimulates vascular calcification.1-8

Vascular smooth muscle cells in the normal artery wall naturally express potent inhibitors of calcification, such as matrix Gla protein; its absence results in spontaneous medial calcification. In calciphylaxis, an imbalance exists between local vascular promoters and inhibitors of the calcification process. For example, promoters include altered vascular smooth muscle cells and the aforementioned osteopontin, while inhibitors include osteoprotegerin and matrix GLA protein. High levels of 125-dihydroxyvitamin D3 deleteriously affect the vascular smooth cell phenotype and can cause medial wall calcification. One interesting observation: despite the high incidence of vascular calcification reported in hemodialysis patients, the incidence of tissue necrosis with vascular calcification has been estimated at just 1 case per 100 hemodialysis patients per year.1,4
The most common clinical feature of calciphylaxis is exquisite pain when the ulcer is touched or when the wound dressing is removed. Initially, violaceous mottling or livedo reticularis in a stellate pattern is observed. Slightly-raised nodules or plaques with deep-purple skin discoloration might also be noted. Often, patients will have multiple lesions, and in 90% of patients, these manifestations are located on the limbs. Other locations have also been reported in anecdotes.

As lesions progress, the interior becomes necrotic. Calcifications are generally observed at the center of the ulcer, and the skin surrounding the ulcer has a firm consistency. Of course, skin ulcerations are associated with disorders other than calciphylaxis. When clinicians are confronted with necrotic skin lesions in association with fever and renal failure, other conditions to be considered in the differential diagnosis include: leucocytoclastic vasculitis, lupus erythematosus, Wegener’s granulomatosis, and polyarteritis nodosa.

Diagnosis is most frequently achieved with incisional biopsy of the ulcer. Punch biopsies may not be adequate because the quantity of tissue obtained is often insufficient. Pathology typically demonstrates calcification within the medial portion of medium-sized arterioles and extensive intimal hyperplasia and fibrosis. The diameters of the affected vessels range from 30 to 600 μm. Microthrombi are commonly seen.

MANAGEMENT

Our patient underwent disarticulation of the left knee because there was no possibility of revascularization, and uncontrolled sepsis was a concern. She improved and was discharged home with a cosmetic prosthesis less than 2 months after her admission. Six months later, she noticed 2 new lesions of calciphylaxis on the right leg. These were controlled with repeated debridement and grafting.

Oclusion of small vessels is a feature of cryoprecipitate disorders as well. Cryoglobulinemia may be present in end-stage renal disease and manifest with a classic palpable purpura or skin necrosis. Protein C or S deficiency, coumarin necrosis, antiphospholipid syndrome, pancreatic panniculitis, cholesterol emboli, subacute bacterial endocarditis, and disseminated intravascular necrosis should also be included in the differential diagnosis.

Diagnosis is most frequently achieved with incisional biopsy of the ulcer. Punch biopsies may not be adequate because the quantity of tissue obtained is often insufficient. Pathology typically demonstrates calcification within the medial portion of medium-sized arterioles and extensive intimal hyperplasia and fibrosis. The diameters of the affected vessels range from 30 to 600 μm. Microthrombi are commonly seen.

The management of calciphylaxis is both medical and surgical. Trigger factors—for example, trauma, calcium supplementation, elevated phosphate levels—should be identified and when possible, eliminated. Some authors advocate increasing the number of hemodialysis treatments for optimizing both calcium and phosphate clearance. Surgical treatment consists of repeated debridement and skin grafting or amputation where necessary. In addition, parathyroidectomy should be considered even in patients without renal failure, though conclusive evidence supporting this approach is still lacking. Calciphylaxis is a highly morbid disease, and it is fatal in 60% to 80% of affected
patients. Those who do not die of septicemia or renal failure frequently require amputation.1,20

References
A 59-year-old African American female had been binging on cocaine for 4 days when she developed a brief episode (<1 minute) of unwitnessed syncope. The incident, which occurred during ambulation, prompted her hospital visit.

The patient had no chest pain, shortness of breath, or palpitations associated with or prior to fainting, and no prodromal or postdromal symptoms were noted. Her past medical history included asthma, alcohol abuse, and substance abuse. She was not taking any medications at the time of admission. Two similar episodes had occurred in the last 6 months, but she did not seek any medical care. The patient was unclear about any cocaine use prior to these 2 syncopal events. There was no family history of similar complaints.

Bradycardia and a left-sided carotid bruit were detected during an otherwise unremarkable physical examination. A tilt test was negative for orthostatic hypotension. Laboratory data did not reveal any electrolyte abnormalities, but a urine drug screen was positive for cocaine and negative for amphetamines, barbiturates, benzodiazepines and phencyclidine. The patient’s ECG on admission showed bradycardia (50 beats per minute), diffuse, deep T-wave inversions, and prolonged QT and QTc intervals of 660 msec and 600 msec respectively (Figure 1). This was compared with a routine ECG obtained during a previous admission, which showed less prominent T-wave inversions in the inferior lateral leads, bradycardia (46 beats per minute), and a normal QTc interval of 414 msec (Figure 2). Serial cardiac enzymes—creatine kinase, creatine kinase-MB, and troponins—were within the normal limits.

A 2-dimensional echocardiogram displayed left ventricular hypertrophy and normal left ventricular systolic function (ejection fraction >55%). Computerized tomography of the head did not disclose any evidence of intracranial masses, bleeds, or midline shift. A carotid angiogram showed 90% stenosis in the left internal carotid artery; a coronary angiogram revealed normal coronary arteries. By the fifth day of hospitalization, the patient’s QTc interval shortened to within normal limits (440 msec). Abnormal T wave morphology persisted.

Requests for reprints should be addressed to Diwakar Jain, MD, Director of Nuclear Cardiology, Drexel University College of Medicine, 245 N 15th Street, MS 470, Philadelphia, PA, 19102.

E-mail address: Diwakar.Jain@drexelmed.edu

0002-9343/$ -see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjmed.2007.01.015
Long QT syndromes, which can be inherited or acquired, are characterized by prolongation of the QTc interval on the ECG (>470 msec in women; >450 msec in men) and an increased predisposition to polymorphic ventricular tachycardia.5 The acquired syndrome can be triggered by drugs, including members of the antiarrythmics, antibiotics, antidepressants, antihistamines, and antipsychotics.1 Electrolyte abnormalities (hypomagnesemia, hypokalemia, and hypocalcemia), stroke, heart failure, HIV infection, myocardial ischemia and infarction can also induce long QT syndrome.1

Cocaine affects cardiovascular electrophysiology in 3 different ways. The drug prevents the reuptake of norepinephrine from the synaptic cleft, accentuating the action of norepinephrine on target organs.4 It also acts like a class I antiarrhythmic agent by blocking sodium channels, which inhibits membrane depolarization.4 In addition, cocaine affects cardiac repolarization by blocking potassium channels encoded by the human ether-a-go-go-related gene, subsequently inhibiting the rapid component of the outward rectifier potassium current. Ventricular repolarization is delayed, an effect manifested on the ECG as a prolonged QTc interval.5 Blockade of the outward rectifier potassium current can also accentuate the remaining repolarization currents, enhancing repolarization differences already present in myocardium. These might manifest on a surface ECG as intensification of T-wave changes, including T-wave inversions and alternans.7

Marked QTc prolongation (> 550 msec) secondary to cocaine abuse is relatively rare and has been described in a few case reports.8,9 Furthermore, in all these cases, QTc prolongation recovered within 72 hours. In our case report, QTc prolongation persisted for 5 days after the last occurrence of cocaine abuse. This observation emphasizes that apart from ischemia-mediated acute myocardial injury, cocaine can have profound and potentially fatal electrophysiological effects on the heart. Since cocaine-induced QTc prolongation can act as a nidus for fatal arrhythmias, perhaps medications known to prolong the QTc interval should be avoided or used cautiously in patients with a history of cocaine abuse.

Figure 1 An ECG done on admission showed sinus bradycardia (50 beats per minute); deep T-wave inversions in leads V3-V6, I, II, III and AVF; a prolonged QTc interval (600 msec); and a prolonged QT interval (660 msec). (The QTc was calculated using Bazett’s formula.)

Figure 2 An ECG performed 4 months prior to the patient’s current hospitalization shows sinus bradycardia (45 beats per minute); T-wave inversions in leads V4-V6, I, II, III, and AVF; and Q waves in leads V1-V3. The QTc interval was within the normal range (414 msec). (The QTc was calculated using Bazett’s formula.)
The patient remained asymptomatic throughout her entire hospital course. She was offered a cocaine detoxification program, but she refused it. A beta-adrenergic blocking drug was not initiated because concomitant use with cocaine would put the patient at risk of serious adverse events, including significant bradycardia. She was advised to undergo stent placement for stenosis of the left carotid artery, but this treatment route also was refused.

References
The Impersonator

Hartmut Lode, MD,a Jutta Wagner, MD,a Hans Scherübl, MDa,b

aCharité-Universitätsmedizin Berlin, Berlin, Germany bVivantes Klinikum Am Urban, Berlin, Germany

PRESENTATION

Sometimes you find things when and where you least expect them. A 29-year-old man born in Yugoslavia presented with a 2-month history of bilateral cervical swelling and sinus discharge. He had no fever, night sweats, or weight loss. A tuberculin skin test was positive. Clinical examination and computed tomography revealed cervical and upper thoracic lymphadenitis with abscess formation on the left side (Figure 1). The abscess was drained, and an antituberculous drug regimen comprised of ethambutol, isoniazid, pyrazinamide, and rifampin was initiated. During the next 2 months, the swelling of the lymph nodes decreased, and the abscess gradually healed. Nonresistant Mycobacterium tuberculosis was ultimately cultured, and HIV infection was excluded. Nine weeks after the start of antituberculous therapy, the patient presented again with fever and a tender swelling in the upper right quadrant of the abdomen.

ASSESSMENT

Computed tomography revealed an 8-cm abscess in the hepatic portal region (Figures 2 and 3). Pus containing acid-fast bacilli was withdrawn from the abscess by needle aspiration. At the time of the initial diagnosis, computed tomography had demonstrated only a 3-cm matted lymphadenopathy adjacent to the inferior vena cava, surrounding the pancreas, and reaching towards the portal region (Figure 4). Thus, this latest tuberculous lesion formed even though chemotherapy achieved regression of the patient’s cervical lymphadenitis.

Tuberculosis in the portal region is challenging to diagnose and to manage. Isolated tuberculous abscesses or pseudotumors in the pancreatic head, portal region, liver, or gall bladder are well-known to mimic pancreatic cancer or carcinoma of the biliary ducts, gall bladder, or liver.1-4 Differentiating between neoplasm and tuberculosis in the portal region is a rare clinical problem in Western countries, where cancer is the most common cause, and in developing countries, where infection is the most frequent culprit. However, in recently developed countries, one etiology is as likely as the other, and the differential diagnosis can become very difficult.

Tuberculous biliary strictures can produce obstructive jaundice suggestive of cholangiocarcinoma or pancreatic cancer.1,5-7 For example, 1 patient who presented with obstructive jaundice produced by a tuberculous abscess in the pancreatic head initially appeared to have pancreatic cancer; he was managed with pancreaticoduodenectomy and antituberculous therapy.1 Vascular involvement secondary to abdominal tuberculosis may also occur. Tuberculous lymphadenopathy in the portal region can lead to encasement and compression of the portal vein resulting in prestenotic dilatation and variceal formation in the upper abdomen.8 In rare cases, portal hypertension due to abdominal tuberculosis may even cause upper gastrointestinal bleeding.9,10

Requests for reprints should be addressed to Hans Scherübl, MD, Klinik für Innere Medizin, Vivantes Klinikum Am Urban, Dieffenbachstr. 1, 10967 Berlin, Germany.
E-mail address: hans.scheruebl@vivantes.de

Figure 1 Cervical and upper thoracic lymphadenitis with abscess formation, as designated by arrows, was evident on computed tomography.
Paradoxical expansion of tuberculomas or tuberculous adenopathy despite effective antituberculous chemotherapy is widely recognized. Interestingly, this development generally does not cause clinical complications. However, as our patient’s case report shows, this statement does not always hold true for tuberculosis in the portal region.\textsuperscript{11}

**MANAGEMENT**

Our patient’s antituberculous therapy was stepped up again with ethambutol, isoniazid, pyrazinamide, and rifampin, and the patient became afebrile within 6 days. He was then discharged and was well enough to return to work. Two months later all hepatotoxic medication had to be stopped for several weeks because of drug-induced hepatitis. Once the hepatitis subsided, therapy was resumed with isoniazid and sparfloxacin and continued for 9 more months. Abscesses were then no longer detectable, but imaging showed stenosis and cavernous transformation of the portal vein (Figure 5). The patient is well and has not developed endoluminally-visible esophageal, gastric, or duodenal varices in the 8 years since.

**References**


CLINICAL RESEARCH STUDY

Increasing Trends in Incidence of Overweight and Obesity over 5 Decades

Nisha I. Parikh, MD, MPH,a Michael J. Pencina, PhD,a,e Thomas J. Wang, MD,a,b Katherine J. Lanier, BS,e Caroline S. Fox, MD,a,c,d Ralph B. D'Agostino, PhD,a,e Ramachandran S. Vasan, MDa,f

aFramingham Heart Study, Framingham, Mass; bDivision of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Mass; cDepartment of Endocrinology, Diabetes, and Hypertension, Brigham and Women’s Hospital, Harvard Medical School, Boston, Mass; National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Md; eDepartment of Mathematics and Statistics, Boston University, Boston, Mass; fPreventive Medicine and Cardiology Sections, Boston Medical Center, Boston University School of Medicine, Boston, Mass.

ABSTRACT

PURPOSE: We evaluated trends in the incidence of overweight and obesity over the past 50 years.

METHODS: We evaluated trends in the incidence of overweight (BMI ≥ 25 kg/m²), obesity (BMI ≥ 30 kg/m²) and stage 2 obesity (BMI ≥ 35 kg/m²) from 1950 to 2000 in Framingham Study participants (n = 6798, 54% women). Individuals aged 40-55 years who attended 2 examinations 8 years apart in each decade were eligible.

RESULTS: The incidences of overweight, obesity, and stage 2 obesity increased across the decades in both sexes (P for trend < .001). For men, the incidence of overweight rose from 21.8% (95% confidence interval [CI], 17.6-26.5) in the 1950s to 35.2% (95% CI, 28.6-42.5) in the 1990s; of obesity from 5.8% (95% CI, 4.4-7.6) to 14.8% (95% CI, 12.2-17.9); and of stage 2 obesity from 0.2% (95% CI, 0.1-0.9) to 5.4% (95% CI, 4.0-7.2). For women, incidence rates of overweight increased from 15.0% (95% CI, 12.3-18.1) to 33.1% (95% CI, 29.0-37.4); of obesity from 3.9% (95% CI, 2.9-5.3) to 14% (95% CI, 11.6-16.7); and of stage 2 obesity from 1.7% (95% CI, 1.1-2.6) to 4.4% (95% CI, 3.2-6.0). Overall, incidence rates of overweight increased 2-fold and that of obesity more than 3-fold over 5 decades, findings that remained robust upon additional adjustment for baseline BMI in each decade.

CONCLUSIONS: The incidence of overweight and obesity increased progressively over the last 5 decades, suggesting that the rising trend in prevalence is not a recent phenomenon. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Body mass index; Obesity; Overweight; Trends; Epidemiology

The epidemic of obesity in the United States is a major public health problem. Excess adiposity increases risk of diabetes, hypertension, cardiovascular diseases, and certain types of cancers. Furthermore, obesity is associated with elevated mortality risks due to cardiovascular disease and all-causes.

The increasing prevalence of overweight and obesity underscores the need to better understand this epidemic. Data from the National Health Examination Survey (NHES I, 1960-1962) and National Health and Nutritional Examination Surveys (NHANES I, 1971-1974; NHANES II, 1976-1980; NHANES III, 1988-1994; NHANES 1999-2000) indicate that the prevalence of obesity (body mass...
index [BMI $\geq 30$ kg/m$^2$] in the United States was relatively stable from 1960-1980, at which point rates escalated.\textsuperscript{10} In comparison, studies have suggested a stable prevalence of overweight (25 $\leq$ BMI $< 30$ kg/m$^2$) until the 1990s (with the exception of women between the ages of 20 and 29 years in whom overweight prevalence increased before the 1990s).\textsuperscript{11} In this context, it is important to note that parallel data on temporal trends in the incidence of overweight and obesity in the community are lacking. Such incidence data are critical for understanding reasons underlying rising prevalence trends and for making projections about future burden of obesity. Additionally, an analysis of incidence trends may help clarify the apparent paradox of rising obesity prevalence despite stable overweight prevalence in the 1980s.

We examined temporal trends in the incidence of overweight and obesity among Framingham Study participants over a 50-year period from 1950 to 2000. We hypothesized that the current epidemic of obesity was paralleled by an epidemic of overweight, which may have been unapparent in the absence of data on incidence trends. We also postulated that the rising trend of excess adiposity is not a recent phenomenon.

**METHODS**

**Sample**

The Framingham Heart Study, a community-based prospective cohort study, began in 1948, with enrollment of 5209 participants (original cohort).\textsuperscript{12,13} In 1971, 5124 individuals who were children of the original cohort (and their spouses) were enrolled into the Framingham Offspring Study. Participants in the original cohort are examined biennially, whereas the offspring cohort is evaluated quadriennially.\textsuperscript{14,15}

For the present investigation, we chose 2 examinations within each calendar decade from 1950 up to 2000 that were approximately 8 years apart (Figure). The availability of only 2 offspring cohort examinations in the 1970s that were 8 years apart constrained us to using observations from only 2 examinations within each decade. Individuals were eligible if they were between ages 40 and 55 years at the first of the 2 examinations in a given decade; attended a follow-up examination 8 years later; and were not underweight (BMI $< 18.5$ kg/m$^2$) at the baseline examination. For primary analyses we excluded underweight individuals because such persons may have had illnesses that prevent weight gain. We chose the 40-55 years group because adequate numbers of individuals for this age range were available in each of the 5 decades. All participants gave informed consent and the study protocol was approved by the Institutional Review Board of the Boston Medical Center.

**Body Mass Index and Risk Factors**

**Measurement**

At each Framingham study examination, height and weight are measured using standardized protocols.\textsuperscript{16,17} BMI was calculated as the weight in kilograms divided by the square of height in meters.

Participants underwent assessment of vascular risk factors at each examination.\textsuperscript{17} Current smoking was defined as regular cigarette smoking in the year preceding the examination. Smoking cessation was defined as a history of quitting smoking subsequent to the baseline examination. A physical activity index\textsuperscript{18,19} was calculated based on responses to a physical activity questionnaire for offspring cohort attendees at examinations 2 (1980s) and 4 (1990s). Dietary daily caloric intake was calculated based on 3-day dietary recall information obtained for the offspring cohort in the 1980s and the 1990s.\textsuperscript{20}

**BMI on Follow-Up: Definitions of Overweight and Obesity**

All eligible participants within a given decade were followed for 8 years to assess the development of the BMI outcomes (see below). Categories of BMI were defined according to established guidelines\textsuperscript{21,22}: normal weight (18.5 to $< 25$ kg/m$^2$), overweight (25 $\leq$ BMI $< 30$ kg/m$^2$), overweight or more ($\geq 25$ kg/m$^2$), obesity ($\geq 30$ kg/m$^2$), and stage 2 obesity ($\geq 35$ kg/m$^2$).

**Statistical Methods**

For each decade from 1950 to 2000, we evaluated the prevalence of BMI categories at the baseline examination in that decade, and assessed trends in prevalence (with generalized estimating equations to account for individuals contributing to more than one decade, 1950s serving as referent). Next, we evaluated the sex-specific 8-year incidence of the following BMI outcomes (eligibility defined by BMI at first examination in each decade):

a. *Overweight* defined as $25 \leq$ BMI $< 30$ kg/m$^2$; participants with BMI $< 25$ kg/m$^2$ were eligible for these analyses. Individuals with BMI $> 30$ kg/m$^2$ at follow-up examination were excluded for estimating incidence of overweight alone.
Overweight or more defined as a BMI $\geq 25$ kg/m$^2$; participants with BMI $<25$ kg/m$^2$ were eligible for these analyses.

Obesity defined as a BMI $\geq 30$ kg/m$^2$; participants with BMI $<30$ kg/m$^2$ were eligible for these analyses.

Stage 2 obesity defined as a BMI $\geq 35$ kg/m$^2$; participants with BMI $<35$ kg/m$^2$ were eligible for these analyses.

We evaluated trends in incidence of BMI outcomes using sex-specific multivariable pooled logistic regression adjusting for age and smoking cessation, the latter 2 being major confounders. In additional analyses, we adjusted for baseline BMI. The 1950s served as referent. Because individuals who developed an outcome of interest could not contribute to future decades, generalized estimating equations were not necessary for these analyses. We repeated analyses to examine trends in incidence of obesity among nonsmokers; trends in overweight among nonsmokers were not analyzed because there were too few people eligible in the referent decade. We analyzed nonsmokers to determine if our results were influenced by the striking decreases in smoking rates across decades in our cohorts. Likewise, we did not analyze trends in incidence of stage 2 obesity among nonsmokers because there were no events (development of BMI $\geq 35$ kg/m$^2$) observed in the 1950s in nonsmokers. We repeated analyses including individuals who were underweight to see if this altered our incidence estimates.

Finally, we studied temporal trends in decade-specific incidence rates of weight loss by examining:

- Overweight individuals who lost weight: proportion of individuals with BMI $\geq 25$ kg/m$^2$ during a decade who subsequently became normal weight (18.5 kg/m$^2$ < BMI < 25 kg/m$^2$).
- Obese individuals who lost weight: the proportion of individuals with BMI $\geq 30$ kg/m$^2$ during the decade who...
subsequently became normal weight or overweight (18.5 kg/m² < BMI < 25 kg/m²).

A 2-tailed P value < .05 was considered statistically significant.

RESULTS

Baseline Characteristics

In our sample, the prevalence of obesity, stage 2 obesity, mean values of height, and weight increased across the decades from the 1950s to 2000 in both sexes (Table 1; P for trend <.001). Over this period, mean BMI increased by about 2.7 kg/m² in men and 1.5 kg/m² in women. Smoking rates decreased in men more substantially than in women (P for trend <.001).

Table 1 Baseline Characteristics of All Participants and Prevalence Rates for Body Mass Index (BMI) Categories in Each Decade

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>921</td>
<td>799</td>
<td>1103</td>
<td>804</td>
<td>822</td>
<td>431</td>
<td>1217</td>
<td>970</td>
<td>1183</td>
<td>839</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight or more (BMI ≥ 25 kg/m²), %*</td>
<td>64.0</td>
<td>66.3</td>
<td>75.1</td>
<td>74.9</td>
<td>77.9</td>
<td>81.4</td>
<td>51.0</td>
<td>40.3</td>
<td>44.5</td>
<td>42.0</td>
</tr>
<tr>
<td>Obesity (BMI ≥ 30 kg/m²), %*</td>
<td>10.2</td>
<td>13.4</td>
<td>19.6</td>
<td>20.8</td>
<td>25.4</td>
<td>32.3</td>
<td>13.9</td>
<td>11.4</td>
<td>13.8</td>
<td>14.2</td>
</tr>
<tr>
<td>Stage 2 obesity (BMI ≥ 35 kg/m²), %*</td>
<td>0.5</td>
<td>1.6</td>
<td>3.0</td>
<td>2.4</td>
<td>5.6</td>
<td>10.0</td>
<td>3.9</td>
<td>2.9</td>
<td>4.1</td>
<td>5.6</td>
</tr>
<tr>
<td>Weight, kg*</td>
<td>77.5</td>
<td>79.6</td>
<td>82.0</td>
<td>83.8</td>
<td>87.0</td>
<td>90.4</td>
<td>64.9</td>
<td>63.7</td>
<td>64.9</td>
<td>65.7</td>
</tr>
<tr>
<td>Height, m*</td>
<td>1.72</td>
<td>1.73</td>
<td>1.73</td>
<td>1.75</td>
<td>1.76</td>
<td>1.77</td>
<td>1.59</td>
<td>1.60</td>
<td>1.59</td>
<td>1.61</td>
</tr>
<tr>
<td>Smoking, %*</td>
<td>76.3</td>
<td>66.7</td>
<td>43.6</td>
<td>44.3</td>
<td>26.5</td>
<td>19.5</td>
<td>37.7</td>
<td>49.2</td>
<td>40.5</td>
<td>32.6</td>
</tr>
</tbody>
</table>

Values are means, unless indicated otherwise.

*P-values <.01 for trend across decades in both sexes (with the exception of BMI ≥ 25 kg/m² in women); P-values are age-adjusted and based on generalized estimating equations to account for participants contributing to more than one decade.

Table 2 Baseline Characteristics of Persons Aged 40-55 Years Eligible in Each Calendar Decade

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>827</td>
<td>692</td>
<td>887</td>
<td>637</td>
<td>613</td>
<td>1048</td>
<td>859</td>
<td>1020</td>
<td>720</td>
<td>731</td>
<td>332</td>
</tr>
<tr>
<td>Persons with BMI &lt; 30 kg/m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight, kg*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, m*</td>
<td>1.72</td>
<td>1.73</td>
<td>1.73</td>
<td>1.75</td>
<td>1.76</td>
<td>1.77</td>
<td>1.59</td>
<td>1.60</td>
<td>1.60</td>
<td>1.61</td>
<td>1.63</td>
</tr>
<tr>
<td>Smoking, %*</td>
<td>77.5</td>
<td>67.8</td>
<td>44.1</td>
<td>43.9</td>
<td>26.3</td>
<td>39.9</td>
<td>51.1</td>
<td>41.0</td>
<td>33.7</td>
<td>24.8</td>
<td></td>
</tr>
</tbody>
</table>

Values are means, unless indicated otherwise.

*P-values <.01 for trend across decades in both sexes; P-values are age-adjusted and based on generalized estimating equations to account for participants contributing to more than 1 decade.
1990s (from 36.0 to 37.5 in men, and from 33.9 to 36.8 in women, \( P < .001 \) for both) paralleled by increases in reported daily caloric intake (2206 to 2363 kilocalories in men and 1549 to 1676 kilocalories in women, \( P < .001 \) for both).

### Incidence of BMI Outcomes

The incidences of overweight, overweight or more, obesity and stage 2 obesity increased from 1950 to 1990 in both sexes (Table 3). In women, the incidence rates of BMI outcomes rose across the 5 decades in a step-wise fashion. In men, incidence rate increases across decades were less graded; rates rose in the 1960s (relative to the 1950s) but then decreased in the 1970s, only to escalate again in the 1980s.

Table 4 displays adjusted trends in incidence of BMI outcomes with rates in the 1950s as referent. Overall, incidence rates of overweight increased over 2-fold and that of obesity more than 3-fold over the 5-decade period in both sexes. In models fitting a linear trend across decades, in women there was a 25% (95% confidence interval [CI], 17%-34%) increase in incidence of overweight per decade, a 34% (95% CI, 24%-46%) increase in incidence of obesity per decade, and a 31% (95% CI, 16%-49%) increase in incidence of stage 2 obesity per decade. In men, there was a 20% (95% CI, 10%-32%) increase in incidence of overweight per decade, a 29% (95% CI, 18%-40%) increase in incidence of obesity per decade, and a 97% (95% CI, 63%-138%) increase in incidence of stage 2 obesity per decade. These results remained robust upon additional adjustment for baseline BMI (\( P = .004 \) for trend across decades for overweight or more incidence in men; \( P < .001 \) for all other BMI outcomes in both sexes). These trends were consistent when we limited our analysis to nonsmokers (Appendix 1).

### Cohort-Specific Incidence Rates in the 1970s

In order to investigate if differences in incidence rates across decades were influenced by cohort effects (original cohort contributing data for the 1950s and 1960s, and off-spring cohort for 1980s and 1990s), we assessed incidence rates separately for eligible participants in the 2 cohorts in the 1970s (the decade at which individuals between ages 40 and 55 years were available in both cohorts; Appendix 2). The incidence rates of BMI outcomes were similar in the 2 cohorts; whereas point estimates in rates varied for the 2 samples, the 95% CI overlapped. Analyses that included underweight individuals did not materially change incidence rates for overweight, obesity and stage 2 obesity (Appendixes 3, 4).

### Temporal Trends in Weight Loss

The proportions of overweight or more individuals who achieved normal weight status and of obese individuals who became nonobese have steadily decreased over the past 50 years (Tables 5, 6).
### Table 4  Trends in Incidence of BMI Outcomes Across Decades (1950s to 1990s): Results of Age- and Smoking Cessation-Adjusted Models

<table>
<thead>
<tr>
<th>Odds Ratio (95% CI)</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence of overweight (25 ≤ BMI &lt; 30 kg/m²)</td>
<td>Referent</td>
<td>1.53* (1.04-2.26)</td>
</tr>
<tr>
<td>Incidence of overweight or more (BMI ≥ 30 kg/m²)</td>
<td>Referent</td>
<td>1.53* (1.04-2.24)</td>
</tr>
<tr>
<td>Incidence of obesity (BMI ≥ 30 kg/m²)</td>
<td>Referent</td>
<td>1.67* (1.11-2.51)</td>
</tr>
<tr>
<td>Incidence of stage 2 obesity (BMI ≥ 35 kg/m²)</td>
<td>Referent</td>
<td>5.60* (1.22-25.66)§</td>
</tr>
</tbody>
</table>

*P < .05.
†P < .01.
‡P < .001.
§Wide 95% CI are the result of the small n in the referent group (2 events in 916 men at risk in 1950s, Table 3).

### Table 5  Incidence Rates of Losing Weight

<table>
<thead>
<tr>
<th>BMI Outcome</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events/no. at risk</td>
<td>58/589</td>
<td>31/530</td>
</tr>
<tr>
<td>Incidence rate (95% CI)</td>
<td>9.8 (7.7-12.5)</td>
<td>5.8 (4.1-8.2)</td>
</tr>
<tr>
<td>BMI &gt; 30 on follow-up†</td>
<td>32/94</td>
<td>20/107</td>
</tr>
<tr>
<td>Events/no. at risk</td>
<td>34.0 (25.2-44.2)</td>
<td>18.7 (12.4-27.2)</td>
</tr>
</tbody>
</table>

*Refers to individuals overweight or more (BMI ≥ 25 kg/m²) during the referent decade who subsequently became normal weight (18.5 kg/m² < BMI < 25 kg/m²).
†Refers to those individuals obese or more (BMI ≥ 30 kg/m²) during the referent decade who subsequently became normal weight or overweight (18.5 kg/m² < BMI < 30 kg/m²).
DISCUSSION

It is estimated that the lifetime risk of being overweight exceeds 70% and that for obesity it exceeds 35%.²³ In the present investigation we assessed if the rising prevalence of overweight and obesity in national cross-sectional surveys was a recent trend or a gradual phenomenon accruing over decades. Our principal findings are 3-fold. First, incidence rates of overweight and obesity increased 2- to more than 3-fold over the last 5 decades in our community-based sample. Second, the incidence rates increased across decades in a fairly monotonic fashion in women. However, incidence rates in men demonstrated a biphasic pattern of increase, with an initial increase in the 1960s and a subsequent one in the 1980s. The reasons for these differing patterns in the 2 sexes are not clear. Third, we noted a striking increase in the prevalence of stage 2 obesity in the 1990s and 2000s, and observed a rising incidence in the 1990s, a matter of great concern.

The reasons for the increases in obesity over the past several decades are likely manifold. Our data demonstrate a decrease in incidence rates of weight loss among both overweight and obese individuals over the past 50 years. Smoking cessation among adults has been related to an increased prevalence of overweight.²⁴ Indeed, smoking rates decreased considerably in our study cohort over the past 50 years. However, adjusting for smoking cessation did not significantly alter the incidence trends in our investigation. Increased mean energy intake among adults between 1970 and 2000 has also been implicated as a potential cause for the growing obesity epidemic.²⁵,²⁶ Such increases in energy intake have been attributed to increases in mean food and beverage portion sizes,²⁷ and consumption of energy-dense fast foods.²⁸ Consistent with these data, the self-reported mean energy intake in our cohort increased between the 1980s and the 1990s.

Decreased physical activity in the United States population has been suggested as a causal factor for greater obesity prevalence. Data from the Behavioral Risk Factor Surveillance System (BRFSS), as well as from our cohort indicate that self-reported physical activity has actually increased moderately from the 1980s into the 2000s.²⁹,³⁰ One explanation for the rising incidence in obesity in the face of higher physical activity may be that increases in calorific intake may have been greater relative to increases in physical activity. An alternative explanation is that the physical activity questionnaire in BRFSS and our study may not have captured adequately overall increases in sedentary lifestyle among adults. Factors contributing to a sedentary lifestyle include less physically demanding occupations,³¹ changing land-use patterns (eg, urban sprawl³²), and increased automotive travel.³³

Temporal Trends in BMI Prevalence and Incidence: Comparison with Published Literature

The baseline prevalence of overweight (or more), obesity, and stage 2 obesity in our sample in each decade closely resembles that for corresponding age groups in the
However, our estimates of prevalence of overweight (or more) and obesity were higher in the 1990s and 2000s than those reported for comparable age groups in BRFSS. Lower prevalence estimates in BRFSS may be due to self-reported height and weight. Previous studies of self-reported weight and height report that overweight participants tend to underestimate their true weight, whereas most participants overestimate their height. Also, inaccurate height reporting increases after age 45 years.

Previous longitudinal studies of BMI and weight categories have not focused on temporal trends in incidence of overweight and obesity during the past 5 decades; therefore, we are unable to compare our results with other studies. Our incidence data complement available information on the prevalence of overweight and obesity from national surveys. The substantial incidence rates of overweight, obesity, and stage 2 obesity in our study population in the 1990s are consistent with a parallel increase in the prevalence of obesity nationwide. Overall, our longitudinal observations over 50 years suggest that the increase in prevalence of overweight and obesity was accompanied by increasing incidence rates for both.

**Strengths and Limitations**

The strengths of our investigation include the use of prospectively collected data over a 50-year period in a community-based study, and the standardized measurements of BMI over this period. Nonetheless, several limitations should be acknowledged. We evaluated only participants aged 40-55 years (an unavoidable constraint of studying BMI trends for participants in a similar age range over 5 decades) that limits the generalizability of our results to persons older or younger, and limits the comparison of our findings with observations in the NHANES samples. Our sample consisted of 2 separate cohorts—original cohort participants contributed observations in the earlier decades (1950s and 1960s), whereas offspring cohort members provided information during the later decades (1980s and 1990s). Birth cohort effects have been described with later birth cohorts demonstrating a greater propensity for obesity.

We found no clear-cut evidence of cohort-related differences in incidence of overweight and obesity. Our study cohort is overwhelmingly white, and our results may not be generalizable to nonwhites. Physical activity and dietary caloric intake measurements were available only for recent decades; consequently, we were unable to assess the contributions of these factors to the rising incidences of obesity. We did not account for temporal changes in the incidence of chronic diseases (like cancer) that may influence BMI trends. Lastly, we were unable to examine trends in the incidence of central adiposity because waist measurements were not obtained at early examinations. NHANES data suggest that the prevalence of abdominal adiposity has increased from 1960-1962 to 1999-2000.

**Implications**

To our knowledge, the present investigation is the first systematic analysis of the incidence of overweight, obesity, and stage 2 obesity over 5 decades in a community-based sample in the United States. Increased incidences of overweight and obesity over the last 50 years suggests that the rising trend of excess adiposity is not a recent phenomenon, although rates have escalated in the last decade. If these patterns of rising incidence continue unabated, the community burden of overweight and obesity may continue to increase over the next decade. Data provided by our study can help in the estimation of the future burden of excess adiposity combining the trajectory of incidence rates over 50 years (average increase per decade of 20%-30% for overweight and 30%-35% for obesity) with the incidence rates observed in the last decade.

**References**


### Appendix 1: Incidence Rates of Obesity (BMI ≥ 30 kg/m²) in Nonsmokers

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events/no. at risk</td>
<td>10/164 19/197 20/409 34/345 53/441</td>
<td>28/588 45/407 47/571 34/467 73/540</td>
</tr>
<tr>
<td>Incidence rate (95% CI)</td>
<td>6.1 (3.3-11.0) 4.9 (3.2-7.5) 3.8 (2.2-6.1) 2.9 (2.1-3.9) 2.1 (1.0-4.2)</td>
<td>4.8 (3.3-6.8) 11.1 (8.4-14.5) 8.2 (6.2-10.8) 8.9 (6.0-12.7) 7.3 (5.0-10.0)</td>
</tr>
<tr>
<td>Odds ratio (95% CI) Referent</td>
<td>1.65 (0.75-3.66) 0.81 (0.37-1.76) 1.69 (0.81-3.51) 2.11* (1.05-4.26)</td>
<td>2.50† (1.53-4.09) 1.80* (1.11-2.93) 1.56 (0.93-2.62) 3.12‡ (1.99-4.91)</td>
</tr>
</tbody>
</table>

†P < .01.  
*P < .05.  
‡P < .001.

### Appendix 2: Cohort-Specific 8-year Crude Incidence Rates* of Overweight, Obesity and Stage 2 Obesity in the 1970s

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incidence rate of overweight (95% CI) 17.8 (11.6-26.2) 22.3 (16.6-29.3)</td>
<td>23.9 (19.0-29.6) 18.2 (14.7-22.2)</td>
</tr>
<tr>
<td></td>
<td>Incidence rate of overweight or more (95% CI) 17.8 (11.6-26.2) 23.2 (17.4-30.2)</td>
<td>24.5 (19.6-30.2) 18.4 (14.9-22.4)</td>
</tr>
<tr>
<td></td>
<td>Incidence rate of obesity (95% CI) 7.4 (5.0-10.8) 6.6 (4.8-8.9)</td>
<td>6.4 (4.4-9.3) 7.3 (5.5-9.7)</td>
</tr>
<tr>
<td></td>
<td>Incidence rate of stage 2 obesity (95% CI) 1.3 (0.5-3.0) 1.0 (0.5-2.1)</td>
<td>2.7 (1.5-4.7) 2.8 (1.8-4.3)</td>
</tr>
</tbody>
</table>

*Incidence rate per 100 persons.
### Appendix 3: Incidence Rates of Overweight, Obesity, and Stage 2 Obesity (for Sample that Included Individuals with BMI <18.5 mg/kg²)

<table>
<thead>
<tr>
<th>BMI Outcome</th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overweight or more</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events/no. at risk</td>
<td>73/342</td>
<td>82/277</td>
<td>58/282</td>
<td>76/208</td>
<td>66/184</td>
<td>90/617</td>
<td>129/599</td>
<td>136/670</td>
<td>129/500</td>
<td>170/500</td>
</tr>
<tr>
<td>Incidence rate (95% CI)</td>
<td>21.3 (17.3-26.0)</td>
<td>29.6 (24.5-35.2)</td>
<td>20.6 (16.2-25.7)</td>
<td>37.5 (31.2-44.3)</td>
<td>35.9 (29.3-43.1)</td>
<td>14.6 (12.0-17.6)</td>
<td>21.5 (18.4-25.0)</td>
<td>20.3 (17.4-23.5)</td>
<td>25.8 (22.2-29.8)</td>
<td>34.0 (30.0-38.3)</td>
</tr>
<tr>
<td>Obesity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events/no. at risk</td>
<td>48/837</td>
<td>67/700</td>
<td>61/894</td>
<td>75/643</td>
<td>91/615</td>
<td>41/1069</td>
<td>60/879</td>
<td>71/1034</td>
<td>58/733</td>
<td>102/745</td>
</tr>
<tr>
<td>Incidence rate (95% CI)</td>
<td>5.7 (4.3-7.5)</td>
<td>9.6 (7.6-12.0)</td>
<td>6.8 (5.3-8.7)</td>
<td>11.7 (9.4-14.4)</td>
<td>14.8 (12.2-17.8)</td>
<td>3.8 (2.8-5.2)</td>
<td>6.8 (5.3-8.7)</td>
<td>6.9 (5.5-8.6)</td>
<td>7.9 (6.2-10.1)</td>
<td>13.7 (11.4-16.4)</td>
</tr>
<tr>
<td>Stage 2 obesity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events/no. at risk</td>
<td>2/926</td>
<td>11/794</td>
<td>12/1077</td>
<td>26/791</td>
<td>42/778</td>
<td>20/1191</td>
<td>15/962</td>
<td>31/1148</td>
<td>27/805</td>
<td>37/857</td>
</tr>
<tr>
<td>Incidence rate (95% CI)</td>
<td>0.2 (0.1-0.9)</td>
<td>1.4 (0.8-2.5)</td>
<td>1.1 (0.6-2.0)</td>
<td>3.3 (2.2-4.8)</td>
<td>5.4 (4.0-7.2)</td>
<td>1.7 (1.1-2.6)</td>
<td>1.6 (0.9-2.6)</td>
<td>2.7 (1.9-3.8)</td>
<td>3.4 (2.3-4.8)</td>
<td>4.3 (3.1-5.8)</td>
</tr>
</tbody>
</table>

### Appendix 4: Trend across Decades (1950s to 1990s): Age and Smoking Cessation Adjusted Incidence of Adiposity Outcomes (for Sample that Included Individuals with BMI <18.5 mg/kg²)

<table>
<thead>
<tr>
<th>Odds Ratio (95% CI)</th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence of overweight Referent</td>
<td>1.53* (1.04-2.25)</td>
<td>1.03 (0.69-1.55)</td>
<td>2.18* (1.47-3.23)</td>
<td>2.26* (1.50-3.41)</td>
<td>&lt;.001</td>
<td>Referent</td>
<td>1.53* (1.14-2.07)</td>
<td>1.42* (1.05-1.90)</td>
<td>1.95* (1.44-2.64)</td>
<td>2.93* (2.19-3.92)</td>
</tr>
<tr>
<td>Incidence of obesity Referent</td>
<td>1.68* (1.12-2.52)</td>
<td>1.24 (0.82-1.86)</td>
<td>2.12* (1.43-3.15)</td>
<td>3.05* (2.09-4.47)</td>
<td>&lt;.001</td>
<td>Referent</td>
<td>1.80* (1.19-2.70)</td>
<td>1.81* (1.22-2.69)</td>
<td>2.10* (1.39-3.17)</td>
<td>3.90* (2.68-5.69)</td>
</tr>
<tr>
<td>Incidence of stage 2 obesity Referent</td>
<td>5.61* (1.22-25.72)</td>
<td>4.54* (1.00-20.58)</td>
<td>14.00* (3.31-59.24)</td>
<td>24.85* (5.99-103.05)</td>
<td>&lt;.001</td>
<td>Referent</td>
<td>0.93 (0.47-1.82)</td>
<td>1.61 (0.91-2.86)</td>
<td>2.04* (1.14-3.67)</td>
<td>2.64* (1.52-4.59)</td>
</tr>
</tbody>
</table>

*P < .05.
†P < .01.
‡P < .001
Gaps in Treatment Among Users of Osteoporosis Medications: The Dynamics of Noncompliance

M. Alan Brookhart, PhD, Jerry Avorn, MD, Jeffrey N. Katz, MD, MS, Joel S. Finkelstein, MD, Marilyn Arnold, ScD, Jennifer M. Polinski, MPH, Amanda R. Patrick, MS, Helen Mogun, MS, Daniel H. Solmon, MD, MPH

Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, Mass.

ABSTRACT

PURPOSE: Cyclical patterns of compliance have been observed with many health-related activities such as dieting and exercise. It is not known whether such patterns of compliance exist among users of chronic medications. We sought to estimate the percentage of patients who restart osteoporosis therapy after a prolonged lapse in medication use and to identify the factors associated with a return to compliance.

METHODS: We studied 26,636 new users of an osteoporosis medication (alendronate, calcitonin, estrogen, raloxifene, or risedronate) who were age 65 or older and had an extended lapse in refill compliance, defined as a period of at least 60 days after the completion of one prescription in which no refill for any osteoporosis medication was obtained. Survival curves were used to estimate the length of time until therapy is resumed. We estimated the association between patient characteristics and the rate of resuming treatment using Cox proportional hazards analysis. We then conducted a case crossover analysis to examine whether certain events occurring during follow-up triggered a return to refill compliance.

RESULTS: Of patients who stopped therapy for at least 60 days, an estimated 30% restarted treatment within 6 months, and 50% restarted within 2 years. Among patients who had at least 6 months of continuous use before their interruption in treatment (n = 5863), 42% restarted therapy within 6 months and 59% within 2 years. Younger patients, women, and those with a history of a fracture were more likely to return after a break in medication use. Recent hip fractures, discharges from nursing homes, and bone mineral density testing also predicted a return to treatment.

CONCLUSION: Extended gaps in treatment are common among users of osteoporosis medications. Because the effectiveness of these drugs used in an interrupted way is unknown, compliance interventions should emphasize the need for continuous medication use. Further research is needed to understand why patients often go for months without refilling prescriptions and also whether similar utilization patterns exist for other chronic medications. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Osteoporosis; Medication compliance; Adherence; Persistence; Drug holiday
step in the design of an effective intervention to improve use of prescription medications.

In many studies of long-term medication use, non-compliance is either explicitly or implicitly treated as an endpoint. These analyses commonly report the percentage of patients still compliant with a therapeutic regimen at various time points or conduct analyses examining associations between patient characteristics and the probability of stopping therapy. While these analytic approaches have led to important characterizations of the problem of non-compliance, they are not designed to reveal dynamic patterns of medication use, such as patients stopping and then restarting therapy.

Cyclical patterns of compliance have been observed with other health-related activities; well-known examples include dieting and exercise. In controlled settings in which medication use is electronically monitored, compliance has been found to be erratic. In such contexts, the term “drug holiday” has been used to describe an interruption in medication use lasting more than 3 days. However, relatively little is known about the extent to which patients in routine care stop and restart drug therapies for chronic diseases.

In a previous study of compliance among new users of osteoporosis medications, we found that 1 year after initiation of therapy, 45% of patients had a period of \( \geq 120 \) days in which their prescription was not refilled. The present study was conducted within this same cohort to understand whether or not patients who discontinue osteoporosis medication for an extended period ultimately resume medication use.

**CLINICAL SIGNIFICANCE**

- Long unexplained interruptions in treatment are common among users of osteoporosis medications.
- Fractures and bone mineral density testing are predictors of a return to medication use among patients who have stopped using medication.
- The effectiveness of osteoporosis drugs used in an interrupted way is unknown.
- Clinicians should stress the need for continuous medication use to patients being treated for osteoporosis.

**METHODS**

**Data**

We studied Medicare beneficiaries aged \( \geq 65 \) years who were concurrently enrolled in the Pharmaceutical Assistance Contract for the Elderly (PACE) of Pennsylvania. To be eligible for PACE, a participant’s annual income must be \( \leq \$17,700 \). The program reimburses the cost of all prescription medications with a co-payment of \( \leq \$9 \). There are no prescribing restrictions on medications used for osteoporosis. Medication information from PACE includes the drug name, dosage, number of pills dispensed, and days supplied. The study database also includes Medicare information on all inpatient and outpatient encounters, including diagnoses, procedures, and tests ordered.

Our study was conducted within an existing retrospective cohort of new users of osteoporosis medication who initiated treatment between 1996 and 2002. The medications that we studied were bisphosphonates, calcitonin, estrogen therapy (except vaginal creams), and raloxifene. Teriparatide was not included because it only became available in the last year of the study. Vitamin D preparations were also excluded because they are often used for other bone diseases. To reduce the likelihood that patients were obtaining medicines through other programs, we required patients to have filled at least one prescription through PACE in each of the two 6-month intervals preceding the filling of their initial osteoporosis prescription. Patients were followed until they became ineligible for PACE, died, or reached the administrative end of follow-up, which occurred on December 31, 2002.

Follow-up time was broken into discrete 60-day intervals. Within each 60-day period for each patient, we computed the proportion of days covered by osteoporosis medications using the dates on which prescriptions were filled and the “days supply” field in the pharmacy claim. If a prescription was refilled before an existing one was completed, the new prescription was assumed to start on the day the previous one should have ended. Within each 60-day interval, the proportion of days covered was calculated by summing the number of days in each interval covered by a prescription for an osteoporosis medication and dividing by the number of days in the interval in which the subject was not hospitalized or in a nursing home.

We restricted the present analysis to the members of the cohort who experienced an extended lapse in refill compliance. This occurred when a patient experienced a 60-day interval in which no days were covered by an osteoporosis medication. Among these patients, we defined the index date as the start of the 60-day period in which no days were covered by an osteoporosis medication. A return to medication use was defined by the filling of any prescription medication used to treat osteoporosis. This allowed patients to switch treatments and still be considered compliant.

The study investigators have Data Use Agreements in place with Center for Medicare and Medicaid Services and PACE. The Partners Healthcare Institutional Review Board approved this research.

**STATISTICAL METHODS**

We used the Kaplan-Meier method to estimate the survival distribution of the time until a patient returned to refill a medication used for treating osteoporosis and Cox proportional hazards models to estimate the hazard of resuming therapy.
using baseline variables defined during the year before initiation. For both analyses, subjects were censored by loss of PACE eligibility, death, or the end of follow-up. Baseline variables included in the Cox model were age, sex, number of co-morbid conditions, number of medications used, and history of the following in the year before initiation of an osteoporosis drug: a fracture, a nursing home visit, an acute care hospitalization, or a bone mineral density test.

We then performed a case-crossover analysis\textsuperscript{19,20} to determine if particular events occurring during the follow-up period triggered transitions back to compliance. The case-crossover approach stratifies the analysis across individuals, so that each patient serves as his or her own control (Figure 1). This design removes the confounding effects of all patient-level variables that are constant across time. The events that we considered were the occurrence of a new fracture, a bone mineral density test, acute care hospitalization, or a nursing home stay. The case-crossover analysis was implemented by comparing the frequency of events in the 60 days immediately before a transition in medication use (the hazard period) with the frequency of events in the period 120 days to 60 days before the transition occurred (the control period). This analysis was necessarily restricted to subjects who had at least 2 consecutive 60-day periods with no days covered by an osteoporosis medication. Because estrogen therapy may be used for indications other than osteoporosis, we repeated all analyses, excluding patients who started on estrogen. We conducted an additional sensitivity analysis in which “stopping” was defined to be 120 days not covered by medication, rather than 60 days. In a final sensitivity analysis, patients were censored by nursing home admission. All statistical analyses were performed in SAS 9.1.\textsuperscript{21}

\section*{RESULTS}

We identified 40,002 patients who initiated a medication used for osteoporosis during 1996-2002; this cohort has been described in previous work.\textsuperscript{8} Of these new users, 26,636 (67\%) had at least one 60-day period in which no days were covered by a prescription for any osteoporosis medication. The characteristics of this sample are given in Table 1. These patients had an average age of 82 years and were predominately white and female, had an average of 3.5 co-morbid conditions, and used 9 different medications. Slightly over 50\% had an acute care hospitalization, and 20\% had a nursing home stay during the year before initiation. Approximately 25\% had a fracture, and a similar proportion underwent bone mineral density testing during the year before the start of treatment. Among these patients, 11,857 (45\%) ultimately returned to medication use, 2994 (11\%) were censored by death, 3908 (15\%) were censored by loss of PACE eligibility, and 7877 (30\%) reached the end of follow-up without returning to regular use.

In Figure 2, we present Kaplan-Meier estimates of the survival function for time until a return to therapy. An

\begin{table}
\centering
\caption{Baseline Characteristics of Subjects in Sample}
\begin{tabular}{ll}
\hline
\textbf{n (\%)} or Mean & \\
\hline
\textbf{n} & 26,636 \\
Female sex & 25,660 (96.3) \\
Age, years & 81.8 (± 6.8) \\
White race & 25,550 (95.9) \\
Number of co-morbid conditions & 3.5 (± 3.1) \\
Number of different medications & 9.1 (± 5.4) \\
Number of physician visits & 10.1 (± 6.9) \\
Time on medication before & 9.6 (± 12.9) \\
interruption (in months) & \\
Acute care hospitalization & 14,178 (53.2) \\
Nursing home residence & 5271 (19.8) \\
Fracture of the hip, wrist, radius, or spine & 6559 (24.6) \\
Bone mineral density testing & 6338 (23.8) \\
Starting medications & \\
Bisphosphonate & 11,459 (43.0) \\
Calcitonin & 8691 (32.6) \\
Estrogen & 4054 (15.2) \\
Raloxifene & 1431 (5.4) \\
Bisphosphonate and Calcitonin & 608 (2.3) \\
Other combinations & 393 (1.5) \\
\hline
\end{tabular}
\end{table}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Kaplan-Meier estimate of the cumulative probability of returning to treatment.}
\end{figure}
estimated 30% of the population returned to regular use within 6 months after initially discontinuing medication and 50% by 2 years. In Figure 3, we present Kaplan-Meier estimates stratified by length of time on medication before the interruption in treatment. Patients with a longer history of medication use returned to treatment at a higher rate. Six months after an interruption in treatment, 42% of patients with 6 months or more of regular use returned to refill a prescription, and 59% returned by 2 years.

Table 2 presents the results of the Cox proportional hazards model examining the association between patient characteristics and the rate at which medication use was resumed. Younger female patients with a history of fracture and few co-morbid conditions were the most likely to return to treatment. Consistent with the Kaplan-Meier analysis, increasing the length of time spent on a medication before the interruption increased the probability of an ultimate return to compliance, with each additional 60-day period spent on medication increasing the probability of return by 3%. Patients who initiated estrogen therapy were the least likely to return to treatment after an interruption. Patients who started on calcitonin or a combination therapy were more likely to return than patients on a bisphosphonate. In Table 3, we present the results of the case-crossover analysis. A fracture or the use of a bone mineral density test was associated with a higher likelihood that a patient would re-start therapy in the subsequent 60-day period. Conversely, nursing home stays were associated with a lower likelihood of returning to therapy.

To assess whether treatment interruptions were due to medication switching, in Table 4 we present a cross-tabulation of drugs used before and after an observed interruption in treatment. For the majority (68%) of the 11,884 patients who were observed to return to treatment, the agent used after an interruption was the same as the agent used before the interruption. Patients on combination therapies often returned to use a single agent.

In the sensitivity analysis in which estrogen users were excluded, the results were almost identical. For example, we estimated that 51% of patients in this population return to treatment by 2 years (compared with 50%). When we re-defined “stopping” to be 120 days without a refill, many people still returned to treatment, but the estimated probability of return was substantially smaller (38% by 2 years). When patients were censored by nursing home admission, slightly more were estimated to return to use (56% by 2 years), but the results from the regressions were largely unchanged.

**DISCUSSION**

Within a population of moderate- to low-income older Americans, we found that long breaks in treatment were common among new users of osteoporosis medication. We previously reported that within 2 years of treatment initiation, a majority of new users experienced a period of at least 120 days after the completion of a prescription in which no refill was obtained. In the present study, we found that many of the same individuals that stopped refilling medication ultimately returned to treatment. Although many previous studies have reported that patient compliance with therapeutic regimens is poor, to our knowledge this is one of the few reports suggesting that

**Table 2** Cox Proportional Hazards Analysis of Baseline Variables Associated with a Return to Medication Use

<table>
<thead>
<tr>
<th>Variables Assessed at Initiation of Therapy</th>
<th>Hazard Ratio</th>
<th>95% LCL</th>
<th>95% UCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in decades</td>
<td>0.66</td>
<td>0.65</td>
<td>0.68</td>
</tr>
<tr>
<td>History of fracture in past year</td>
<td>1.18</td>
<td>1.12</td>
<td>1.23</td>
</tr>
<tr>
<td>History of bone mineral density testing in past year</td>
<td>1.02</td>
<td>0.97</td>
<td>1.06</td>
</tr>
<tr>
<td>History of a nursing home visit in past year</td>
<td>1.04</td>
<td>0.98</td>
<td>1.10</td>
</tr>
<tr>
<td>History of hospitalization in past year</td>
<td>1.01</td>
<td>0.97</td>
<td>1.05</td>
</tr>
<tr>
<td>Female sex</td>
<td>1.63</td>
<td>1.45</td>
<td>1.84</td>
</tr>
<tr>
<td>Number of physician visits (per 5 visits)</td>
<td>1.04</td>
<td>1.03</td>
<td>1.06</td>
</tr>
<tr>
<td>Number of comorbid conditions (per 5 conditions)</td>
<td>0.71</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td>Number of other prescription medications (per 5 medications)</td>
<td>1.01</td>
<td>0.99</td>
<td>1.03</td>
</tr>
<tr>
<td>Length of time spent on osteoporosis medication</td>
<td>1.03</td>
<td>1.02</td>
<td>1.03</td>
</tr>
<tr>
<td>Type of osteoporosis medication used*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calcitonin</td>
<td>1.22</td>
<td>1.16</td>
<td>1.27</td>
</tr>
<tr>
<td>Estrogen</td>
<td>0.67</td>
<td>0.63</td>
<td>0.71</td>
</tr>
<tr>
<td>Raloxifene</td>
<td>0.93</td>
<td>0.85</td>
<td>1.01</td>
</tr>
<tr>
<td>Bisphosphonate + Calcitonin</td>
<td>1.20</td>
<td>1.06</td>
<td>1.35</td>
</tr>
<tr>
<td>Other combination therapy</td>
<td>1.17</td>
<td>1.01</td>
<td>1.34</td>
</tr>
</tbody>
</table>

LCL = lower confidence limit; UCL = upper confidence limit. Bisphosphonate is the reference category.
many patients return to refill prescriptions after long periods of apparent nonuse.

The process of patients returning to therapy is associated with a variety of subject characteristics and events. Female patients and those with a history of fracture were the most likely to return to refill prescriptions. Patients with a prior fracture have a strong indication for osteoporosis treatment and possibly viewed themselves at particular risk. However, older patients and those with many co-morbid conditions were less likely to return to treatment. For this frail group, it is possible that preventive therapies were a lower priority than treatment for other more acute medical conditions. Several factors during the period of medication nonuse were strong predictors of returning to therapy, including a new fracture and bone mineral density testing. Similar to a fracture during the baseline period, the occurrence of a fracture after discontinuation presents a compelling reason for a patient to return to therapy and puts the patient in contact with the health care system. Use of bone mineral density testing was possibly a marker for concern about osteoporosis by the physician or the patient.

Recent nursing home stays were negatively associated with a return to therapy. This finding is consistent with other studies that have observed that few patients receive osteoporosis treatment while in nursing homes. The low use of osteoporosis treatment in nursing homes may reflect a reluctance of nursing home staff to use nonessential treatments or possibly a concern about the use of bisphosphonates in patients who may be supine most of the day. However, it underscores the fact that many patients at high risk of fracture are likely to have stopped treatment in this environment.

We have conceptualized compliance as a dichotomous state, but medication use is complex. Patients who appear to be completely noncompliant may be taking their medications continuously but infrequently. This is supported by the observations that many people who appear to have stopped return immediately and that fewer patients return when a break in treatment is defined to be 120 days. Nevertheless, the strong association between events such as an incident fracture and a new refill strongly suggests that compliance behavior can change abruptly. For some patients, however, the behavior change may be from “under use” to “regular use”, rather than from “non-use” to “use.”

We acknowledge other limitations of our study related to our use of administrative health care utilization data. First, we are unable to assess the reasons for starting or stopping a medication. Thus, some of the lapses in treatment that we have observed could be appropriate and physician-directed. For example, the interruption could represent a switch in therapy because of an adverse event or concern about one. Or it is possible that a physician concerned about a potential adverse reaction might suspend treatment rather than immediately switch a patient to a new drug. Second, we do not capture data on IV bisphosphonates or over-the-counter calcium and vitamin D. Patients who are temporarily switched to one of these therapies would be assumed to have experienced a gap in treatment. Finally, it is possible

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Results from Case-crossover Analysis of Events Predicting a Return to Medication Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Occurring in 60-day Period</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>Before Transition in Medication Use Status</td>
<td></td>
</tr>
<tr>
<td>Fracture</td>
<td>1.30</td>
</tr>
<tr>
<td>Bone mineral density test</td>
<td>1.26</td>
</tr>
<tr>
<td>Nursing home stay</td>
<td>0.90</td>
</tr>
<tr>
<td>Hospitalization</td>
<td>1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Cross-tabulation of the Drugs Used Before and After the Break in Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Drug</td>
<td>Bisphosphonate</td>
</tr>
<tr>
<td>Bisphosphonate</td>
<td>3293*</td>
</tr>
<tr>
<td></td>
<td>69.2%*</td>
</tr>
<tr>
<td>Calcitonin</td>
<td>829</td>
</tr>
<tr>
<td></td>
<td>18.4%</td>
</tr>
<tr>
<td>Estrogen</td>
<td>335</td>
</tr>
<tr>
<td></td>
<td>20.2%</td>
</tr>
<tr>
<td>Bisphosphonate &amp; Calcitonin</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>20.2%</td>
</tr>
<tr>
<td>Other combination</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>30.2%</td>
</tr>
<tr>
<td>Raloxifene</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>24.5%</td>
</tr>
<tr>
<td>Total</td>
<td>4713</td>
</tr>
<tr>
<td></td>
<td>39.8%</td>
</tr>
</tbody>
</table>

*Reported percentages are row percents.
that some of the interruptions in use that we have observed could have resulted from patients developing stockpiles at home (eg, from physician samples). Future research in this area should focus on developing a greater understanding of the reasons why patients often go for extended periods of time between refills.

For those involved with efforts to improve patient compliance with medication, our results yield insight into the process of compliance, identify particular patient groups most likely to return from temporary interruptions in therapy, and also point to events that predict transitions in medication use. These findings also have specific relevance to physicians treating patients for osteoporosis; namely, that prolonged breaks from osteoporosis medications are probably common. Because the effectiveness of osteoporosis medications used in an interrupted way is unknown, physicians should stress to their patients the importance of continuous medication use. When persistently low bone mineral density test results are observed, physicians should attempt to determine whether the patient has been consistently using their prescribed medication before changing treatment or dosing.

Although our study points out an underappreciated problem with long-term medication use, it also gives some reasons for optimism: patients frequently return to refill prescriptions after long episodes of apparent noncompliance. Cross-sectional estimates of compliance based on refill data may paint an overly pessimistic picture of long-term medication use. These estimates will miss the potentially numerous patients who have not had their prescription refilled recently but who will ultimately return to regular use. Although it is clear that long-term compliance with osteoporosis treatment is poor by any measure, our study suggests the possibility that there are many patients who, if properly motivated, could become more regular medication users.

References

Development of a Risk Score for Colorectal Cancer in Men

Jane A. Driver, MD, MPH,a J. Michael Gaziano MD, MPH,a,b,e Rebecca P. Gelber, MD, MPH,e I-Min Lee, MBBS, ScD,b,c Julie E. Buring, ScD,a,b,c,d Tobias Kurth, MD, ScD,a,b,c

*aThe Brigham and Women’s Hospital Divisions of Aging and bPreventive Medicine, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School; cDepartment of Epidemiology, Harvard School of Public Health; dDepartment of Ambulatory Care and Prevention, Harvard Medical School; eMassachusetts Veterans Epidemiology Research Information Center, Boston VA Healthcare System, Boston, Mass.

ABSTRACT

BACKGROUND: Colorectal cancer is a common and preventable disease for which screening rates remain unacceptably low.

METHODS: We developed a risk scoring system for the development of colorectal cancer among participants in the Physician’s Health Study, a prospective cohort of 21,581 US male physicians who were all free of cancer. Predictors of colorectal cancer were self-reported and identified from the baseline questionnaire. Logistic regression was used to determine the independent predictors of incident colorectal cancer over the follow-up period. Risk scores were created from the sum of the odds ratios of the final predictors and used to divide the cohort into categories of increasing relative risk.

RESULTS: During 20 years of follow-up, 381 cases of colon cancer and 104 cases of rectal cancer developed in the cohort. Age, alcohol use, smoking status, and body mass index were independent significant predictors of colorectal cancer. The point scores were used to define 10 risk groups. Those in the highest risk group (9-10 points) had an odds ratio of 15.29 (6.19-37.81) for colorectal cancer compared with those with the lowest risk. We further stratified scores into 3 risk classes. Compared with those at the lowest relative risk, the odds ratio for colorectal cancer was 3.07 (2.46-3.83) in the intermediate risk group and 5.75 (4.44-7.44) in the highest risk group.

CONCLUSIONS: We developed a simple scoring system for colorectal cancer that identifies men at increased relative risk on the basis of age and modifiable factors. This tool should be validated in other populations. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Colon cancer; Rectal cancer; Risk score; Risk prediction; Behavioral risk factors
SUBJECTS AND METHODS

Study Cohort
This analysis used baseline data from the Physician’s Health Study (PHS), a randomized, placebo-controlled trial of aspirin and beta-carotene for the prevention of cardiovascular disease and cancer among 22,071 US male physicians. A detailed description of the trial cohort, methods, and findings of the study has been reported. Baseline data were collected at study entry in 1982. Participants ranged in age from 40 to 84 years; were apparently healthy; had no history of cardiovascular disease, cancer (with the exception of nonmelanomatous skin cancer), or other serious illnesses; and had no indication or contraindication for aspirin or other pain medication use. Ninety-two percent of study participants were white. Regular users of aspirin were not considered for the trial. We excluded those participants who reported a history of cancer before the receipt of the baseline questionnaire (22 men) and those who had missing information on smoking, height, weight, exercise, alcohol use, and history of diabetes (468 men), leaving 21,581 men for the analysis.

Definitions of Potential Predictors
We performed a literature review and considered all well-accepted and biologically plausible risk factors for colorectal cancer available in the baseline data as potential predictors. The following self-reported baseline variables were included in the analysis: categoric age in years (40-49, 50-59, 60-69, and ≥70); body mass index (BMI) (<25.0 kg/m², 25.0-29.9 kg/m², and ≥30.0 kg/m²); smoking status (never, past, 1 pack/day, 2 packs/day); alcohol use (rarely, weekly, and daily); intake of vegetables (rarely vs ≥ weekly); intake of multivitamins, vitamin C, and vitamin E (never vs past/current); intake of cold cereal (≥ once per week vs < once per week); physical activity vigorous enough to work up a sweat (> monthly vs rarely or never); and history of diabetes (yes/no).

Ascertainment of Colon and Rectal Cancer Cases
The study outcome was the development of colon or rectal cancer during the 20-year follow-up period. Study participants completed mailed questionnaires every 6 months during the first year and then annually. Nonfatal cases of colon and rectal cancer were reported by the participants on follow-up questionnaires, and fatal cases were reported by family members or next of kin. Cases of colorectal cancer were confirmed by review of medical records and pathology reports by an EndPoints Committee of study physicians. By March 2004, we had identified 381 cases of colon cancer and 104 cases of rectal cancer.

CLINICAL SIGNIFICANCE
- We developed a risk scoring system for the development of colorectal cancer based on age, smoking history, alcohol use, and body mass index in a cohort of more than 21,000 men.
- Those in the highest risk group (9-10 points) had an OR of 15.29 (6.19-37.81) for colorectal cancer compared with those with the lowest risk.
- The system can be used to help men understand their relative risk of colorectal cancer based on modifiable risk factors and inform discussions about screening and lifestyle modification.

Statistical Analyses
Descriptive statistics, expressed as proportions for categoric variables and means for continuous variables, were used to compare the characteristics of those with and without the outcome. To determine the univariate association between the baseline factors and the development of colorectal cancer, we used t tests for continuous variables and chi-square tests for categoric variables. All statistical calculations were performed using SAS software (v. 9.1; SAS Institute Inc, Cary, NC); a 2-tailed P value of less than .05 was considered statistically significant. Factors that were found to have a significant univariate relationship to the outcome were included in a multivariable logistic regression model. Only predictors that retained their significance in the adjusted model remained in the final model. Each predictor in the final model was assigned a point value that corresponded to its odds ratio (OR) rounded to the nearest whole integer. The points were then summed to create a risk score for each participant. The cohort was divided into risk categories that corresponded to each risk score.

Discrimination of the model, risk scores, and risk categories was assessed by means of the area under the receiver operating characteristic curve (AUC). The AUC is the probability that the predicted value for the subject who developed the outcome will be greater than that for the subject who did not develop the outcome. The Hosmer-Lemeshow goodness-of-fit statistic was used to assess the reliability of the models, where P greater than .05 indicates adequate calibration. Validation of the models was performed in the derivation cohort by bootstrap resampling with replacement using 200 iterations of the study cohort.

We compared the AUC of the model containing the predictors with that of the risk score and risk categories using the method of DeLong et al. We examined the calibration of the risk categories by calculating the observed versus the expected number of events in each risk class. The predicted probability of the outcome for each member of the cohort was calculated on the basis of the final predictor model. Within each risk category, these risks were summed to produce an average predicted risk, which was compared
with the observed incidence of the outcome in the same risk class.

Because prior research suggests that risk factors for colon and rectal cancer may differ, we developed a separate model for colon cancer and compared its accuracy and performance with that of the colorectal model.

RESULTS

Colorectal Model

A total of 485 of the 21,581 subjects in the PHS cohort developed colorectal cancer during the 20-year follow-up period. Those who developed colorectal cancer were significantly older (mean age 61 vs 54 years), were more likely to be overweight or obese, drank alcohol regularly, exercised rarely, and had a history of smoking. Vegetable intake, cold cereal intake, and vitamin use were not significant predictors on univariate analyses. Baseline characteristics of the groups and results of the univariate analyses are displayed in Table 1.

Four variables remained statistically significant in the multivariable logistic regression model: age, BMI, history of smoking, and weekly or daily alcohol use. The efficacy of the multivariable model was compared with that obtained by a risk factor sum. The score was

| Table 1 Baseline Characteristics of 21,581 Men with 20 Years of Follow-up and Incident Cases of Colon and Rectal Cancer |
|---|---|---|---|---|
| Characteristics | Prevalence n (%) | Colon Cancer n (%) | Rectal Cancer n (%) | p* |
| Age (y) | | | | |
| 40-49 | 8871 (41.1) | 52 (13.7) | 29 (27.9) | <.001 |
| 50-59 | 7314 (33.9) | 117 (30.7) | 41 (39.4) | .01 |
| 60-69 | 3993 (18.5) | 145 (38.0) | 21 (20.2) | <.001 |
| 70+ | 1403 (6.5) | 67 (17.6) | 13 (12.5) | .02 |
| History of smoking | | | | |
| Never | 10,689 (49.5) | 138 (36.2) | 42 (40.4) | <.001 |
| Past | 8521 (39.5) | 197 (51.7) | 44 (42.3) | .02 |
| 1 pack/d | 830 (3.9) | 12 (3.2) | 3 (2.9) | |
| 2 packs/d | 1541 (7.1) | 34 (8.9) | 15 (14.4) | |
| Alcohol use | | | | |
| Rarely or never | 5607 (26.0) | 80 (21.0) | 17 (16.4) | .007 |
| Weekly | 10,594 (49.1) | 182 (47.8) | 51 (49.0) | .02 |
| Daily | 5380 (24.9) | 119 (31.2) | 36 (34.6) | |
| Body mass index | | | | |
| <25.0 kg/m² | 12,431 (57.6) | 190 (49.9) | 54 (51.9) | .008 |
| 25.0-29.9 kg/m² | 8246 (38.2) | 171 (44.9) | 42 (40.4) | .15 |
| ≥30 kg/m² | 904 (4.2) | 20 (5.2) | 8 (7.7) | |
| History of diabetes | | | | |
| No | 21,067 (97.6) | 363 (95.3) | 102 (98.1) | .003 |
| Yes | 514 (2.4) | 18 (4.7) | 2 (1.9) | .76 |
| Exercise | | | | |
| ≥Monthly | 18,660 (86.5) | 316 (82.9) | 88 (84.6) | .042 |
| Rarely or never | 2921 (13.5) | 65 (17.1) | 16 (15.4) | .58 |
| Vegetable intake | | | | |
| ≥Weekly | 12,453 (57.7) | 213 (55.9) | 65 (62.5) | .47 |
| ≤Rarely (1-3×/mo) | 9125 (42.3) | 168 (44.1) | 39 (37.5) | .32 |
| Cold cereal intake | | | | |
| ≥Once/wk | 13,945 (66.9) | 243 (68.1) | 64 (64.0) | .63 |
| <Once/wk | 6910 (33.1) | 114 (31.9) | 36 (36.0) | .54 |
| Multivitamin use | | | | |
| Past/current | 7706 (35.8) | 145 (38.2) | 40 (38.8) | .34 |
| Never | 13,801 (64.2) | 235 (61.8) | 63 (61.2) | .52 |
| Vitamin C use | | | | |
| Past/current | 4957 (23.0) | 88 (23.2) | 24 (23.1) | .93 |
| Never | 16,613 (77.0) | 292 (76.8) | 80 (76.9) | .98 |
| Vitamin E use | | | | |
| Past/current | 2233 (10.4) | 45 (11.8) | 10 (9.7) | .35 |
| Never | 19,323 (89.6) | 336 (88.2) | 93 (90.3) | .83 |

Effective sample size of some groups may vary because of missing variables.

*p values from chi-square tests of association between each characteristic and the outcome of incident cases of colon or rectal cancer.
calculated from the sum of the ORs of the final predictors as follows: risk score = (2 x age 50-59 years) + (4 x age 60-69 years) + (6 x age ≥ 70 years) + (1 x BMI ≥ 25-30 kg/m²) + (2 x BMI ≥ 30 kg/m²) + (1 x history of past or current smoking) + (1 x alcohol use ≥ weekly). One score was given for each risk factor, and the risk factor sum for each individual ranged from 0 to 10. There was no statistically significant difference between the AUC of the risk factor score (c = 0.692) and the predictor model (c = 0.695), with a P value for the comparison of .075.

The point scores were used to define 10 risk groups. Points 9 and 10 were collapsed into 1 because of the small number in group 10. There was no statistically significant difference between the AUC of the risk factor score (c = 0.692) and the predictor model (c = 0.695), with a P value for the comparison of .075.

Table 2

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Unadjusted OR (95% CI)</th>
<th>Adjusted Beta Coefficient</th>
<th>OR (95% CI)</th>
<th>Prediction Score Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>2.40 (1.83-3.14)</td>
<td>0.81</td>
<td>2.25 (1.72-2.95)</td>
<td>2</td>
</tr>
<tr>
<td>60-69</td>
<td>4.71 (3.60-6.16)</td>
<td>1.48</td>
<td>4.40 (3.36-5.77)</td>
<td>4</td>
</tr>
<tr>
<td>≥70</td>
<td>6.56 (4.79-8.99)</td>
<td>1.83</td>
<td>6.25 (4.55-8.60)</td>
<td>6</td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.68 (1.40-2.03)</td>
<td>0.35</td>
<td>1.42 (1.17-1.72)</td>
<td>1</td>
</tr>
<tr>
<td>Body mass index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.0-29.9 kg/m²</td>
<td>1.32 (1.10-1.60)</td>
<td>0.23</td>
<td>1.26 (1.05-1.52)</td>
<td>1</td>
</tr>
<tr>
<td>≥30 kg/m²</td>
<td>1.60 (1.07-2.38)</td>
<td>0.48</td>
<td>1.62 (1.09-2.42)</td>
<td>2</td>
</tr>
<tr>
<td>Alcohol use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥Once/wk</td>
<td>1.41 (1.13-1.77)</td>
<td>0.31</td>
<td>1.36 (1.08-1.71)</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.79 (1.14-2.83)</td>
<td>0.19</td>
<td>1.21 (0.76-1.92)</td>
<td>.....</td>
</tr>
<tr>
<td>Rarely exercise</td>
<td>1.29 (1.01-1.64)</td>
<td>0.06</td>
<td>1.06 (0.83-1.36)</td>
<td>.....</td>
</tr>
</tbody>
</table>

OR = odds ratio; CI = confidence interval.

A second risk class model was created to identify those at relatively low (0-3 points), intermediate (4-6 points), or relatively high (7-10 points) risk of colorectal cancer. The 20-year observed risk of colorectal cancer was 1% in the lowest risk group, 3% in the intermediate risk group, and 6% in the highest risk group (Table 4).

### Comparison of Predictors for Colon and Rectal Cancer

Participants with rectal cancer were older than the cohort mean (age 57 vs 54 years) but significantly younger than those with colon cancer (mean age 57 vs 61 years). In the multivariable analysis, colon cancer was associated with the following baseline factors: each age category, history of smoking, and BMI of 25 kg/m² or greater. In contrast, rectal cancer was associated only with the oldest age category

Table 3

<table>
<thead>
<tr>
<th>No. Points</th>
<th>No. (%) Patients</th>
<th>Predicted n with CRC</th>
<th>Observed n with CRC</th>
<th>Predicted OR (95% CI) of CRC</th>
<th>20-year Cumulative Risk of CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>977 (5)</td>
<td>5</td>
<td>6</td>
<td>1.00</td>
<td>0.006</td>
</tr>
<tr>
<td>1</td>
<td>3090 (14)</td>
<td>23</td>
<td>28</td>
<td>1.48 (0.61-3.59)</td>
<td>0.009</td>
</tr>
<tr>
<td>2</td>
<td>3981 (18)</td>
<td>41</td>
<td>37</td>
<td>1.52 (0.64-3.61)</td>
<td>0.009</td>
</tr>
<tr>
<td>3</td>
<td>3422 (16)</td>
<td>52</td>
<td>47</td>
<td>2.25 (0.96-5.29)</td>
<td>0.014</td>
</tr>
<tr>
<td>4</td>
<td>3356 (16)</td>
<td>75</td>
<td>75</td>
<td>3.70 (1.61-8.52)</td>
<td>0.022</td>
</tr>
<tr>
<td>5</td>
<td>2656 (12)</td>
<td>81</td>
<td>85</td>
<td>5.35 (2.33-12.29)</td>
<td>0.032</td>
</tr>
<tr>
<td>6</td>
<td>1953 (9)</td>
<td>82</td>
<td>86</td>
<td>7.46 (3.25-17.12)</td>
<td>0.044</td>
</tr>
<tr>
<td>7</td>
<td>1268 (6)</td>
<td>66</td>
<td>58</td>
<td>7.76 (3.33-18.05)</td>
<td>0.046</td>
</tr>
<tr>
<td>8</td>
<td>600 (3)</td>
<td>37</td>
<td>39</td>
<td>11.25 (4.73-26.74)</td>
<td>0.065</td>
</tr>
<tr>
<td>9-10</td>
<td>278 (1)</td>
<td>22</td>
<td>24</td>
<td>15.29 (6.19-37.81)</td>
<td>0.086</td>
</tr>
</tbody>
</table>

CRC = colorectal cancer; CI = confidence interval; OR = odds ratio.
(≥70 years), current smoking of at least 2 packs/day, and daily alcohol use. Table 5 compares the predictors of colon and rectal cancer.

**Colon Cancer Prediction Model**

A predictive model for colon cancer was constructed using independent risk factors from the multivariable model: age category, smoking status, and BMI category. The AUC of this model was only slightly superior to the model predicting colorectal cancer (c = 0.717 vs 0.695), but the goodness-of-fit test showed it to perform less well than the colorectal model (Hosmer-Lemeshow statistic 0.43 vs 0.91).

**DISCUSSION**

In this large prospective cohort of apparently healthy men, we found that a model using age, alcohol use, smoking status, and BMI predicted the development of colorectal cancer over a 20-year follow-up period. Not surprisingly, increasing age was the strongest predictor of colorectal cancer. However, when age was excluded from the model, the other factors remained statistically significant. Those in the highest risk group had a 15-fold increased odds of developing colorectal cancer compared with those in the lowest risk group.

The risk score presented in Table 3 can serve as a complement to age-based screening strategies by helping to motivate those in the highest risk classes. For example, a 59-year-old man with none of the risk factors in the model has an OR of 1.5 for developing colorectal cancer compared with those with the lowest risk, whereas a man of the same age who is obese, smokes, and drinks regularly has an OR of 7.5. Previous work has shown that information about individualized colon cancer risk can lead to a reduction in multiple behavioral risk factors in patients with a history of colon adenoma.14

Once validated in other populations, our model might have applications such as estimating risk for cost-benefit analyses or the stratification of populations into risk groups for differential screening strategies. Table 4 offers an example of how the risk scores could be used to stratify a population into different risk levels.

**Table 4** Performance of 3 Risk Groups for Colorectal Cancer

<table>
<thead>
<tr>
<th>Risk Group</th>
<th>No. (%) Patients</th>
<th>Predicted n with CRC</th>
<th>Observed n with CRC</th>
<th>Predicted OR (95% CI) of CRC</th>
<th>Observed 20-year Risk CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest (0-3 points)</td>
<td>11,470 (53)</td>
<td>121</td>
<td>118</td>
<td>1.00</td>
<td>1%</td>
</tr>
<tr>
<td>Intermediate (4-6 points)</td>
<td>7965 (37)</td>
<td>238</td>
<td>246</td>
<td>3.07 (2.46-3.83)</td>
<td>3%</td>
</tr>
<tr>
<td>Highest (7-10 points)</td>
<td>2146 (10)</td>
<td>126</td>
<td>121</td>
<td>5.75 (4.44-7.44)</td>
<td>6%</td>
</tr>
</tbody>
</table>

CRC = colorectal cancer; CI = confidence interval; OR = odds ratio.

**Table 5** Comparison of Risk Factors for Colon and Rectal Cancer*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Colon Cancer n = 381</th>
<th>Rectal Cancer n = 104</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (compared with cohort)</td>
<td>Unadjusted n = 381</td>
<td>Adjusted n = 355</td>
</tr>
<tr>
<td>Age (y)</td>
<td>60.78 (59.85-61.70)</td>
<td>57.34 (55.30-59.38)</td>
</tr>
<tr>
<td>50-59</td>
<td>2.76 (1.99-3.83)</td>
<td>2.54 (1.81-3.56)</td>
</tr>
<tr>
<td>60-69</td>
<td>6.39 (4.65-8.79)</td>
<td>5.92 (4.25-8.25)</td>
</tr>
<tr>
<td>≥70</td>
<td>8.51 (5.89-12.27)</td>
<td>7.99 (5.41-11.80)</td>
</tr>
<tr>
<td>History of smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>1.81 (1.45-2.25)</td>
<td>1.50 (1.19-1.89)</td>
</tr>
<tr>
<td>1 PPD</td>
<td>1.12 (0.62-2.03)</td>
<td>1.06 (0.57-1.98)</td>
</tr>
<tr>
<td>2 PPD</td>
<td>1.73 (1.18-2.52)</td>
<td>1.53 (1.02-2.29)</td>
</tr>
<tr>
<td>Alcohol use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly</td>
<td>1.21 (0.93-1.57)</td>
<td>1.30 (0.98-1.72)</td>
</tr>
<tr>
<td>Daily</td>
<td>1.56 (1.17-2.08)</td>
<td>1.19 (0.88-1.62)</td>
</tr>
<tr>
<td>BMI &gt;25.0 kg/m²</td>
<td>1.37 (1.12-1.68)</td>
<td>1.38 (1.11-1.70)</td>
</tr>
</tbody>
</table>

PPD = packs per day; BMI = body mass index.

*Multivariable model also adjusted for history of diabetes, exercise level, vegetable intake, cold cereal intake, vitamin C, vitamin E, and multivitamin intake. Only significant variables for colon or rectal cancer are shown.
population at “average risk.” Future studies could better define appropriate cutoffs for the point scores when validating them in other populations.

Our model is meant to be used to help a man estimate his risk class of developing colorectal cancer on the basis of his age and significant lifestyle predictors. The risk score may not be used to estimate absolute risk. Among those for whom screening is recommended on the basis of age, consideration of the relative risk of disease may be more motivating and informative to patients than discussion of absolute risk.

Our approach was to build a simple and parsimonious model focusing on modifiable predictors, rather than an exhaustive one. The inclusion of highly correlated variables does not increase model prediction. We chose predictors that were widely accepted, independent, and easily described.

To our knowledge, only 1 prior attempt has been made to use risk factor information to create a risk index for large bowel cancer: the Harvard Cancer Risk Index. The simplicity of our model and its ease of use in a clinical encounter make it an important complement to this project. Although the Harvard Index included only those aged 40 to 70 years, our model includes men aged more than 70 years, a population in whom individualized information on risk is particularly helpful. Validation of the discrimination of our model by bootstrap resampling (c = 0.70) generated similar results to the external validation of other cancer risk-assessment models.

The risk factors for colon cancer identified in our study are consistent with the findings of other multivariable analyses in men, women, and both genders. In our cohort, alcohol use and smoking were shared risk factors for both colon and rectal cancer. These findings are similar to those found in another large prospective study of men. Stratification by age showed an increasing effect of smoking by age group and the strongest effect of alcohol in the youngest group (data not shown). The model was unstable in the youngest age groups because of the low number of colorectal cancers.

To control for aspirin use, we adjusted for randomized aspirin assignment in the final models. There was no association between randomized aspirin use during the study and the development of colorectal cancer. These findings agree with prior work showing no association between aspirin or nonsteroidal anti-inflammatory drug use and colorectal cancer in the PHS cohort.

Although there are differences between the predictors of colon and rectal cancer, a tool to predict both types of large bowel cancer has more utility in the clinical setting. When we compared a model limited to colon cancer with the combined model, the AUC was only slightly improved and the goodness-of-fit was less. We thus chose to use a model that predicts both proximal and distal cancers.

The strengths of our study include its prospective nature, large number of outcome events and participants, and long follow-up period. Cases of colorectal cancer were confirmed only after medical record review. Finally, the homogeneous nature of our cohort reduces potential confounding by differences due to factors such as access to medical care and socioeconomic status. However, a number of methodologic limitations need to be considered.

We had no baseline information on family history of colon or rectal cancer. Inclusion of such information may have increased the predictive ability of our model, but we would not expect it to alter the value of the relative risks of the factors we chose to include. We were also not able to control for colorectal cancer screening. However, because the PHS was a prevention trial among physicians, differences in screening patterns are likely not as significant as in other populations.

We could not externally validate the prediction rule for colorectal cancer and decided against data splitting to have more power identifying potential risk factors. We internally validated the model and chose only biologically plausible risk factors, so we are confident that our prediction model is generalizable to other male populations.

Our model was built on only baseline information; thus, our ability to accurately describe exposures is limited compared with studies that take time of exposure into account. We chose a logistic regression model to better simulate a screening decision based on information obtained at 1 clinical encounter; however, when we re-ran the analyses using the proportional hazards model, the same predictors were identified.

Finally, caution should be used when generalizing the results of our model because it was derived in a cohort of predominantly white male physicians. Although the absolute risk of colorectal cancer in a cohort of relatively healthy physicians is likely lower than that of the general population, the relative risks associated with the predictors should be similar because the biological mechanism by which these factors are associated with colorectal cancer is not expected to be different in our cohort.

Although the prediction model should be validated in other populations of both men and women, the risks for colorectal cancer included in the model are common to both sexes.

In summary, we were able to create a risk score to predict the 20-year risk of colorectal cancer in a prospective cohort of apparently healthy US men. This model has potential clinical utility and should be validated in other populations, particularly in women.

ACKNOWLEDGMENTS

We are grateful to the staff of the PHS and to the 22,071 dedicated physicians who made this project possible.

References


Smoking, Alcohol Consumption, and Raynaud’s Phenomenon in Middle Age

Lisa G. Suter, MD,a,b Joanne M. Murabito, MD, ScM,c,d David T. Felson, MD, MPH,e Liana Fraenkel, MD, MPHa,b

aVA Connecticut Healthcare System, West Haven, Conn; bYale University School of Medicine, New Haven, Conn; cNational Heart, Lung, and Blood Institute’s Framingham Heart Study, Framingham, Mass; dBoston University School of Medicine, Section of General Internal Medicine, Boston, Mass; eBoston University School of Medicine, Clinical Epidemiology, Boston, Mass.

ABSTRACT

BACKGROUND: Data suggest Raynaud’s phenomenon shares risk factors with cardiovascular disease. Studies of smoking, alcohol consumption, and Raynaud’s have produced conflicting results and were limited by small sample size and failure to adjust for confounders. Our objective was to determine whether smoking and alcohol are independently associated with Raynaud’s in a large, community-based cohort.

METHODS: By using a validated survey to classify Raynaud’s in the Framingham Heart Study Offspring Cohort, we performed sex-specific analyses of Raynaud’s status by smoking and alcohol consumption in 1840 women and 1602 men. Multivariable logistic regression analyses were used to examine the relationship of Raynaud’s to smoking and alcohol consumption.

RESULTS: Current smoking was not associated with Raynaud’s in women but was associated with increased risk in men (adjusted odds ratio [OR] 2.59, 95% confidence interval [CI], 1.11-6.04). Heavy alcohol consumption in women was associated with increased risk of Raynaud’s (adjusted OR 1.69, 95% CI, 1.02-2.82), whereas moderate alcohol consumption in men was associated with reduced risk (adjusted OR 0.51, 95% CI, 0.29-0.89). In both genders, red wine consumption was associated with a reduced risk of Raynaud’s (adjusted OR 0.59, 95% CI, 0.36-0.96 in women and adjusted OR 0.30, 95% CI, 0.15-0.62 in men).

CONCLUSIONS: Our data suggest that middle-aged women and men may have distinct physiologic mechanisms underlying their Raynaud’s, and thus sex-specific therapeutic approaches may be appropriate. Our data also support the possibility that moderate red wine consumption may protect against Raynaud’s.

© 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Raynaud’s phenomenon; Risk factors; Cigarette smoking; Alcohol consumption; Cardiovascular disease; Gender differences
cardiovascular disease. Some studies reported associations of smoking with Raynaud’s in men but not in women, whereas others noted that smoking was significantly associated with Raynaud’s in women but not in men. Of five studies examining alcohol consumption and Raynaud’s, none found significant associations, three were limited by sample size, and none examined alcohol type. Therefore, we sought to determine whether smoking and alcohol were associated with Raynaud’s in a large cross-sectional analysis of a middle-aged population.

**PARTICIPANTS AND METHODS**

**Participants**

Subjects were participants in the Framingham Offspring Study cohort without occupational vibratory tool use. This is a white, community-based population of children of original Framingham subjects. From September 1998 to October 2001 (Examination 7), participants were administered a standardized instrument, developed in a population-based sample, with 100% sensitivity and specificity in distinguishing participants with and without Raynaud’s, using physician impression as the accepted standard. A detailed medical record review of a majority sample of subjects by a trained rheumatologist revealed no clinical or laboratory findings consistent with a diagnosis of an underlying connective tissue disease, nor were any cases of Buerger’s disease identified in an ancillary study chart review. Boston University Medical Center and Yale School of Medicine institutional review boards approved our study.

**Data Collection**

Our methods used validated criteria for Raynaud’s originally defined by Maricq and Weinrich (presence of at least three of four criteria in the past 12 months).

**Smoking Status.** Cigarette smoking status was defined as never (no history of smoking), past (those who smoked at any point before examination), or current (those who smoked regularly within the last 12 months). Number of cigarettes smoked per day was examined as an ordinal variable (none; 1-10 per day; 11-20 per day; >20 per day) to look for dose-response relationships.

**Alcohol Consumption.** Self-report of the number of drinks per week over the past year was recorded for beer, white wine, red wine, and liquor. Total absolute alcohol consumption in ounces per week was calculated using a formula: $[(0.44) \times (\text{number of beers per week}) + (0.40) \times (\text{number of glasses of wine per week}) + (0.57) \times (\text{number of cocktails per week})]$. On the basis of prior work with this cohort, we hypothesized that moderate alcohol consumption may be protective for Raynaud’s, as it is for cardiovascular disease, whereas heavy consumption may be deleterious. Alcohol consumption was defined as none, moderate, or heavy, with a priori cut points based on prior data and guidelines:

- None: $<1$ oz absolute alcohol per week for both women and men
- Moderate: $\geq 1$ but $\leq 3.5$ oz per week for women and $\geq 1$ but $\leq 7$ oz per week for men
- Heavy: $>3.5$ oz per week for women and $>7$ oz per week for men

Subgroup analyses by light-moderate ($\geq 1$ but $\leq 3.5$ oz per week) and heavy-moderate ($>3.5$ but $\leq 7$ oz per week) alcohol consumption were also performed. Analyses by type of alcohol (beer, liquor, white or red wine) were used as dichotomous variables (none vs any), defined using the average number of drinks per week ($12$ oz of beer, $1$ oz of liquor, and $4$ oz of either type of wine).

Covariates, selected because of known or suspected associations with Raynaud’s phenomenon, smoking, and alcohol consumption, were defined as follows:

**Age and Body Mass Index.** Age and body mass index (BMI) were examined as continuous variables.

**Use of Antihypertensive Medications.** The current use of antihypertensive medications was examined as current use of vasodilating agents (eg, calcium channel blockers and angiotensin-blocking agents) and current use of beta-blocker medications.

**Cardiovascular Disease.** The presence of cardiovascular disease was defined as a history of angina, coronary insufficiency, myocardial infarction, congestive heart failure, intermittent claudication, stroke, or transient ischemic attack. A 3-physician investigatory panel (or neurologists for cerebrovascular outcomes) made final determinations as to diagnosis based on hospital records and physicians’ reports.

**Menopausal and Hormone Status.** In women, menopause was defined as the absence of menstrual periods for 12 months or more before examination for surgical, natural,
or other nonhormonal causes. Because of a possible association between unopposed estrogen exposure and Raynaud’s,23 the combined effect of menopausal status and hormone exposure was examined using a categoric variable (premenopausal; postmenopausal on current unopposed estrogen; postmenopausal on combination hormone therapy; postmenopausal on no current hormone therapy).

Statistical Analysis
Because of the discordant results of Raynaud’s risk-factor studies in women versus men,1,18,31 all analyses were sex-specific. Bivariate analyses of Raynaud’s by smoking status, alcohol consumption, and covariates (age, BMI, current use of vasodilating or beta-blocker medications, cardiovascular disease, and, in women, menopausal and hormone status) were performed using Cochran-Armitage for trend, Student t, chi-square, and Fisher’s exact tests, as appropriate. We used 3-way tables to look for interactions among smoking, alcohol, and Raynaud’s.

Multivariable analyses were performed using logistic regression analysis and adjusting for statistically significant (P < .1) and clinically relevant covariates. (All other analyses used P < .05 for significance.) Smoking and alcohol consumption were coded as dummy variables (“never smoker” and “no alcohol consumption” as referents). Possible interactions between smoking and alcohol consumption included that alcohol might potentiate the effects of smoking on Raynaud’s or attenuate these effects through vasodilation or other mechanisms. To determine the impact of cardiovascular disease on the relationship among smoking, alcohol, and Raynaud’s, we repeated the analyses excluding the cardiovascular disease term and performed subgroup analyses by presence of prevalent cardiovascular disease. All analyses were performed using the SAS statistical package, version 8.02 (SAS Institute, Cary, NC).

RESULTS
Population Characteristics
In our study of 1840 women and 1602 men, the prevalence of Raynaud’s was 6.1% (n = 113) in women and 4.2% (n = 68) in men. More than 36% of women and 45% of men with Raynaud’s in our study cohort recalled an age of onset after 50 years. Table 1 lists participants’ baseline characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Women N = 1840 (%)</th>
<th>Men N = 1602 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (mean ± SD)</td>
<td>61.4 ± 9.2</td>
<td>62.5 ± 9.4</td>
</tr>
<tr>
<td>BMI (mean ± SD)</td>
<td>27.7 ± 5.6</td>
<td>28.6 ± 4.3</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>708 (38.5)</td>
<td>466 (29.1)</td>
</tr>
<tr>
<td>Past</td>
<td>884 (48.0)</td>
<td>938 (58.6)</td>
</tr>
<tr>
<td>Current</td>
<td>248 (13.5)</td>
<td>198 (12.5)</td>
</tr>
<tr>
<td>Smoked more than 1 pack per day</td>
<td>158/248 (63.7)</td>
<td>142/198 (71.7)</td>
</tr>
<tr>
<td>Alcohol consumption per week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>737 (40.0)</td>
<td>503 (31.4)</td>
</tr>
<tr>
<td>Moderate†</td>
<td>820 (44.6)</td>
<td>868 (54.2)</td>
</tr>
<tr>
<td>Heavy‡</td>
<td>283 (15.4)</td>
<td>231 (14.4)</td>
</tr>
<tr>
<td>Alcohol consumption by type§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any red wine</td>
<td>506 (27.5)</td>
<td>570 (35.6)</td>
</tr>
<tr>
<td>Any white wine</td>
<td>716 (38.9)</td>
<td>428 (26.7)</td>
</tr>
<tr>
<td>Any beer</td>
<td>263 (14.3)</td>
<td>744 (46.4)</td>
</tr>
<tr>
<td>Any liquor</td>
<td>445 (24.2)</td>
<td>554 (34.6)</td>
</tr>
<tr>
<td>Antihypertensive medication use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current vasodilator use</td>
<td>390 (21.2)</td>
<td>572 (35.7)</td>
</tr>
<tr>
<td>Current beta-blocker use</td>
<td>311 (16.9)</td>
<td>448 (28.0)</td>
</tr>
<tr>
<td>History of cardiovascular disease</td>
<td>253 (13.8)</td>
<td>516 (32.3)</td>
</tr>
<tr>
<td>Menopausal and hormone status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td>336 (18.3)</td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and no hormone therapy</td>
<td>993 (54.2)</td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and unopposed estrogen</td>
<td>241 (13.2)</td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and combination therapy</td>
<td>263 (14.4)</td>
<td>-</td>
</tr>
</tbody>
</table>

SD = standard deviation; BMI = body mass index.
*All P values were less than .05 for the comparison between women and men.
†For women: 1 oz up to 3.5 oz per week. For men: 1 oz up to 7 oz per week.
‡For women: >3.5 oz per week. For men: >7 oz per week.
§Any: ≥12 oz of beer, 1 oz of liquor or 4 oz of either type of wine per week.
||Data missing for 7 women.
We found no relationship between smoking and Raynaud’s in women (Table 2). There were no current women smokers with both cardiovascular disease and Raynaud’s; 17.9% (5/28) of women with both Raynaud’s and cardiovascular disease noted that their symptoms limited their daily activities versus only 9.4% (8/85) of women with Raynaud’s but no cardiovascular disease ($P = 0.01$).

In men, 2.6% of never smokers had Raynaud’s versus 4.6% of past smokers and 6.6% of current smokers. Also, 5.4% of men who currently smoked 11 to 20 cigarettes per day had Raynaud’s versus 16.0% of men who currently smoked more than 20 cigarettes per day ($P$ for trend $< 0.001$).

### Table 2 Proportion* of Prevalent Raynaud’s Phenomenon or Mean Values Among Those With and Without Raynaud’s Phenomenon by Covariates, Including Smoking Status and Alcohol Consumption, in Women and Men†

<table>
<thead>
<tr>
<th>Variable</th>
<th>Women n/N (%)</th>
<th>P value</th>
<th>Men n/N (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (mean ± SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raynaud’s phenomenon</td>
<td>62.5 ± 8.7</td>
<td>.2</td>
<td>64.5 ± 7.1</td>
<td>.05</td>
</tr>
<tr>
<td>No Raynaud’s phenomenon</td>
<td>61.4 ± 9.2</td>
<td></td>
<td>62.4 ± 9.5</td>
<td></td>
</tr>
<tr>
<td>Body mass index (mean ± SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raynaud’s phenomenon</td>
<td>25.6 ± 4.3</td>
<td>&lt;.001</td>
<td>26.4 ± 3.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No Raynaud’s phenomenon</td>
<td>27.8 ± 5.7</td>
<td></td>
<td>28.7 ± 4.3</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never§</td>
<td>40/708 (5.7)</td>
<td>.9</td>
<td>12/466 (2.6)</td>
<td>.01</td>
</tr>
<tr>
<td>Past</td>
<td>60/884 (6.8)</td>
<td>43/938 (4.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>13/248 (5.2)</td>
<td>13/198 (6.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol consumption per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>43/737 (5.8)</td>
<td>.08</td>
<td>31/503 (6.2)</td>
<td>.2</td>
</tr>
<tr>
<td>Moderate</td>
<td>42/820 (5.1)</td>
<td></td>
<td>24/868 (2.8)</td>
<td></td>
</tr>
<tr>
<td>Heavy</td>
<td>28/283 (9.9)</td>
<td></td>
<td>13/231 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Alcohol consumption by type¶</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red wine consumption per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>92/1334 (6.9)</td>
<td>.03</td>
<td>59/1032 (5.7)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Any</td>
<td>21/506 (4.2)</td>
<td></td>
<td>9/570 (1.6)</td>
<td></td>
</tr>
<tr>
<td>White wine consumption per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>59/1124 (5.3)</td>
<td>.05</td>
<td>54/1174 (4.6)</td>
<td>.2</td>
</tr>
<tr>
<td>Any</td>
<td>54/716 (7.5)</td>
<td></td>
<td>14/428 (3.3)</td>
<td></td>
</tr>
<tr>
<td>Beer consumption per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>90/1577 (5.7)</td>
<td>.06</td>
<td>39/858 (4.6)</td>
<td>.5</td>
</tr>
<tr>
<td>Any</td>
<td>23/263 (8.8)</td>
<td></td>
<td>29/744 (3.9)</td>
<td></td>
</tr>
<tr>
<td>Liquor consumption per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>79/1395 (5.7)</td>
<td>.1</td>
<td>47/1048 (4.5)</td>
<td>.5</td>
</tr>
<tr>
<td>Any</td>
<td>34/445 (7.6)</td>
<td></td>
<td>21/554 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Current vasodilator use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>32/390 (8.2)</td>
<td>.06</td>
<td>27/572 (4.7)</td>
<td>.5</td>
</tr>
<tr>
<td>No</td>
<td>81/1450 (5.6)</td>
<td></td>
<td>41/1030 (4.0)</td>
<td></td>
</tr>
<tr>
<td>Current beta-blocker use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25/311 (8.0)</td>
<td>.1</td>
<td>23/448 (5.1)</td>
<td>.3</td>
</tr>
<tr>
<td>No</td>
<td>88/1529 (5.8)</td>
<td></td>
<td>45/1154 (3.9)</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>28/253 (11.1)</td>
<td>&lt;.001</td>
<td>30/516 (5.8)</td>
<td>.03</td>
</tr>
<tr>
<td>Absent</td>
<td>85/1587 (5.4)</td>
<td></td>
<td>38/1086 (3.5)</td>
<td></td>
</tr>
<tr>
<td>Menopausal and hormone status**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td>22/336 (6.6)</td>
<td>.7</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and no hormone therapy</td>
<td>63/993 (6.3)</td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and unopposed estrogen</td>
<td>14/241 (5.8)</td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Postmenopausal and combination therapy</td>
<td>14/263 (5.3)</td>
<td></td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

SD = standard deviation.

*For example, n = 40 of N = 708 women who were never smokers (5.7%) had prevalent Raynaud’s phenomenon.
†All P values obtained using Cochran-Armitage for trend test.
‡Unless otherwise noted.
§When smoking was examined as a dichotomous (ever/never) variable, there was no change in these findings in either gender.
¶None: <1 oz per week for women and men. Moderate: 1 oz up to 3.5 oz per week for women and 1 oz up to 7 oz per week for men. Heavy: >3.5 oz per week for women and >7 oz per week for men.
¶¶Any: ≥12 oz of beer, 1 oz of liquor, or 4 oz of either type of wine per week.
**Data missing for 7 women.

### Bivariate Analyses

**Smoking.** We found no relationship between smoking and Raynaud’s in women (Table 2). There were no current women smokers with both cardiovascular disease and Raynaud’s; 17.9% (5/28) of women with both Raynaud’s and cardiovascular disease noted that their symptoms limited their daily activities versus only 9.4% (8/85) of women with Raynaud’s but no cardiovascular disease ($P = 0.01$).

In men, 2.6% of never smokers had Raynaud’s versus 4.6% of past smokers and 6.6% of current smokers. Also, 5.4% of men who currently smoked 11 to 20 cigarettes per day had Raynaud’s versus 16.0% of men who currently smoked more than 20 cigarettes per day ($P$ for trend $< 0.001$).
Among moderate male drinkers, smoking remained associated with increased Raynaud’s prevalence (9.9%) compared with nondrinkers and moderate drinkers (5.8% and 5.1%, respectively), but without a clear dose response (10.7% Raynaud’s prevalence in women who drank >3.5 but ≤7 oz per week vs 8.2% in those who drank >7 oz per week) (Table 2). White wine, beer, and liquor consumption were associated with higher Raynaud’s prevalence, but not significantly. Women who drank red wine had a lower prevalence of Raynaud’s than those who did not. This association did not seem to be due to a dose-response relationship, because women who drank red wine were actually more likely to be heavy alcohol consumers than non-red wine drinkers (30.6% vs 21.4%, P < .001).

Interaction Among Smoking, Alcohol, and Raynaud’s.

Regardless of smoking status, women who consumed heavy amounts of alcohol had a higher prevalence of Raynaud’s than women who abstained or drank moderately (Table 3). Further, women who were heavy drinkers and current smokers did not show increased Raynaud’s prevalence compared with heavy-drinking nonsmokers.

Among moderate male drinkers, smoking remained associated with increasing Raynaud’s prevalence (P for trend < .001). There were no cases of Raynaud’s among the 55 men who smoked 1 to 10 cigarettes per day, limiting our ability to draw conclusions regarding threshold effects.

**Alcohol.** Heavy alcohol consumption in women was associated with increased Raynaud’s prevalence (9.9%) compared with nondrinkers and moderate drinkers (5.8% and 5.1%, respectively), but without a clear dose response (10.7% Raynaud’s prevalence in women who drank >3.5 but ≤7 oz per week vs 8.2% in those who drank >7 oz per week) (Table 2). White wine, beer, and liquor consumption were associated with higher Raynaud’s prevalence, but not significantly. Women who drank red wine had a lower prevalence of Raynaud’s than those who did not. This association did not seem to be due to a dose-response relationship, because women who drank red wine were actually more likely to be heavy alcohol consumers than non-red wine drinkers (30.6% vs 21.4%, P < .001).

Examining alcohol consumption in men demonstrated a V-shaped relationship between Raynaud’s prevalence and alcohol consumption. Light-moderate (≥1 but ≤3.5 oz per week) and heavy-moderate (>3.5 but ≤7 oz per week) male drinkers had similar rates of Raynaud’s (3.0% vs 2.4%, respectively). Men who drank red wine, most commonly moderate drinkers, had a lower prevalence of Raynaud’s than those who did not, whereas men who consumed other types of alcohol did not have a significantly different Raynaud’s prevalence than their nondrinking counterparts.

**Multivariable Analyses**

**Women.** Because there was no association between smoking and Raynaud’s in women, multivariable analyses were only performed to evaluate the relationship between alcohol consumption and Raynaud’s (Table 4). In a logistic regression model adjusting for BMI, the use of vasodilators, and history of cardiovascular disease, heavy alcohol consumption was independently associated with increased odds of Raynaud’s (adjusted odds ratio [OR] 1.69, 95% confidence interval [CI], 1.02-2.82; nondrinkers as referent). After adjustment for BMI, use of vasodilator medications, and history of cardiovascular disease, consumption of red wine

**Table 3** Proportion* of Prevalent Raynaud’s Phenomenon by Smoking Status and Level of Alcohol Consumption in Women and Men†

<table>
<thead>
<tr>
<th>Smoking</th>
<th>None n/N (%)</th>
<th>Moderate n/N (%)</th>
<th>Heavy n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>18/341 (5.3)</td>
<td>16/303 (5.3)</td>
<td>6/64 (9.4)</td>
</tr>
<tr>
<td>Past</td>
<td>21/298 (7.1)</td>
<td>22/427 (5.2)</td>
<td>17/159 (10.7)</td>
</tr>
<tr>
<td>Current</td>
<td>4/98 (4.1)</td>
<td>4/90 (4.4)</td>
<td>5/60 (8.3)</td>
</tr>
<tr>
<td>Men†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>4/126 (3.2)</td>
<td>6/305 (2.0)</td>
<td>2/35 (5.7)</td>
</tr>
<tr>
<td>Moderate</td>
<td>26/319 (8.2)</td>
<td>11/478 (2.3)</td>
<td>6/141 (4.3)</td>
</tr>
<tr>
<td>Heavy</td>
<td>1/58 (1.7)</td>
<td>7/85 (8.2)</td>
<td>5/55 (9.1)</td>
</tr>
</tbody>
</table>

*For example, n = 18 (5.3%) of N = 341 female never smoker, nondrinkers had prevalent Raynaud’s phenomenon.
†None of the comparisons were statistically significant (P < .05) in women. In men, only the comparison of Raynaud’s phenomenon prevalence by smoking status among moderate drinkers ("Moderate" column) demonstrated a significant trend of increasing Raynaud’s phenomenon prevalence by increasing smoking exposure (P = .005).‡None: <1 oz per week. Moderate: 1 oz up to 3.5 oz per week. Heavy: >3.5 oz per week.
§None: <1 oz per week. Moderate: 1 oz up to 7 oz per week. Heavy: >7 oz per week.

**Table 4** Multivariable Logistic Regression Analysis of Raynaud’s Phenomenon and Alcohol Consumption in Women and Smoking Status and Alcohol Consumption in Men

<table>
<thead>
<tr>
<th>Variable</th>
<th>Adjusted OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy alcohol consumption</td>
<td>1.69</td>
<td>1.02-2.82</td>
</tr>
<tr>
<td>Moderate alcohol consumption</td>
<td>0.89</td>
<td>0.57-1.40</td>
</tr>
<tr>
<td>Men†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>2.59</td>
<td>1.11-6.04</td>
</tr>
<tr>
<td>Past smoker</td>
<td>1.69</td>
<td>0.86-3.30</td>
</tr>
<tr>
<td>Heavy alcohol consumption</td>
<td>0.85</td>
<td>0.42-1.70</td>
</tr>
<tr>
<td>Moderate alcohol consumption</td>
<td>0.51</td>
<td>0.29-0.89</td>
</tr>
</tbody>
</table>

OR = odds ratio; CI = confidence interval.

*Adjusting for BMI, current use of vasodilator medications, and presence of cardiovascular disease with nondrinkers as referent group.
†Adjusting for age in years, BMI, and presence of cardiovascular disease with never smokers and nondrinkers as referent groups.
remained independently associated with reduced odds of Raynaud’s (adjusted OR 0.59, 95% CI, 0.36-0.96; no red wine as referent).

**Men.** In a multivariable logistic regression model adjusting for age, BMI, and history of cardiovascular disease, current smoking status remained independently associated with prevalent Raynaud’s (adjusted OR 2.59, 95% CI, 1.11-6.04; never smokers as referent) (Table 4). Neither past smoking exposure nor heavy alcohol consumption was significantly associated with increased Raynaud’s prevalence, but moderate alcohol consumption retained its protective effect with an adjusted OR of 0.51 (95% CI, 0.29-0.89; nondrinkers as referent). Including an interaction term did not alter these findings. Examining red wine consumption and smoking in a model adjusting for age, BMI, and history of cardiovascular disease produced an adjusted OR for Raynaud’s among red wine drinkers of 0.30 (95% CI, 0.15-0.62) and for current smoking of 2.71 (95% CI, 1.17-6.23).

**Effects of Cardiovascular Disease.** Because of associations of smoking and alcohol consumption with cardiovascular disease and the possibility that these interactions might confound our analysis, we performed analyses with and without adjustment for cardiovascular disease. We also performed subgroup analyses among those with and without cardiovascular disease. Removing cardiovascular disease from the model in women (ie, adjusting for only BMI and vasodilator use) decreased the adjusted OR for Raynaud’s among heavy drinkers to 1.56 (95% CI, 0.94-2.59; nondrinkers as referent). In women without cardiovascular disease, the adjusted OR for Raynaud’s among heavy drinkers was 1.88 (95% CI, 1.06-3.33), whereas the adjusted odds of Raynaud’s in heavy drinkers with cardiovascular disease was 1.15 (95% CI, 0.34-3.98). Removing cardiovascular disease from the model in men (ie, adjusting for only age and BMI) had little impact on Raynaud’s risk (adjusted OR for current smoking 2.69, 95% CI, 1.16-6.27; never smokers as referent; adjusted OR for moderate alcohol consumption 0.49, 95% CI, 0.28-0.85; nondrinkers as referent). However, in men without cardiovascular disease, there were no significant associations between smoking or moderate alcohol consumption and Raynaud’s, whereas the adjusted OR for Raynaud’s in smokers and moderate alcohol consumers with cardiovascular disease was 17.17 (95% CI, 1.92-153.30) and 0.39 (95% CI, 0.16-0.91), respectively.

**DISCUSSION**

We found no association between smoking and Raynaud’s in women, whereas in men, current smoking was associated with an increased odds of Raynaud’s. Among women, heavy alcohol consumption was associated with increased Raynaud’s prevalence. Red wine consumption was associated with reduced Raynaud’s prevalence, and moderate consumption was not associated with Raynaud’s. Among men, heavy alcohol consumption was not associated with prevalent Raynaud’s, whereas moderate alcohol and red wine consumption were associated with reduced odds of Raynaud’s.

The lack of association of Raynaud’s with smoking in women is consistent with the majority of prior reports and only differs from a study of ethnic Slavs with both primary and secondary Raynaud’s. Similarly, our finding that smoking was associated with an increased Raynaud’s prevalence in men is consistent with previously published reports that used validated criteria for Raynaud’s classification and gender-specific analyses, and examined community populations, but differs from published data that included individuals with secondary causes of Raynaud’s. Although causal statements cannot be made because of the cross-sectional nature of our study, our findings support the possibility that smoking has distinct vascular effects in women versus men.

The association of smoking and Raynaud’s in men correlated strongly to the presence of cardiovascular disease. Although increased Raynaud’s prevalence was also associated with the presence of cardiovascular disease in women, the absence of any current women smokers with both cardiovascular disease and Raynaud’s suggests that women with both might be more likely than men to quit smoking.

The association of heavy alcohol consumption with increased odds of Raynaud’s in women, but not men, contrasts with earlier reports. This is likely due to our large sample, adjustment for confounders, and use of validated classification criteria. Although women may drink to relieve their symptoms, our prior work suggests that Raynaud’s is relatively mild in this cohort, and thus alternative explanations, such as decreased use of protective behaviors and altered neurovascular reflexes, are more likely. Our results in men confirm those of Palesch et al, demonstrating a V-shaped relationship of alcohol consumption to Raynaud’s prevalence, with moderate (or light-moderate) consumption associated with decreased Raynaud’s prevalence.

The association of red wine consumption with reduced odds of Raynaud’s is consistent with prior reports linking red wine consumption to reductions in cardiovascular disease and may indicate that Raynaud’s represents a common physiologic pathway with cardiovascular disease in this age group. Our results indicate that smoking may attenuate the beneficial effect of moderate alcohol consumption on Raynaud’s prevalence in men and that the protective effect of moderate alcohol and/or red wine intake on vascular disease may also play a role in the expression of Raynaud’s in middle-aged men.

Several strengths of our study merit comment. Our study sample was community-based and unselected for factors known to be associated with Raynaud’s, a validated instrument was used to ascertain the presence of Raynaud’s, and adjustment was made for a range of relevant confounders.
Conversely, our study has important limitations. In addition to the limitations of a cross-sectional analysis, antinuclear antibody results were not available; however, a meta-analysis of 639 patients with primary Raynaud’s found that antinuclear antibodies had only a 30% positive predictive value for development of a connective tissue disease, and it is unlikely that our cohort included many individuals with unrecognized connective tissue disease. Because of the scarcity of women in our study, our study may have been underpowered to find a significant relationship between smoking and Raynaud’s in women, but our data are consistent with published literature. Also, our failure to find an association between smoking and Raynaud’s in women might have been because of a gender difference in cumulative smoking exposure; however, there were no significant gender differences in the number of cigarettes smoked per day, and initiation of smoking in later life is rare in this cohort. Despite our large sample size, we were limited in our ability to perform subgroup analyses and, thus, to look for further effect, modification from cardiovascular disease, or fully ascertain the interaction of smoking and alcohol on Raynaud’s prevalence. Finally, the generalizability of our results is limited because our study cohort was white and predominantly middle-aged.

CONCLUSION

There was no association between smoking and Raynaud’s in women. Heavy alcohol consumption was associated with increased Raynaud’s in women, particularly among women without cardiovascular disease. Among men with cardiovascular disease, smoking was independently associated with an increase in Raynaud’s prevalence. Moderate alcohol consumption was associated with a reduction in Raynaud’s prevalence in men, but smoking may attenuate this effect. Red wine consumption seemed to be protective against Raynaud’s in both women and men. Our data suggest that middle-aged women and men have distinct physiologic mechanisms underlying their Raynaud’s and, thus, may need sex-specific therapeutic approaches. Our data also support the possibility that moderate red wine consumption may protect against Raynaud’s and offer a focus for future Raynaud’s research.

References


CLINICAL RESEARCH STUDY

Cost-effectiveness of Treatment for Hepatitis C in an Urban Cohort Co-infected with HIV

Nicole G. Campos, MSc,a Joshua A. Salomon, PhD,b Julie C. Servoss, MD, MPH,c David P. Nunes, MD,d Jeffrey H. Samet, MD, MA, MPH,e Kenneth A. Freedberg, MD, MSc,f,g Sue J. Goldie, MD, MPHg

aProgram in Health Policy, Harvard University, Cambridge, Mass; bDepartment of Population and International Health, Harvard School of Public Health, Harvard University Initiative for Global Health, Cambridge, Mass; cGastrointestinal Unit, Massachusetts General Hospital, Harvard Medical School, Boston; dSection of Gastroenterology, and eSection of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, Boston, Mass; fDivisions of General Medicine and Infectious Diseases and the Harvard Center for AIDS Research (CFAR), Massachusetts General Hospital, Harvard Medical School, Boston; and gProgram in Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, Mass.

ABSTRACT

PURPOSE: Recent clinical trials have evaluated treatment strategies for chronic infection with hepatitis C virus (HCV) in patients co-infected with human immunodeficiency virus (HIV). Our objective was to use these data to examine the cost-effectiveness of treating HCV in an urban cohort of co-infected patients.

METHODS: A computer-based model, together with available published data, was used to estimate lifetime costs (2004 US dollars), life expectancy, and incremental cost per year of life saved (YLS) associated with 3 treatment strategies: (1) interferon-alfa and ribavirin; (2) pegylated interferon-alfa; and (3) pegylated interferon-alfa and ribavirin. The target population included treatment-eligible patients, based on an actual urban cohort of HIV-HCV co-infected subjects, with a mean age of 44 years, of whom 66% had genotype 1 HCV, 16% had cirrhosis, and 98% had CD4 cell counts >200 cells/mm³.

RESULTS: Pegylated interferon-alfa and ribavirin was consistently more effective and cost-effective than other treatment strategies, particularly in patients with non-genotype 1 HCV. For patients with CD4 counts between 200 and 500 cells/mm³, survival benefits ranged from 5 to 11 months, and incremental cost-effectiveness ratios were consistently less than $75,000 per YLS for men and women of both genotypes. Due to better treatment efficacy in non-genotype 1 HCV patients, this group experienced greater life expectancy gains and lower incremental cost-effectiveness ratios.

CONCLUSIONS: Combination therapy with pegylated interferon-alfa and ribavirin for HCV in eligible co-infected patients with stable HIV disease provides substantial life-expectancy benefits and appears to be cost-effective. Overcoming barriers to HCV treatment eligibility among urban co-infected patients remains a critical priority. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Hepatitis C virus (HCV); Human immunodeficiency virus (HIV); Cost-effectiveness; Peginterferon-alfa and ribavirin; Clinical guidelines; Treatment eligibility

Among the estimated 950,000 persons infected with human immunodeficiency virus (HIV) in the United States, approximately 30% are co-infected with the hepatitis C virus (HCV).1,2 While highly active antiretroviral therapy (HAART) has essentially transformed HIV to a chronic disease, co-infected patients are increasingly vulnerable to complications of chronic liver disease, including cirrhosis and liver failure. Compared with HCV mono-infected patients, they tend to have higher levels of HCV RNA and
to progress more rapidly to cirrhosis and end-stage liver disease.\textsuperscript{3} Mortality attributable to end-stage liver disease has steadily increased since 1996, and in some HIV patient populations it is now the leading cause of death.\textsuperscript{4} The impact of HCV on HIV progression is more controversial.\textsuperscript{5-7}

In clinical trials among patients with HCV mono-infection, combination therapy with pegylated interferon-alfa and ribavirin has produced sustained virologic response rates ranging from 54\% to 63\%.\textsuperscript{8-10} Recently, 4 randomized controlled trials evaluated combination therapy with pegylated interferon-alfa and ribavirin compared with interferon-alfa and ribavirin in patients with HIV-HCV co-infection.\textsuperscript{11-14} The largest of these trials, the AIDS Pegasys Ribavirin International Co-infection Trial (APRICOT), was conducted at 95 centers in 19 countries with 868 subjects and yielded a sustained virologic response rate of 40\%.\textsuperscript{11} Based on the APRICOT findings, in February 2005 the United States (US) Food and Drug Administration approved pegylated interferon-alfa-2a and ribavirin for the treatment of HCV in patients with HIV.\textsuperscript{15}

To the best of our knowledge, only one cost-effectiveness analysis evaluating the treatment of HCV in HIV-HCV co-infected patients has been published. Kuehne et al demonstrated that combination therapy for histologically moderate HCV in co-infected patients resulted in an increase in quality-adjusted life expectancy while incurring costs comparable with other well-accepted clinical interventions.\textsuperscript{16} However, this analysis was performed before randomized controlled trials had established approximate treatment efficacy rates in HIV-HCV co-infected patients, and the APRICOT trial rates of sustained virologic response were generally lower than the lower bounds of the sensitivity analysis performed by Kuehne and colleagues. Since this prior cost-effectiveness analysis was conducted, considerable progress has been made in discerning treatment efficacy rates and relative risk estimates for progression of liver disease in HIV-HCV co-infected patients.

Our objective was to use recent prospective data regarding eligibility for interferon-based treatment, the impact of HIV on the progression of HCV-related liver disease, and demonstrated treatment efficacy from clinical trials to consider the potential health benefits, economic costs, and cost-effectiveness of treatment for HCV among an urban cohort of co-infected patients with stable HIV disease.

**METHODS**

**Overview**

We modified an existing Markov model of HCV\textsuperscript{17} to reflect co-infection with HIV and examined the cost-effectiveness of the following strategies for HCV treatment in the treatment-eligible segment of an urban co-infected cohort\textsuperscript{18}: combination therapy with interferon-alfa-2a and ribavirin; monotherapy with pegylated interferon-alfa-2a; and combination therapy with pegylated interferon-alfa-2a and ribavirin.

Population characteristics (mean age, Metavir score distribution, and mean CD4 cell count) for the modeled cohort were derived from a subgroup of the Hepatitis and AIDS Liver Outcomes (HALO) Study cohort that was co-infected with HIV and HCV, and eligible for treatment (Table 1).\textsuperscript{18}

We followed the recommendations of the US Panel on Cost-Effectiveness in Health and Medicine,\textsuperscript{19} adopting a societal perspective (although we excluded patient time costs) and discounting all costs and clinical consequences at a rate of 3\% per year. The comparative efficiencies of alternative treatment strategies were measured by the incremental cost-effectiveness ratio, defined as the additional cost of a specific treatment strategy divided by its additional health benefit, expressed here as

![Table 1](image-url)
years of life saved (YLS). The incremental ratio for a strategy was computed in comparison with the next most effective option after eliminating strategies that were dominated (more costly and less effective than other options) or ruled out by extended (weak) dominance (strategies with higher incremental cost-effectiveness ratios than more effective options). We conducted sensitivity analyses to assess the influence of varying uncertain parameters and adopting alternative assumptions on our results.

**Model**

A deterministic state-transition Markov model (DATA 4.0; TreeAge Software Inc., Williamstown, Mass) was used to simulate the natural history of HCV infection in patients co-infected with HIV. Early stages of liver disease were classified using the Metavir scoring system, which characterizes the extent of fibrosis that results as damaged liver cells are repaired, including no fibrosis (F0), portal fibrosis without septa (F1), portal fibrosis with few septa (F2), numerous septa without cirrhosis (F3), and cirrhosis (F4). Advanced stages of liver disease were defined clinically as compensated cirrhosis, decompensated cirrhosis (including separate states for ascites, variceal hemorrhage and hepatic encephalopathy), and primary hepatocellular carcinoma (Figure 1). Monthly transition probabilities were derived from the literature and allowed individuals to move through different health states over time.

We made the following assumptions:

- HCV infection may resolve through successful treatment, implying clearance of HCV RNA
- Patients without a sustained viral response to HCV treatment (as defined by an HCV RNA level of ≤50 IU per milliliter 24 weeks after completion of therapy) received no clinical benefit and were subject to pretreatment rates of HCV-related liver disease progression
- Patients in early stages of liver disease who experienced a sustained viral response were no longer at risk for HCV-related liver disease
- Patients with cirrhosis received treatment, and if it was effective they regressed to Metavir stage F3; if treatment was ineffective they were still at risk for decompensated cirrhosis
- Response to treatment was conditional on genotype
- The rate of fibrosis progression in the absence of effective treatment was conditional on age and sex, and remained the same for transition to the next higher Metavir stage
- All patients were assumed to have a stable CD4 cell count between 200 and 500 cells/mm³ and to be receiving HAART for HIV; and
• Patients with decompensated cirrhosis were eligible for liver transplantation.24,25

data
Selected parameters used in the model are shown in Table 2. We used baseline age- and sex-specific rates of progression from chronic HCV infection to cirrhosis based on an empirically calibrated model of chronic HCV in mono-infected patients.17,26 We modified these based on data from studies comparing the relative progression in HIV co-infected patients versus mono-infected patients.27-30 We assumed that progression from cirrhosis to decompensated cirrhosis was similar in co-infected patients and mono-infected patients, and used rates derived from a cost-effectiveness study in mono-infected patients.31 Excess mortality due to HIV was based on data from the Multicenter AIDS Cohort Study, from which CD4-specific rates were derived to parameterize a natural history model of HIV/AIDS published by Freedberg et al32; an additive relationship to age- and sex-standardized mortality rates was assumed. We then compared the model’s predictions of cirrhosis prevalence with a published study by Di Martino et al that was not used for natural history parameter estimation.33 Among an HIV-HCV co-infected cohort of former intravenous drug users with an average 10.6-year duration of HCV infection and a mean CD4 count of 482 cells/mm³, 59% of whom had received interferon monotherapy (overall efficacy of 6.4%) and 73% of whom were male, 8.75% had cirrhosis at baseline and 17.5% had cirrhosis at follow-up 4.7 years later.33 Our model predicted that in a cohort of patients with this sex and treatment profile and duration of HCV infection, 16% had cirrhosis over the same follow-up period.

Annual costs of care related to chronic HCV infection and liver disease included detailed estimates of resource utilization, including hospitalizations, outpatient visits, laboratory tests, medications, and interventions.31 Treatment costs were based on average wholesale drug prices34 combined with previously published cost estimates for clinic visits, laboratory tests, and the treatment of adverse events.35 The annual costs of HIV care for patients on HAART with CD4 count between 200 and 500 cells/mm³ were obtained from a recently published model for HIV screening.36,37

In the base case we assumed treatment for 48 weeks,11-13 and dosages resembled those in APRICOT: for interferon and ribavirin, 3 million IU interferon alfa-2a subcutaneously 3 times/week plus 800 mg ribavirin/day; for pegylated interferon, 180 μg of pegylated interferon alfa-2a subcutaneously weekly; and for pegylated interferon and ribavirin, pegylated interferon as described above plus 800 mg ribavirin/day. We made the conservative assumption that all patients completed the full course of medication. In sensitivity analyses, we explored a second treatment protocol and assumed that patients without an early virologic response at week 12 discontinued treatment. The percentage of such patients was drawn from APRICOT and varied by treatment arm. We assessed early treatment withdrawal due to adverse events or abnormal laboratory values in a sensitivity analysis.

Results
Table 3 shows the discounted lifetime costs, life expectancy, and incremental cost per YLS for each treatment strategy, stratified by sex and genotype. For the men in our modeled cohort, the average discounted life expectancy without treatment was 11.6 years and lifetime costs were $240,300. In men with genotype 1, treatment for HCV provided incremental gains ranging from 1.0 to 5.2 months compared with no therapy. Combination therapy with pegylated interferon and ribavirin dominated all other strategies because it was both more effective and had a lower (more attractive) cost-effectiveness ratio. Compared with no therapy, its incremental cost-effectiveness ratio was $73,000 per YLS. For men with non-genotype 1 HCV, treatment provided incremental gains ranging from 3.0 to 10.7 months compared with no therapy. Again, combination therapy with pegylated interferon and ribavirin was the dominant strategy. Compared with no therapy, its incremental cost-effectiveness ratio was $39,700 per YLS. Results in women were very similar.

In an exploratory analysis assuming cessation of treatment when no early virologic response occurs, the rank ordering of strategies remained the same. The incremental cost-effectiveness ratio for combination therapy with pegylated interferon and ribavirin became more attractive—lower by 19% in men with genotype 1 ($59,300 per YLS) and by 17% in men with non-genotype 1 ($33,100 per YLS). Again, results in women were similar.

Results were relatively insensitive to varying parameters across plausible ranges for treatment efficacies among patients with cirrhosis and treatment risks. Results were most sensitive to variation in the annual excess death rate due to HIV, fibrosis progression rates and treatment efficacies in noncirrhotic patients. Results were moderately sensitive to drug costs. For men with genotype 1 infection, when the excess mortality due to HIV was reduced by 97%, the incremental cost-effectiveness ratio decreased to $41,400 (base case: $73,000 per YLS). When excess mortality was increased 11-fold to represent death rates in patients with a history of severe opportunistic infections, the incremental cost-effectiveness ratio increased so greatly that treatment was dominated by nontreatment.

We conducted a 2-way sensitivity analysis in which we varied the effectiveness of combination therapy with pegylated interferon and ribavirin, and the relative risk of fibrosis progression due to co-infection with HIV (Figure 2). When treatment efficacy exceeded 50%, cost-effectiveness ratios were consistently less than $50,000 per YLS, regardless of the relative risk of fibrosis progression. When efficacy was >25%, ratios were consistently <$100,000 per YLS across the entire plausible range of relative risk assumptions. However, when efficacy was <25%, the relative risk of progres-
Table 2  Base Case Values for Model Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Base Case</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HCV natural history parameters</strong>&lt;sup&gt;*&lt;/sup&gt;&lt;sup&gt;,17,26-29,31&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Fibrosis progression in men, age (years)&lt;sup&gt;†&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.149</td>
</tr>
<tr>
<td>50</td>
<td>0.159</td>
</tr>
<tr>
<td>60</td>
<td>0.298</td>
</tr>
<tr>
<td>≥70</td>
<td>0.406</td>
</tr>
<tr>
<td>Fibrosis progression in women, age (years)&lt;sup&gt;†&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.108</td>
</tr>
<tr>
<td>50</td>
<td>0.135</td>
</tr>
<tr>
<td>60</td>
<td>0.208</td>
</tr>
<tr>
<td>≥70</td>
<td>0.284</td>
</tr>
<tr>
<td>Cirrhosis to ascites</td>
<td>0.025</td>
</tr>
<tr>
<td>Cirrhosis to hepatic encephalopathy</td>
<td>0.004</td>
</tr>
<tr>
<td>Cirrhosis to hepatocellular carcinoma</td>
<td>0.015</td>
</tr>
<tr>
<td>Cirrhosis to variceal hemorrhage</td>
<td>0.011</td>
</tr>
<tr>
<td>Liver transplant probability</td>
<td>0.031</td>
</tr>
<tr>
<td>Mortality rate per person</td>
<td></td>
</tr>
<tr>
<td>Ascites</td>
<td>0.110</td>
</tr>
<tr>
<td>Hepatic encephalopathy (first year)</td>
<td>0.680</td>
</tr>
<tr>
<td>Hepatic encephalopathy (subsequent years)</td>
<td>0.400</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>0.433</td>
</tr>
<tr>
<td>Transplant (first year)</td>
<td>0.210</td>
</tr>
<tr>
<td>Transplant (subsequent years)</td>
<td>0.057</td>
</tr>
<tr>
<td>Variceal hemorrhage (first year)</td>
<td>0.400</td>
</tr>
<tr>
<td>Variceal hemorrhage (subsequent years)</td>
<td>0.130</td>
</tr>
<tr>
<td><strong>Treatment parameters</strong>&lt;sup&gt;11&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Treatment response probability</td>
<td></td>
</tr>
<tr>
<td>Interferon-alfa + ribavirin</td>
<td></td>
</tr>
<tr>
<td>Genotype 1</td>
<td>0.07</td>
</tr>
<tr>
<td>Non-genotype 1</td>
<td>0.18</td>
</tr>
<tr>
<td>Pegylated interferon-alfa</td>
<td></td>
</tr>
<tr>
<td>Genotype 1</td>
<td>0.14</td>
</tr>
<tr>
<td>Non-genotype 1</td>
<td>0.31</td>
</tr>
<tr>
<td>Pegylated interferon-alfa + ribavirin</td>
<td></td>
</tr>
<tr>
<td>Genotype 1</td>
<td>0.29</td>
</tr>
<tr>
<td>Non-genotype 1</td>
<td>0.58</td>
</tr>
<tr>
<td>Probability of early virologic response (EVR)</td>
<td></td>
</tr>
<tr>
<td>Interferon-alfa + ribavirin</td>
<td>0.38</td>
</tr>
<tr>
<td>Pegylated interferon-alfa</td>
<td>0.55</td>
</tr>
<tr>
<td>Pegylated interferon-alfa + ribavirin</td>
<td>0.71</td>
</tr>
<tr>
<td>Treatment mortality probability</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Costs (2004 US $)</strong>&lt;sup&gt;31,34-37&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Treatment Protocol 1&lt;sup&gt;‡&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Interferon-alfa + ribavirin</td>
<td>$15,568</td>
</tr>
<tr>
<td>Pegylated interferon-alfa</td>
<td>$19,305</td>
</tr>
<tr>
<td>Pegylated interferon-alfa + ribavirin</td>
<td>$27,880</td>
</tr>
<tr>
<td>Treatment Protocol 2§</td>
<td></td>
</tr>
<tr>
<td>Interferon-alfa + ribavirin</td>
<td>$8,769</td>
</tr>
<tr>
<td>Pegylated interferon-alfa</td>
<td>$13,109</td>
</tr>
<tr>
<td>Pegylated interferon-alfa + ribavirin</td>
<td>$22,022</td>
</tr>
<tr>
<td><strong>Costs of annual HCV care</strong></td>
<td></td>
</tr>
<tr>
<td>Chronic HCV</td>
<td>$1,440</td>
</tr>
<tr>
<td>Compensated cirrhosis</td>
<td>$1,017</td>
</tr>
<tr>
<td>Ascites</td>
<td>$4,280</td>
</tr>
<tr>
<td>Variceal hemorrhage, first year</td>
<td>$23,669</td>
</tr>
<tr>
<td>Variceal hemorrhage, subsequent years</td>
<td>$4,632</td>
</tr>
<tr>
<td>Hepatic encephalopathy, first year</td>
<td>$15,192</td>
</tr>
<tr>
<td>Hepatic encephalopathy, subsequent years</td>
<td>$3,519</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>$40,828</td>
</tr>
<tr>
<td>Liver transplant, first year</td>
<td>$134,458</td>
</tr>
<tr>
<td>Liver transplant, subsequent years</td>
<td>$23,481</td>
</tr>
<tr>
<td><strong>Costs of annual HIV care</strong></td>
<td></td>
</tr>
<tr>
<td>CD4 count 200-500/mm&lt;sup&gt;3&lt;/sup&gt;</td>
<td>$5,096</td>
</tr>
<tr>
<td>Three-drug antiretroviral therapy</td>
<td>$13,752</td>
</tr>
</tbody>
</table>

*Annual rates per person are presented here except where specified, but these were converted to monthly probabilities in the model.
†Fibrosis progression rates are assumed to be linearly interpolated in these ranges. Progression rates were the same for Metavir stages F0 to F4.
‡All patients receive the full 48-week course of therapy.
§Patients without an early virologic response receive only 12 weeks of therapy; protocol assumes nondrug costs by week 12 are two-thirds of nondrug costs for 48 weeks of therapy.
sion had slightly more influence on the cost-effectiveness of treatment.

Results were sensitive to the discount rate used. With no discounting, the incremental cost-effectiveness ratio was approximately 60% lower than in the base case, while a discount rate of 5% resulted in a ratio that was 140% higher than the base case. Varying the cost of pegylated interferon and ribavirin between 50% and 150% of the base-case result in cost-effectiveness ratios ranging from $56,300 to $88,500 per YLS.

DISCUSSION

We found that treatment with pegylated interferon and ribavirin for chronic HCV infection in an urban cohort co-infected with HIV, with characteristics similar to the treatment-eligible subgroup in the HALO Study, will provide substantial life expectancy gains. These gains are greatest in patients with non-genotype 1 infection. Combination therapy with pegylated interferon and ribavirin was the most effective and cost-effective treatment strategy regardless of genotype or sex.

There were substantial differences in the cost-effectiveness ratios between patients with genotype 1 and non-genotype 1 HCV, mainly attributable to differences in treatment response rates. There were only small differences in the cost-effectiveness of treatment for men and women.

Four trials to date have evaluated the efficacy of HCV treatment in HIV-HCV co-infected patients; we used efficacy results from the largest multicenter trial, APRICOT, for the base case analysis.\textsuperscript{11-14} Each trial found treatment to be more efficacious in non-genotype 1 patients, but the magnitude of the differences varied by trial, as shown in Figure 2. Because cost-effectiveness ratios are sensitive to treatment efficacy, a wide range of cost-effectiveness ratios is possible. Despite our conservative assumptions, our cost-effectiveness results may be particularly favorable because APRICOT efficacies are high relative to other trials’ results. However, the APRICOT study’s treatment protocol requiring 800 mg of ribavirin daily has become a relatively standard clinical practice. The AIDS Clinical Trials Group (ACTG) trial, on the other hand, administered ribavirin according to a dose-escalation schedule.\textsuperscript{13} The Agence Nationale de Recherches sur le Sida (ANRS) HCO2-RIBAVIC trial’s exclusion criteria for subjects was less stringent than the other trials, allowing patients with alcohol intake up to 40 grams per day for women or 50 grams per day for men to participate; 21% of this study population had psychiatric disorders, and 40% had bridging fibrosis or cirrhosis.\textsuperscript{12} These factors may explain the higher efficacies demonstrated in the APRICOT study, but may also indicate that the RIBAVIC trial results might be more applicable to the urban cohort modeled here if treatment eligibility criteria are relaxed.

In contrast to results from a recent cost-effectiveness analysis in HCV mono-infected patients, the present study suggests that monotherapy with pegylated interferon and combination therapy with interferon and ribavirin are dominated strategies in co-infected populations.\textsuperscript{17} Our results also differ from those of Kuehne et al.\textsuperscript{16} because this earlier cost-effectiveness analysis of HIV-HCV co-infected pa-

| Table 3 | Cost-effectiveness of 4 Strategies by Treatment Protocol, Sex, and Genotype for Patients with CD4 Count 200-500 Cells/mm$^3$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy*</td>
<td>Cost, 2004 US$</td>
<td>Life Expectancy, Years</td>
<td>Incremental Cost per YLS</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Genotype 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>$240,300</td>
<td>11.63</td>
<td>–</td>
</tr>
<tr>
<td>Interferon + ribavirin</td>
<td>$256,400</td>
<td>11.71</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon</td>
<td>$261,100</td>
<td>11.83</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon + ribavirin</td>
<td>$271,700</td>
<td>12.06</td>
<td>$73,000</td>
</tr>
<tr>
<td><strong>Non-genotype 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>$240,300</td>
<td>11.63</td>
<td>–</td>
</tr>
<tr>
<td>Interferon + ribavirin</td>
<td>$257,900</td>
<td>11.88</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon</td>
<td>$263,400</td>
<td>12.09</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon + ribavirin</td>
<td>$275,600</td>
<td>12.52</td>
<td>$39,700</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Genotype 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>$252,200</td>
<td>12.28</td>
<td>–</td>
</tr>
<tr>
<td>Interferon + ribavirin</td>
<td>$268,400</td>
<td>12.37</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon</td>
<td>$273,200</td>
<td>12.48</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon + ribavirin</td>
<td>$284,000</td>
<td>12.73</td>
<td>$70,700</td>
</tr>
<tr>
<td><strong>Non-genotype 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No treatment</td>
<td>$252,200</td>
<td>12.28</td>
<td>–</td>
</tr>
<tr>
<td>Interferon + ribavirin</td>
<td>$270,000</td>
<td>12.55</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon</td>
<td>$275,700</td>
<td>12.76</td>
<td>†</td>
</tr>
<tr>
<td>Pegylated interferon + ribavirin</td>
<td>$288,400</td>
<td>13.20</td>
<td>$39,300</td>
</tr>
</tbody>
</table>

*Assumes 48 weeks of HCV therapy for all patients. †-dominated strategy.
and subject characteristics. In a previous analysis with treatment efficacy data and identify influential parameters. Implications of our analysis are restricted to a specific target population of treatment-eligible patients with stable HIV disease and stable CD4 cell counts between 200 and 500 cells/mm$^3$. A more sophisticated model of co-infection that fully represents the natural history of HIV disease will be necessary to explore important questions regarding the optimal timing of treatment for chronic HCV relative to antiretroviral therapy for HIV disease, and potential interactions or additive toxicities between treatments. Other limitations of our study include the uncertainty surrounding many of the model’s parameters. Also, we did not consider health-related quality of life in co-infected patients, and all of our costs were literature-based.

We chose to use population characteristics and treatment eligibility criteria from an urban cohort that has been previously described. While the treatment-eligible portion of the HALO Study cohort was in many respects similar to the population studied by APRICOT, efficacy as demonstrated in a randomized controlled trial is not the same as effectiveness in a typical clinical setting. Prospective studies currently underway with the HALO Study and other cohorts will provide insight to treatment effectiveness in a particular population of HIV-HCV patients. It is important to note that only 30% of the co-infected subgroup of the HALO Study cohort was eligible for HCV treatment; the remainder were not eligible for a variety of reasons, including unstable social circumstances, concern about potential adverse effects, and concern about ability to work during the treatment course.

As treatment-eligible co-infected patients are not currently the norm, further studies are needed to establish the effectiveness of combination HCV therapy in populations with low eligibility for treatment. Overcoming barriers to HCV treatment eligibility and initiation in HIV-HCV co-infected patients remains a priority, now that combination therapy has been demonstrated efficacious in certain populations. For co-infected patients with stable HIV disease, treatment appears to be not only life-prolonging but cost-effective as well.

**References**


NSAID Use and Progression of Chronic Kidney Disease

Katherine Gooch, MSc,a,b,c Bruce F. Culleton, MD, MSc,b,d Braden J. Manns, MD, MSc,a,b,d,e Jianguo Zhang, MSc,d Helman Alfonso, PhD,f Marcello Tonelli, MD, SM,a,b,g Cy Frank, MD,c,h Scott Klarenbach, MD, MSc,a,b,g Brenda R. Hemmelgarn, PhD, MDb,d,e

aInstitute of Health Economics, Edmonton, Alberta; bAlberta Kidney Disease Network, Edmonton; cAlberta Bone and Joint Health Institute, Calgary; dDepartment of Medicine and eDepartment of Community Health Sciences, University of Calgary, Alberta; fDepartment of Public Health, Curtin University of Technology, Perth, Australia; gDepartment of Medicine, University of Alberta, Edmonton; and hDepartment of Surgery, University of Calgary, Alberta, Canada.

ABSTRACT

PURPOSE: The effects of nonselective and selective cyclooxygenase-2 specific (COX-2) nonsteroidal anti-inflammatory drug (NSAID) use on the progression of chronic kidney disease (CKD) is uncertain. Due to the high prevalence of both CKD and NSAID use in older adults, we sought to determine the association between NSAID use and the progression of CKD in an elderly community-based cohort.

METHODS: All subjects ≥66 years of age who had at least one serum creatinine measurement in 2 time periods (July-December, 2001 and July-December, 2003) were included. Multiple logistic regression analyses, including covariates for age, sex, baseline estimated glomerular filtration rate (eGFR), diabetes, and comorbidity were used to explore the associations of NSAID use on the primary (decrease in eGFR of ≥15 mL/min/1.73²) and secondary (mean change in eGFR) outcomes.

RESULTS: A total of 10,184 subjects (mean age 76 years; 57% female) were followed for a median of 2.75 years. High-dose NSAID users (upper decile of cumulative NSAID exposure) experienced a 26% increased risk for the primary outcome (odds ratio [OR] 1.26, 95% confidence interval [CI], 1.04-1.53). A linear association between cumulative NSAID dose and change in mean GFR also was seen. No risk differential was identified between selective and nonselective NSAID users.

CONCLUSIONS: High cumulative NSAID exposure is associated with an increased risk for rapid CKD progression in the setting of a community-based elderly population. For older adult patients with CKD, these results suggest that nonselective NSAIDs and selective COX-2 inhibitors should be used cautiously and chronic exposure to any NSAID should be avoided. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Chronic kidney disease; Progressive kidney disease; Non-steroidal anti-inflammatory drug; Cyclooxygenase-2 inhibitor; Elderly

Chronic kidney disease is a worldwide public health problem with an increasing incidence and prevalence, particularly in the elderly population.1-7 Nonsteroidal anti-inflammatory drugs (NSAIDs) have been identified as nephrotoxic agents with both acute and chronic effects on kidney function. While the short-term biological effects of sodium retention, edema, and acute renal failure with NSAIDs are well documented, there are limited scientific data reporting the safety of these drugs on kidney function when NSAIDs are taken chronically or when they are taken by patients with pre-existing kidney disease. Existing data regarding long-term NSAID exposure is inconsistent, with earlier studies suggesting an increased risk for adverse kidney related outcomes,8-12 though more recent reports have failed to confirm these risks.13-15

The frequency of NSAID use, including nonselective conventional NSAIDs and selective cyclooxygenase-2 (COX-2) inhibitors, has increased in the last several years. Potential factors responsible for this increase include over-the-counter availability, an aging population with concomi-
itant musculoskeletal disorders, and the perceived superior gastrointestinal safety profile associated with COX-2 inhibitors. In particular, COX-2 inhibitor use has increased substantially since their introduction in 1999. Of the more than 25% Albertan senior population who were prescribed at least 1 NSAID during a 1-year period, 68% received a COX-2 inhibitor.16

Given the high prevalence of both chronic kidney disease and NSAID use in the elderly, and the uncertainty of the chronic kidney disease risk associated with NSAIDs, this study sought to examine the association between NSAID use and progression of chronic kidney disease in a community-based cohort of elderly subjects. This study also sought to determine if this association differed for conventional nonselective NSAIDs versus selective COX-2 inhibitors.

**METHODS**

**Study Population**

The Conjoint Ethics Review Board at the University of Calgary approved this study. A cohort of elderly subjects aged ≥66 years were identified from the Calgary Laboratory Services database in Calgary, Alberta, Canada. This laboratory provides testing for the entire Calgary Health Region (catchment population 1.1 million) using a single regional laboratory and standardized methods that are recalibrated routinely against reference samples. To be eligible for inclusion in this study, participants required at least one serum creatinine measurement in 2 study periods: July 1, 2001 to December 31, 2001; and July 1, 2003 to December 31, 2003. To reduce the impact of episodes of acute renal failure, laboratory measurements associated with a hospital admission were not included. Subjects were also excluded from the cohort if they had more than 12 outpatient serum creatinine measurements in either of the 6-month observation periods, as they were likely to represent patients with acutely unstable kidney function. Subjects receiving dialysis at study entry were also excluded.

**Measurement of Kidney Function and Definition of Outcomes**

Glomerular filtration rate was used to estimate kidney function using the abbreviated Modification of Diet in Renal Disease (MDRD) equation, which includes variables for age, sex, race and serum creatinine.17 Although data for race were not available for the cohort, less than 1% of the Alberta population is African American.18 Therefore, the impact at the population level of eliminating race from the estimate of glomerular filtration rate was expected to be minimal. Furthermore, given the study’s focus on change in glomerular filtration rate, information on race was not needed. Because of concerns about the validity of the MDRD equation for subjects with higher levels of kidney function,19 subjects with baseline glomerular filtration rate values exceeding 90 mL/min/1.73² were also excluded.

Serum creatinine measurements were analyzed in the same laboratory, thus eliminating the potential for inter-laboratory measurement variation. However, because of possible intra-laboratory variation in measurement resulting from changes in calibration of serum creatinine assays, measurements between the 2 time periods were assessed and calibrated in the following manner. First, a subset of healthy subjects (defined as subjects with no prescriptions for medications commonly used to treat cardiovascular disease, hypertension or diabetes mellitus in the year before the index glomerular filtration rate) younger than 80 years of age was identified. The median serum creatinine measurement for these subjects, by 1-year age increments, for the 2001 and the 2003 time periods was calculated. The difference between measurements for the 2 periods was calculated, and the average of the differences determined. To correct for the systematic differences in serum creatinine measurements evident from this analysis, 2.0 umol/L (0.02 mg/dL) was subtracted from the serum creatinine measurements in 2003.

The primary outcome was rapid progression of kidney disease, defined as a decrease in glomerular filtration rate ≥15 mL/min/1.73². Progression was calculated as the difference in the subject mean glomerular filtration rate for the 2 time periods: July 1, 2001 to December 31, 2001 and July 1, 2003 to December 31, 2003. A progression of ≥15 mL/min/1.73² was approximately the 88th percentile for progression within the entire cohort. The change in mean glomerular filtration rate (as a continuous variable) over the study period was chosen as a priori as a secondary outcome.

**Measure of Exposure**

Using the unique provincial health care number for each subject, laboratory data were linked to the provincial administrative Alberta Blue Cross database to obtain information on prescription drug use for the exposure period July 1, 2000 to March 31, 2003. All residents of Alberta aged 65 years and older receive insured health services including coverage for prescription drugs. NSAID exposure (for the period 1 year before the initial glomerular filtration rate measurement up to March 31, 2003) was defined using 2 approaches. The first was a broad approach based on the

---

**CLINICAL SIGNIFICANCE**

- High NSAID exposure is associated with a rapid decrease in kidney function.
- Traditional non-selective NSAIDs and cyclo-oxygenase-2 specific NSAIDs are associated with a similar risk of kidney function decline.
- Chronic exposure to any NSAID should be avoided in older adult patients with chronic kidney disease.
presence or absence of exposure, while the second included a measure of the cumulative dose of NSAID received during the exposure period. For the broad categorization, there were 4 mutually exclusive categories: nonusers, which had no use of any NSAID; nonselective NSAIDs and COX-2 inhibitors combined, at least one prescription for a nonselective NSAID, as well as at least one prescription for a COX-2 inhibitor; nonselective NSAIDs only, at least one prescription for a nonselective NSAID, with no prescriptions for COX-2 inhibitors; COX-2 inhibitors only, at least one prescription for a COX-2 inhibitor with no prescription for a nonselective NSAID.

The second approach to the exposure definition took into account the cumulative NSAID dose, using the anatomical chemical therapeutic code and defined daily dose to standardize NSAID exposure in the following manner:

\[
\text{Drug exposure} = \text{drug strength} \times \text{drug quantity/defined daily dose}
\]

Total drug exposure was combined to obtain a cumulative dose of NSAID exposure for each subject and was categorized into nonusers (no use of any NSAID during the study period); low-dose users (cumulative dose < 90th percentile); and high-dose users (cumulative dose ≥ 90th percentile). Aspirin was excluded from the exposure classification as it is routinely available over the counter and, therefore, was not captured in the Alberta Blue Cross database.

Measure of Covariates

Other variables of interest included patient age and sex, diabetes, and overall comorbidity status. Subjects were identified as having diabetes if they received at least 1 prescription for insulin or an oral hypoglycemic agent in the year before their first serum creatinine measurement. A measurement of comorbidity status, based on the use of prescription drugs, was calculated using the Chronic Disease Score (CDS) as described by Clark et al. The CDS is a validated weighted index of prescription medications, whereby higher scores are a result of more classes of medications dispensed, especially medications used to treat serious diseases.

Statistical Methods

Baseline characteristics by type of NSAID user are presented as means and standard deviations for normally distributed continuous variables and percent prevalence for dichotomous variables. Given the skewed nature of the comorbidity score, these data are presented as median with interquartile range. The statistical significance of the differences in baseline characteristics across categories of NSAID use was determined by chi-squared test, analysis of variance, and Kruskal-Wallis analysis, where appropriate. The association between NSAID use and the risk of rapid progression was assessed using multivariate logistic regression, adjusting for age, sex, baseline glomerular filtration rate, diabetes and comorbidity score. Similar analyses were performed using cumulative NSAID dose as a categorical independent variable, using non-NSAID users as the reference group. Finally, multiple linear regression analyses were undertaken to assess the association of cumulative NSAID exposure (defined daily dose as a continuous variable) with decline in mean glomerular filtration rate. Given their clinical importance, age, sex, baseline glomerular filtration rate, diabetes, and comorbidity score were included in all adjusted models. Assumptions for the logistic and linear regression models were tested and met. All analyses were conducted using SAS (version 8.01, SAS Institute Inc., Cary, NC) or STATA (version 8, STATA Corp., College Station, Tex).

RESULTS

There were 12,641 subjects ≥66 years of age identified who had at least 1 outpatient measurement of serum creatinine in each of the 2 defined time periods. As outlined in Figure 1, a total of 2545 patients were excluded as they did not meet the inclusion criterion, for a final study cohort of 10,184.

Baseline subject characteristics by type of NSAID are shown in Table 1. Baseline subject characteristics varied by type of NSAID, with traditional NSAID users tending to be younger, male, and have a lower comorbidity score and mean glomerular filtration rate than COX-2 inhibitor users and the combined NSAID and COX-2 inhibitor users. NSAID use was inversely associated with the subject’s age and comorbidity score and directly associated with a higher mean baseline glomerular filtration rate. Females were more likely than males to use any NSAID.

There were 1353 (13.3%) subjects who experienced the primary outcome of a decrease in glomerular filtration rate ≥ 15 mL/min/1.73 m² over a median of 2.75 years follow-up. In the multivariate logistic regression analysis with the NSAID type as a categorical variable, there was a significant interaction between mean glomerular filtration rate and NSAID type (P = .03). Results of this analysis are therefore presented stratified by patient mean glomerular filtration rate (Table 2). In this analysis, among subjects with a mean glomerular filtration rate of 60-89 mL/min/1.73 m², COX-2 inhibitor users had a 25% increased risk of rapid progression of kidney disease (odds ratio [OR] 1.25, 95% confidence interval [CI], 1.05-1.47) and traditional NSAID users a 29% increased risk (OR 1.29, 95% CI, 1.02-1.63) compared with non-NSAID users. There was no association between NSAID use and rapid progression of chronic kidney disease for the other 2 categories of mean glomerular filtration rate.

NSAID Cumulative Dose and Chronic Kidney Disease Progression

Figure 2 illustrates the distribution of NSAID cumulative dose (defined daily dose) for study subjects that were exposed to at least one NSAID or COX-2 inhibitor. After adjusting for age, sex, mean glomerular filtration rate, diabetes, and comorbidity, high dose NSAID users (cumulative...
dose ≥90th percentile for all subjects) were 26% more likely to develop the primary outcome than nonusers (OR 1.26, 95% CI, 1.04-1.53). The relationship between NSAID cumulative dose and progression of chronic kidney disease was further explored in multiple linear regression analysis with change in mean glomerular filtration rate as the dependent variable, adjusting for study mean glomerular filtration rate, age, sex, diabetes and comorbidity. For each 100-unit

![Figure 1](https://example.com/figure1.png)

**Figure 1** Formation of study cohort and reasons for exclusion (GFR = glomerular filtration rate).

### Table 1 Baseline Subject Characteristics by NSAID Use

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Non-NSAID Users (n = 5304)</th>
<th>Nonselective NSAID and COX-2 Inhibitor Users (n = 1167)</th>
<th>Nonselective NSAID Users Only (n = 1110)</th>
<th>COX-2 Inhibitor Users Only (n = 2603)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>76.3 ± 6.9</td>
<td>75.3 ± 6.2</td>
<td>74.9 ± 6.3</td>
<td>76.2 ± 6.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Females (%)</td>
<td>54.8</td>
<td>60.3</td>
<td>50.6</td>
<td>64.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Glomerular filtration rate, mL/min/1.73m²</td>
<td>64.5 ± 16.9</td>
<td>63.3 ± 16.1</td>
<td>62.4 ± 17.5</td>
<td>64.5 ± 16.2</td>
<td>.0004</td>
</tr>
<tr>
<td>Chronic Disease Score</td>
<td>2011 (1432 – 2772)</td>
<td>2520 (1864 – 3468)</td>
<td>2283 (1590 – 3182)</td>
<td>2304 (1590 – 3262)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>16.0</td>
<td>17.7</td>
<td>18.7</td>
<td>16.7</td>
<td>.13</td>
</tr>
</tbody>
</table>

Age and glomerular filtration rate expressed as mean ± standard deviation; Chronic Disease Score expressed as median and interquartile range.

### Table 2 Multivariate Adjusted Odds Ratios for Rapid Progression of Kidney Disease and Type of NSAID Use, by Stage of Kidney Disease

<table>
<thead>
<tr>
<th>Stage of Chronic Kidney Disease (Glomerular Filtration Rate in mL/min/1.73m²)</th>
<th>Glomerular filtration rate 60-89</th>
<th>Glomerular filtration rate 30-59</th>
<th>Glomerular filtration rate &lt; 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category of NSAID exposure</td>
<td>n</td>
<td>Odds ratio* (95% CI)</td>
<td>n</td>
</tr>
<tr>
<td>No NSAID se</td>
<td>3475</td>
<td>Ref</td>
<td>1590</td>
</tr>
<tr>
<td>Nonselective NSAID and COX-2 inhibitor use</td>
<td>722</td>
<td>1.10 (0.87-1.40)</td>
<td>411</td>
</tr>
<tr>
<td>Nonselective NSAID use only</td>
<td>671</td>
<td>1.29 (1.02-1.63)</td>
<td>374</td>
</tr>
<tr>
<td>COX-2 inhibitor use only</td>
<td>1705</td>
<td>1.25 (1.05-1.47)</td>
<td>816</td>
</tr>
</tbody>
</table>

*Adjusted for age, sex, diabetes, baseline glomerular filtration rate and comorbidity score.
increase in defined daily dose there was an associated decrease in glomerular filtration rate of 0.08 mL/min/1.73 m² (95% CI, 0.01 to 0.16; P = .04) over the study period. Figure 3 illustrates this linear relationship between NSAID cumulative dose and change in mean glomerular filtration rate.

**DISCUSSION**
In this prospective community-based study of over 10,000 elderly subjects a small, but statistically significant, increased risk of rapid progression of chronic kidney disease was found only among subjects with a baseline mean glomerular filtration rate between 60 and 89 mL/min/1.73 m² who were exposed to nonselective NSAIDs only or to COX-2 inhibitors only. NSAID use was not associated with an increased risk of rapid progression of chronic kidney disease for subjects with lower levels of kidney function, nor was any risk differential identified between the nonselective and selective NSAIDs. This data should be interpreted with caution given the conservative nature of the
analysis. Many patients classified as NSAID users had very limited exposure to NSAIDs, thus increasing the probability of missing an important association between prolonged use of NSAIDs and deterioration in kidney function.

Supporting this argument was the effect of NSAID cumulative dose on both rapid chronic kidney disease progression and change in mean glomerular filtration rate. In this analysis, high NSAID use, defined as the upper 90th percentile of defined daily dose, was associated with a 26% increase in risk of rapid progression, compared with non-NSAID use. To put this into context, Table 3 provides examples of several NSAID quantities that would be defined as “high NSAID use.” These examples assume exposure to the NSAID for a duration of 2.75 years. For example, a patient consuming on average 120 mg of celecoxib per day for 1004 days would be defined as having high exposure. This finding was further supported by the linear regression analysis whereby each 100 increment in NSAID defined daily dose was associated with a decrease in mean glomerular filtration rate of 0.08 mL/min/1.73 m². A defined daily dose of 100 is equivalent to 100 tablets of celecoxib 200 mg, or 600 tablets of ibuprofen 200 mg.

The dose response effect seen in this study is consistent with the findings of Perneger et al (n = 1077), where the risk of end-stage renal disease was increased more than 8-fold with high lifetime NSAID use (>5000 tablets) compared with subjects using NSAIDs infrequently. Sandler et al (n = 554) also found a 2-fold increased risk of end-stage renal disease in patients reporting daily NSAID use. Several studies of NSAID use and incident acute renal failure have also reported a dose response effect.

Despite the results from Perneger and Sandler, the effect of NSAID use on chronic kidney disease progression remains controversial, largely because of 2 recent publications. In analysis of data from the Nurses’ Health Study (n = 1697), Curhan and colleagues found no association between NSAID exposure and the 11-year risk of kidney function decline. Given that NSAID exposure was determined by questionnaire, it is possible that recall bias affected the study results. Furthermore, the study participants were all female, younger in age (mean age 56.5 years), and had better baseline kidney function (mean glomerular filtration rate 88 mL/min/1.73 m²) than this study cohort, and were therefore at substantially less risk for kidney function decline. A 14-year follow-up study of 4494 male physicians within the United States also failed to show any association between NSAID use and kidney function loss. Similar to the Nurses’ Health Study, questionnaires were used to determine prior NSAID exposure, and study subjects were at a very low risk for progressive kidney disease (mean age 48.9 years and baseline percent diabetes 1.1%).

Although the current study provides important information on the relationship between NSAID exposure and decrease in kidney function, the results should be interpreted within the context of the study’s limitations. Specifically, glomerular filtration rate was not measured directly but was estimated using a serum creatinine measurement. Also, there was no attempt to calibrate serum creatinine measurements with the Cleveland Clinic laboratory from which the MDRD equation was derived. Such calibration is critical to estimate prevalence of kidney disease. The primary interest, however, was in the change of kidney function, whereby such calibration was deemed unnecessary.

There are also limitations as a result of the study design. First, exposure assessment ceased on March 31, 2003, while the outcome assessment was undertaken after that (July 01, 2003 to December 31, 2003). As the study was interested in the chronic long-term effects of NSAID use, this is unlikely to adversely bias the study results. In fact, this difference in the timing of exposure and outcome assessment would likely reduce any impact of acute renal failure episodes on the study outcomes. Second, the use of laboratory data to define the study cohort limited the study to subjects who sought medical care and had a serum creatinine measurement. Also, there was no attempt to calibrate serum creatinine measurements obtained (bias by indication). As the study sample was based on the elderly, who are more likely to access the healthcare system and have laboratory testing performed, this is unlikely to substantially bias the study results. This is further supported by a similar age distribution in this study to that reported for the general population of the Calgary Health Region (data not shown). Data from a cohort identified by laboratory-based case finding is also easily generalized to primary care practice. Third, although the prescription drug database eliminates recall bias, exposure bias may still exist. This may occur if the dispensed NSAID is not consumed or if over-the-counter NSAIDs, such as ibupro-
fen, are used. This potential for exposure misclassification however, would be expected to bias the results towards the null. Finally, the possibility of residual confounding cannot be excluded. The results of the study were unable to be adjusted directly for blood pressure, although antihypertensive medication use contributes to the comorbidity score. Also, the possibility that NSAID use was triggered by a predisposing condition that directly impacts kidney function cannot be excluded.

Despite these limitations, the current study has several strengths. First, the study involved many NSAIDs used in today’s clinical practice, unlike prior reports. The study was able to assess the effects of nonselective NSAID and COX-2 inhibitor exposure separately, an assessment not possible in previous studies. Second, recall bias was eliminated with the use of computerized drug prescription data. In contrast, several other large observational studies used questionnaires to determine exposure to NSAIDs. Third, the size of the cohort (over 10,000 elderly subjects) and its community-based setting increases the generalizability of the study results to community-dwelling elderly individuals. This is particularly relevant given the prevalence of chronic kidney disease and musculoskeletal disorders in this population.

In conclusion, the results of this study show that high cumulative NSAID exposure is associated with an increased risk for rapid chronic kidney disease progression, as well as a decrease in mean glomerular filtration rate in the setting of a community-based elderly population. For older adult patients with chronic kidney disease, these results suggest that nonselective NSAIDs and selective COX-2 inhibitors should be used cautiously and chronic exposure to any NSAID should be avoided.

References

CLINICAL RESEARCH STUDY

Assessment of FIBROspect II to Detect Hepatic Fibrosis in Chronic Hepatitis C Patients

Atif Zaman, MD, MPH,a Hugo R. Rosen, MD,a Ken Ingram, PA-C,a Christopher L. Corless, MD, PhD,a Esther Oh,b Katie Smithb
aOregon Health & Science University, Portland; and bPrometheus Laboratories, San Diego, Calif.

ABSTRACT

BACKGROUND: The degree of liver fibrosis in patients with Hepatitis C (HCV) provides important prognostic information; however, the only current method available to obtain this information is by performing a liver biopsy. Liver biopsies are invasive, associated with complications, and costly. There has been recent interest in developing a panel of serum markers that can reliably predict the presence of fibrosis and, thus, obviate the need for a liver biopsy. Our objective was to prospectively validate a panel of serum fibrosis markers (FIBROspectSM II) that has been recently developed.

METHODS: Serum was obtained from 108 consecutive HCV (15% with HCV/ETOH) patients seen in a hepatology clinic at a single tertiary care center at the time of liver biopsy. The performance of FIBROspect II (consisting of 3 fibrosis markers: hyaluronic acid, tissue inhibitor of metalloproteinases 1, and alpha-2-macroglobulin) in differentiating mild (F0-F1) from significant (F2-F4) fibrosis was assessed by comparing the panel results with performed liver biopsy.

RESULTS: The prevalence of significant fibrosis in the study group was 36.1%. The diagnostic value of the serum marker panel to detect significant fibrosis as assessed by area under the receiver operating characteristic (ROC) curve was 0.826. Performance characteristics are as follows: sensitivity 71.8%, specificity 73.9%, positive predictive value 60.9%, negative predictive value 82.3%, and overall accuracy of 73.1%.

CONCLUSION: This prospective study supports the clinical utility of serum markers in detecting fibrosis and validates the performance of FIBROspect II in a prospective cohort of patients. The high negative predictive value of the test provides a reliable alternative to rule out severe fibrosis. © 2007 Elsevier Inc. All rights reserved.

KEYWORDS: Hepatitis C; Hepatic fibrosis

The majority of chronic liver diseases are typified by hepatic injury and inflammation, which then leads to fibrogenesis.1 Fibrogenesis is considered a dynamic process characterized by the formation of the constituents of the extracellular matrix, which is a mixture of glycoproteins and proteoglycans in a complex network. Fibrosis is felt to be a physiologic mechanism, initially beneficial in helping to limit the extension of the inflammatory reaction, but as injury persists it is detrimental to the liver. Collagen and matrix proteins that constitute fibrosis are largely produced by activated stellate cells. Over time, the inflammation seen in the liver due to chronic injury can fluctuate, worsening or improving over time. However, fibrosis, in a setting of chronic persistent injury, is believed to be progressive and largely irreversible.2,3 Ultimately, progressive fibrosis leads to architectural distortion of the liver, and cirrhosis. Therefore, the progression of fibrosis determines prognosis in patients with chronic liver disease.

Currently, the gold standard for determining the degree of hepatic fibrosis is a liver biopsy. Several scoring systems...
to assess fibrosis have been proposed, including the Knodell score, the Ishak score, and the Metavir score. The Metavir system has been carefully validated in chronic hepatitis C patients and is used increasingly in practice. None to mild fibrosis is generally considered F0-F1, and F2-F4 is considered significant fibrosis—where F4 is cirrhosis. Liver biopsies, however, are invasive and have associated morbidity including pain, intrahepatic hemorrhage, intrahepatic or subcapsular hematomas, hemobilia, bile peritonitis, and pneumothorax. Complications requiring hospitalization occur in 1%-3% of patients, and mortality rates are between 1 in 10,000 and 1 in 12,000. The diagnostic limitations of a percutaneous liver biopsy include inter- and intra-observer variability of the fibrosis staging and sampling error. Although factors that improve the diagnostic accuracy of a liver biopsy include use of a Trucut 15-gauge needle rather than Menghini-type needles, multiple passes, biopsy core sizes ≥2 cm, and use of pathologists familiar with liver biopsy readings, even in the ideal situation a single pass percutaneous liver biopsy incorrectly stages fibrosis in 20% of patients.

Because of the invasiveness of a liver biopsy and the issues outlined above, there has been great interest in developing noninvasive markers of fibrosis. Recently, a panel of fibrosis markers has been developed and consists of hyaluronic acid (HA), tissue inhibitor of metalloproteinases 1 (TIMP-1), and alpha-2-macroglobulin with very good performance characteristics in predicting the absence/presence of significant hepatic fibrosis. However, this serum panel was developed using banked serum samples. Further, none to mild fibrosis is generally considered F0-F1, and F2-F4 is considered significant fibrosis. Therefore, the aim of this study was to prospectively validate this panel of serum markers from an independent cohort. Therefore, the aim of this study more, this serum panel has not been validated using samples seen in a hepatology clinic at a university hospital.

CLINICAL SIGNIFICANCE

- Liver biopsy is an expensive and an invasive procedure, but it is the current gold standard for determining hepatic fibrosis in hepatitis C patients.
- The serum fibrosis panel FibroSpect II can effectively identify people who do not have fibrosis, but it may not be useful in differentiating between intermediate stages of fibrosis.

METHODS

Study Subjects

Consecutive patients with hepatitis C seen in the hepatology clinic at Oregon Health and Science University were asked to participate in this study. All patients were anti-HCV (hepatitis C) positive and had detectable plasma HCV-RNA by polymerase chain reaction. The study was approved by the Institutional Review Board at Oregon Health and Science University. The period of enrollment was between October 2001 and June 2003. All patients at the time of liver biopsy underwent a blood draw for a serum sample. All serum samples were shipped frozen to Prometheus Laboratories for analysis.

Serum Fibrosis Marker Assays

Serum levels of 3 fibrosis markers (FIBROspectrum II®, Prometheus Laboratories, San Diego, Calif) were determined by technologists blinded to clinical, laboratory and histological findings. Serum hyaluronic acid was measured in an enzyme-linked sandwich assay using HA-binding protein (Corgenix, Westminster, Conn). Tissue inhibitor of metalloproteinase-1 (TIMP-1) was measured by a sandwich ELISA (Amersham Pharmacia Biotech, Piscataway, NJ). Alpha2-macroglobulin was measured by nephelometry (Beckman Coulter, Brea, Calif). The analytical performance of these 3 assays has been validated in a clinical laboratory, with intra- and inter-assay variability of 2%-11% and 1%-11%, respectively.

Histologic Analysis

All liver biopsies were read by a single pathologist (C.C.) using the Metavir scoring classification where the stage of fibrosis was assessed on a 5-point scale (0 = no fibrosis, 1 = portal fibrosis without septa, 2 = few septa, 3 = numerous septa without cirrhosis, 4 = cirrhosis). Significant fibrosis was considered to be stage ≥2. A random sample of 1/3 of the liver biopsies was re-read by the pathologist to verify the reproducibility of the histologic findings. Excellent reproducibility of the histologic findings was observed with a κ = 0.8. Therefore, the original Metavir scores were used in the final analysis. A liver biopsy was considered to be adequate if the core sample was >15 mm and had more than 5 portal tracts in the specimen; all study samples were considered adequate by these criteria.

Statistical Analysis

Patient baseline characteristics are summarized using descriptive statistics and reported as mean ± standard deviation or proportions. The FibroSpectrum II index (0-100) was generated for each sample from a logistic regression model of the 3 markers that was previously established to discriminate F0-F1 from F2-F4 (Metavir) fibrosis. The diagnostic value of the algorithm in the study cohort was assessed by area under the ROC curve. Samples with an index ≥42 were classified as consistent with significant fibrosis, and those with an index from 0 to 41 were considered to be consistent with no/mild fibrosis to determine clinical performance of the algorithm. Statistical analyses were performed with SAS software version 8.0 (SAS Institute Inc., Cary, NC) and SPSS software version 11.5 (SPSS Inc., Chicago, Ill).

RESULTS

One hundred eight consecutive hepatitis C patients seen at the Oregon Health and Science University Hepatology
Spectroscopic imaging was no correlation between ALT and AST levels, and FIBROspect II results. Overall, nearly 2/3 of subjects had no/mild fibrosis, while only 13% of subjects had advanced fibrosis (stage 3 or 4) on liver biopsy.

### Performance Characteristics of FIBROspect II Index

Table 2 describes the clinical performance characteristics of FIBROspect II. The sensitivity and specificity were 71.8% and 73.9%, respectively, in a study population with an overall prevalence of significant fibrosis of 36.1%. For this prevalence, the positive and negative predictive values (PPV and NPV) were 60.9% and 82.3%, respectively. Figure 1 describes how the performance characteristics of the assay vary based on the prevalence of significant fibrosis (F2-F4). As the baseline prevalence decreases, the NPV increases. Thus, at a baseline prevalence of 20% significant fibrosis, the NPV is over 90%. The overall accuracy of the assay does not change over a wide prevalence range, nor does the sensitivity or specificity. The likelihood for a positive result is 2.75 (95% CI, 1.79-4.33), and 0.38 (95% CI, 0.22-0.61) for a negative result. When the cutoff of significant fibrosis was changed to F3-F4, the sensitivity, specificity, positive predictive value, and negative predictive value changed to 81.8%, 62.9%, 20%, and 96.8%, respectively.

The ROC curve of the test cohort is described in Figure 2. The area under the curve (AUC) was 0.826, and as a predictive tool it was significantly higher than chance alone (ie, AUC = 0.5). FIBROspect II test is based on a logistic regression index (range 0-100) determined by the predictive model. Figure 3 describes the frequency of Metavir fibrosis stages over the range of FIBROspect II index scores in the study population. Low index values (0-19) correctly identified no/mild fibrosis (F0-F1) in 90% of cases. On the other hand, a very high index score (80-100) correctly identified significant fibrosis (F2-F4) in nearly 80% of cases.

### DISCUSSION

The current study validates the performance characteristics of the FIBROspect II assay using a US cohort of HCV patients. Serum was collected prospectively at the time of liver biopsy. The overall performance of the diagnostic test was very good based upon the area under the ROC curve (a value of 0.826). Negative and positive predictive values were good to very good, depending upon the baseline prevalence of significant fibrosis (F2-F4) of the population. In general, the lower the baseline prevalence of significant fibrosis is, the better the negative predictive value of the test. Furthermore, extremely low and high index values of the test were excellent at correctly predicting the presence or absence of significant fibrosis. In our cohort, 19 patients had very low index value (<20) and in these patients a liver biopsy could be avoided. These findings are similar to a study by Patel et al where the FIBROspect II assay had similar performance characteristics—AUC of 0.831 and positive and negative predictive values of 74.3% and 75.8%, respectively. Thus, our study further validates this serum fibrosis panel, but using prospectively collected samples.

Numerous surrogate markers for hepatic fibrosis have been studied. There are direct biochemical markers of fibrosis, such as procollagen type III N-terminal peptide (PIIINP) and hyaluronic acid (HA). Because pathological accumulation of extracellular matrix is a result of alterations

### Table 1: Demographic, Laboratory and Histologic Characteristics of Hepatitis C Patients Undergoing a Liver Biopsy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study Cohort (n = 108)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>44 (24-60)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>70 (65%)</td>
</tr>
<tr>
<td>Female</td>
<td>38 (35%)</td>
</tr>
<tr>
<td>ALT in IU/L (SD)</td>
<td>81 (76)</td>
</tr>
<tr>
<td>AST in IU/L (SD)</td>
<td>59 (46)</td>
</tr>
<tr>
<td>Total bilirubin in mg/dL (SD)</td>
<td>1.1 (3.6)</td>
</tr>
<tr>
<td>Albumin in mg/dL (SD)</td>
<td>3.9 (0.3)</td>
</tr>
<tr>
<td>Metavir fibrosis stage</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>14 (13%)</td>
</tr>
<tr>
<td>1</td>
<td>55 (51%)</td>
</tr>
<tr>
<td>2</td>
<td>25 (23%)</td>
</tr>
<tr>
<td>3</td>
<td>12 (11%)</td>
</tr>
<tr>
<td>4</td>
<td>2 (2%)</td>
</tr>
</tbody>
</table>

ALT = alanine aminotransferase; AST = aspartate aminotransferase; SD = standard deviation.

### Table 2: Clinical Performance of the FIBROspect II Algorithm Using an Index Threshold of 42 to Classify Each Subject as Consistent with Mild (F0-F1) or Significant (F2-F4) Fibrosis

<table>
<thead>
<tr>
<th>Biopsy</th>
<th>F2-F4</th>
<th>F0-F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIBROspect II positive*</td>
<td>28</td>
<td>18</td>
</tr>
<tr>
<td>FIBROspect II negative*</td>
<td>11</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>69</td>
</tr>
<tr>
<td>Prevalence of F2-F4</td>
<td>36.1%</td>
<td>95% CI</td>
</tr>
</tbody>
</table>

Sensitivity 71.8% 55.1% 85.0%
Specificity 73.9% 61.9% 83.8%
PPV 60.9% 45.3% 74.9%
NPV 82.3% 70.5% 90.8%
Accuracy 73.1% 63.8% 81.2%

PPV = positive predictive value; NPV = negative predictive value.
*At an index threshold of 42.
in the synthesis and degradation of matrix proteins, matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) have also been studied as candidate markers. Matrix metalloproteinases constitute a family of zinc endopeptidases capable of degrading collagens, and TIMPs regulate its activity. Therefore, an imbalance in MMPs and TIMPs may be an important determinant of hepatic fibrosis. Proinflammatory markers such as LPS (lipopolysaccharide), TNF-alpha (transforming growth factor), and HGF (human growth factor) have also been evaluated. Other markers that have been studied include gamma glutamyltranspeptidase (GGT), total bilirubin, aspartase aminotransferases (AST), platelet count, YKL-40, apolipoprotein A1, and alpha 2 macroglobulin.

Recently, well-described serum fibrosis marker panels in addition to FIBROspect II include Fibrotest and the AST-
to-platelet ratio index (APRI). Forns et al used simple parameters including age, GTT, cholesterol level, and platelet count to predict hepatic fibrosis.11 Two cutoff points were determined (a low point of 4.2 and a high point of 6.9). A score <4.2 in 96% of patients accurately identified mild fibrosis. However, a score >6.9 correctly diagnosed significant fibrosis (stage 2-4) only 30% of the time. The APRI assay also uses more standard laboratory tests to assess hepatic fibrosis—the AST-to-platelet ratio index.12 This index was able to predict significant fibrosis accurately in 51% and cirrhosis in 81% of patients with hepatitis C. Imbert-Bismut et al developed a panel of fibrosis markers13 that consists of 5 biomarkers, including apolipoprotein A1, haptoglobin, alpha 2 macroglobulin, GTT, and total bilirubin. Their assay allowed the classification of 12% of patients as having no significant fibrosis (F0-F1 fibrosis) with a negative predictive value of 100%, and 34% of patients having significant fibrosis (F2-F4 fibrosis) with 90% positive predictive value.

At first glance, these studies suggest that the serum fibrosis panels are quite good for assessing the extent of liver fibrosis. However, there are some concerns regarding the currently available tests. Many of the assays have indeterminate values that limit their clinical utility. For example, the panel developed by Forns et al11 has a lower and upper cutoff limit of 4.2 and 6.9, respectively. If the assay value falls between these 2 values, it is considered an indeterminate value. Ultimately, in their study, only 51% of patients could have their fibrosis classified. In addition, many of these tests were studied in patients referred to tertiary care centers or were a part of treatment trials for hepatitis. Therefore, the performance of these assays may be different in routine clinical practice.

There are several issues that need to be kept in mind regarding serum fibrosis markers. Performance characteristics such as negative and positive predictive values are greatly affected by the baseline prevalence in the study population. In turn, the performance of these assays is dependent on the prevalence of hepatic fibrosis in the population being studied. In this study, the FIBROspect II assay was studied in patients referred to a tertiary care center. Therefore, the performance of this assay may be different in routine clinical practice. However, in clinical practice the prevalence of significant fibrosis is likely to be lower and thus, the assay will likely be more useful in excluding significant fibrosis. Also, these assays are generally good at differentiating between advanced fibrosis and minimal or no fibrosis but are poor at differentiating between the intermediate grades of fibrosis (F1-F3)—this is the case with the FibroSpect II assay as well. In addition, these serum assays can only be as good as the gold standard for determining hepatic fibrosis, the liver biopsy. And, as mentioned above, the overall sensitivity of a single pass percutaneous liver biopsy is only 80%. In this study, all biopsy samples were considered adequate and, thus, sampling error was minimized. Another diagnostic limitation of a percutaneous liver biopsy includes inter- and intra-observer variability of the fibrosis staging. This was minimized in this study by using a single hepatopathologist. In addition, when a third of the biopsies were blindly read again, there was high agreement with the initial reading (κ = 0.8). Finally, the FIBROspect II assay was validated using a fibrosis cutoff where F2-F4 was considered significant fibrosis. This cutoff is useful when determining which hepatitis C patients to treat, where typically patients with F2 or higher fibrosis are felt to be appropriate candidates for interferon-based therapy. However, in practice F2 fibrosis or less is considered mild fibrosis. As shown in the results, when the significant fibrosis cutoff was changed to F3-F4, the sensitivity of the test increased but the specificity decreased.

In summary, this study validates the performance characteristics of a noninvasive serum fibrosis marker—FIBROspect II using a prospective cohort of patients with chronic hepatitis C. Similar to other noninvasive fibrosis tests, FIBROspect II may not be useful in differentiating between the intermediate stages of fibrosis. However, the high negative predictive value of this assay in low prevalence populations of F2-F4 fibrosis may allow clinicians to avoid a liver biopsy in certain situations, such as patients who refuse a biopsy, or have a contraindication (hemophiliacs). Simple laboratory tests such as prothrombin time, albumin, total bilirubin, and platelet counts in conjunction with an abdominal imaging study should continue to be utilized to identify overt cirrhosis. Furthermore, FIBROspect II may be useful in assessing fibrosis progression in patients who have had a baseline liver biopsy in the past, including those patients with non-HCV-related liver diseases—although this must be studied prospectively before it can be recommended. It will also be important to prospectively validate these results using a community-based setting with different relative prevalence of different stages of HCV in order to further assess its clinical utility. Moreover, it will be of interest to determine how the FIBROspect II might change in patients receiving antiviral therapy.
References

Physician Preferences and Attitudes Regarding Industry Support of CME Programs

Paul S. Mueller, MD,a C. Christopher Hook, MD,b Scott C. Litin, MDa

affiliation

Division of General Internal Medicine, and "Division of Hematology, Mayo Clinic, Rochester, Minn.

Pharmaceutical and other health care-related companies spend approximately $12 to $15 billion per year ($8000-$15 000 per year, per physician) on marketing.1 One marketing approach used by many pharmaceutical companies is to provide financial support of continuing medical education (CME) programs.1,3 In recent years, this support has increased. Ten years ago, 17% of CME funding came from industry; today, that number is 40%.4,5 Between 1992 and 2001, industry support of medical school-sponsored CME quintupled.6 Organizations that conduct CME programs claim that without financial support from industry, programs must rely on registration fees, which, when combined with travel expenses, would make the programs unaffordable for many participants.2

Physicians attend CME programs for many reasons, including fulfilling state medical licensure requirements, maintaining hospital privileges and specialty society memberships, and obtaining new knowledge and skills.7,8 Many physicians also regard CME courses as their most valuable source for clinical information.7 However, evidence suggests that CME programs sponsored by industry not only may be more biased (in favor of the sponsoring companies’ products) than programs not sponsored by industry3,5 but also may influence physicians’ professional behavior (eg, increased prescriptions of the sponsor’s medication).3,5 These findings raise the ethical concern of industry influence on physicians who participate in CME programs.

CME PROGRAM SURVEY METHODS

Each year, Mayo Clinic College of Medicine School of CME (Rochester, Minn) conducts comprehensive internal medicine CME courses, some of which receive financial support from industry. Over the years, directors of these courses have received mixed feedback from participants regarding industry support. To better understand physician preferences and attitudes regarding industry support of CME programs, an anonymous survey of physicians attending 4 Mayo Clinic internal medicine CME courses was conducted in 2004. Pharmaceutical and other health care-related companies financially support 2 of the courses (Mayo Internal Medicine Board Review and Mayo Clinical Reviews) with unrestricted grants and in accordance with Accreditation Council for Continuing Medical Education (AC-CME) guidelines. Financial support is given directly to the course organizers and is used to defray the costs of the courses; the companies have no role in planning or conducting the courses. These courses allow industry-sponsored exhibits adjacent to the meeting room. The other 2 courses (Mayo Selected Topics in Internal Medicine and Mayo Practice of Internal Medicine) are not supported by industry and have no exhibitors.

Respondents completed the survey during a break period in the middle of each course. The one-page survey instrument asked participants their age, sex, and years in practice. The survey also included 4 specific questions regarding industry support of CME activities. The results of the survey comprise the data set of this study.

Between-group responses to questions were compared using the Pearson chi-squared test. Ordinal logistic regression was used to fit multivariate models. The response ordering was “yes,” “no,” and “no preference”
A P value less than .05 was considered significant. All analyses were conducted using JMP 4.0.4 software (SAS Institute, Inc., Cary, NC). Permission to perform an analysis of the surveys was granted by the Mayo Clinic Institutional Review Board in accordance with federal regulations.

RESULTS

Of 1603 physicians attending the courses, 1130 (70.5%) completed the survey. Most of the 1603 attended only one course; only 19 (1.2%) attended 2 courses, and none attended 3 or 4 courses. (It is unknown whether the physicians who attended 2 courses completed the survey once, twice, or at all). Of the 1130 survey respondents, 671 (59.4%) attended a course sponsored by industry.

Table 1 highlights the characteristics of the participants who completed the survey. In the multivariate analyses, age and years in practice were found to highly relate to each other. Therefore, the rest of the article only presents data related to years in practice.

In response to the question “What type of CME course (industry supported or not) do you prefer to attend?”, 58.3% of respondents indicated no preference (Table 2). Among those physicians who indicated a course preference, the majority preferred non-industry-supported courses. Responses to this question differed significantly by years in practice (P < .001). Experienced physicians (in practice more than 30 years) preferred industry-supported courses more than less experienced physicians. Responses also differed significantly by course attended (P < .001). More participants of non-industry-supported courses preferred that type of course than participants of industry-supported courses. In the multivariate model, more than 30 years in practice (as opposed to 10 or fewer years in practice) and attending an industry-supported course were independent predictors of response to this question (P < .001 for both).

In regard to preference for a CME course with or without exhibitors (question 2), one-half of the respondents indicated no preference, whereas the other one-half was evenly divided (24.8% each) in their preference for courses with or without exhibitors (Table 3). Responses to this question differed significantly by sex and years in practice. More men preferred courses with exhibitors, whereas more women preferred courses without exhibitors (P = .005). Compared with more experienced respondents, physicians in practice fewer than 31 years preferred courses without exhibitors (P < .001). Responses also differed significantly by course attended (P < .001). More participants of non-industry-supported courses preferred courses without exhibitors. In the multivariate model, more than 20 years in prac-

| Table 1 Characteristics of 1130 Physician Attendees Who Completed the Survey |
|---------------------------------|-----------------|-----------------|-----------------|
| Characteristic                  | No. of Respondents (%) |
| Sex                            | 1015 (89.8)       |
| Male                           | 786 (77.4)        |
| Female                         | 229 (22.6)        |
| Age, years                     | 1121 (99.2)       |
| ≤ 40                           | 189 (16.9)        |
| 41-50                          | 331 (29.5)        |
| 51-60                          | 324 (28.9)        |
| ≥ 61                           | 277 (24.7)        |
| Years in practice              | 1030 (91.2)       |
| 0-10                           | 204 (19.8)        |
| 11-20                          | 288 (28.0)        |
| 21-30                          | 289 (28.1)        |
| ≥ 31                           | 249 (24.1)        |
| Type of course attended        | 1130 (100)        |
| Not supported by industry      | 459 (40.6)        |
| Supported by industry          | 671 (59.4)        |

<table>
<thead>
<tr>
<th>Table 2 Responses to Question 1—Preferred Course to Attend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic of Respondents</td>
</tr>
<tr>
<td>All respondents (n = 1121)</td>
</tr>
<tr>
<td>Sex (n = 1009)</td>
</tr>
<tr>
<td>Male (n = 783)</td>
</tr>
<tr>
<td>Female (n = 226)</td>
</tr>
<tr>
<td>Years in practice (n = 1022)</td>
</tr>
<tr>
<td>0-10 (n = 202)</td>
</tr>
<tr>
<td>11-20 (n = 285)</td>
</tr>
<tr>
<td>21-30 (n = 287)</td>
</tr>
<tr>
<td>≥ 31 (n = 248)†</td>
</tr>
<tr>
<td>Type of course attended (n = 1121)</td>
</tr>
<tr>
<td>Not supported by industry (n = 456)</td>
</tr>
<tr>
<td>Supported by industry (n = 665)†</td>
</tr>
</tbody>
</table>

*Pearson chi-squared test.
†Significant predictor of response in the multivariate model.
tice (as opposed to 10 or fewer years in practice) and attending an industry-supported course were independent predictors of response to this question (all $P < .001$). Sex was not a predictor in the multivariate model.

In response to question 3, “Do you believe CME courses should accept industry support if doing so reduces the overall cost of the course?”, 62.3% of respondents answered “yes” (Table 4). Responses differed significantly by sex, with more women answering “no” or “no preference” ($P = .02$), and by years in practice, with increasing years of experience directly associated with answering “yes” ($P = .001$). Responses to this question also differed significantly by course attended ($P < .001$). A majority of industry-supported course attendees (71.5%) answered “yes,” whereas the attendees of the non-industry-supported courses were more evenly divided (48.7% answered “yes,” 33.1% answered “no”). However, in the multivariate model, only attending an industry-supported course was an independent predictor of response to this question ($P < .001$).

For the final question, “Is it your impression that the contents of CME courses supported by industry tend to be biased in favor of the supporting companies?”, 53% of participants responded “no” (Table 5). Responses to this question differed significantly by sex, years in practice, and course attended. More men than women answered “no” ($P = .01$). Compared with physicians in practice longer than 30 years, a greater percentage of less experienced respondents answered “yes” ($P = .001$). A majority of non-industry-supported course participants (52.7%) answered “yes,” whereas a majority (63.7%) of the participants of industry-supported courses answered “no” ($P < .001$). In the multivariate model, only 21 to 30 years in practice and attending an

### Table 3 Responses to Question 2—Preference for Exhibitors

<table>
<thead>
<tr>
<th>Characteristic of Respondents</th>
<th>With Exhibitors</th>
<th>Without Exhibitors</th>
<th>No Preference</th>
<th>$P$ Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>All respondents (n = 1121)</td>
<td>24.8</td>
<td>24.8</td>
<td>50.4</td>
<td>.005</td>
</tr>
<tr>
<td>Sex (n = 1108)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n = 781)</td>
<td>26.6</td>
<td>22.4</td>
<td>51.0</td>
<td></td>
</tr>
<tr>
<td>Female (n = 227)</td>
<td>18.9</td>
<td>31.7</td>
<td>49.3</td>
<td></td>
</tr>
<tr>
<td>Years in practice (n = 1022)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>0-10 (n = 202)</td>
<td>20.3</td>
<td>29.2</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>11-20 (n = 286)</td>
<td>17.8</td>
<td>30.8</td>
<td>51.4</td>
<td></td>
</tr>
<tr>
<td>21-30 (n = 285)†</td>
<td>17.2</td>
<td>25.3</td>
<td>57.5</td>
<td></td>
</tr>
<tr>
<td>≥31 (n = 249)†</td>
<td>45.8</td>
<td>11.2</td>
<td>43.0</td>
<td></td>
</tr>
<tr>
<td>Type of course attended (n = 1121)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Not supported by industry (n = 456)</td>
<td>15.1</td>
<td>43.9</td>
<td>41.0</td>
<td></td>
</tr>
<tr>
<td>Supported by industry (n = 665)†</td>
<td>31.4</td>
<td>11.7</td>
<td>56.8</td>
<td></td>
</tr>
</tbody>
</table>

*Pearson chi-squared test.
†Significant predictor of response in the multivariate model.

### Table 4 Responses to Question 3—Should Industry Support Be Used to Reduce Costs?

<table>
<thead>
<tr>
<th>Characteristic of Respondents</th>
<th>Yes</th>
<th>No</th>
<th>No Preference</th>
<th>$P$ Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>All respondents (n = 1126)</td>
<td>62.3</td>
<td>23.4</td>
<td>14.3</td>
<td>.02</td>
</tr>
<tr>
<td>Sex (n = 1011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n = 784)</td>
<td>64.5</td>
<td>21.9</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Female (n = 227)</td>
<td>54.2</td>
<td>27.8</td>
<td>18.1</td>
<td></td>
</tr>
<tr>
<td>Years in practice (n = 1027)</td>
<td></td>
<td></td>
<td></td>
<td>.001</td>
</tr>
<tr>
<td>0-10 (n = 204)</td>
<td>55.4</td>
<td>28.9</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>11-20 (n = 286)</td>
<td>60.1</td>
<td>27.3</td>
<td>12.6</td>
<td></td>
</tr>
<tr>
<td>21-30 (n = 288)†</td>
<td>62.9</td>
<td>22.6</td>
<td>14.6</td>
<td></td>
</tr>
<tr>
<td>≥31 (n = 249)†</td>
<td>72.3</td>
<td>13.3</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>Type of course attended (n = 1126)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Not supported by industry (n = 456)</td>
<td>48.7</td>
<td>33.1</td>
<td>18.2</td>
<td></td>
</tr>
<tr>
<td>Supported by industry (n = 670)†</td>
<td>71.5</td>
<td>16.9</td>
<td>11.6</td>
<td></td>
</tr>
</tbody>
</table>

*Pearson chi-squared test.
†Significant predictor of response in the multivariate model.
industry-supported course were independent predictors of response to this question ($P < .001$ each).

**DISCUSSION**

This survey was designed to assess physician preferences and attitudes regarding industry support of CME programs. A majority of participants indicated no preference for industry-supported or non-industry-supported courses, and one-half indicated no preference for courses with or without exhibitors. A majority believed CME courses should accept industry support to reduce the cost of courses, and about one-half thought that courses supported by industry were not biased. Organizers of CME programs should find these results interesting and helpful in planning future courses.

Results indicate that the preferences and attitudes regarding industry support of CME programs vary significantly according to the number of years in practice and type of course attended (industry supported or not). Compared with less experienced physicians, those in practice more than 30 years were more likely to prefer industry-supported courses and courses with exhibitors and believed CME courses should accept industry support. Several explanations most likely account for these findings. Physicians who have been in practice for many years have had more opportunities to attend industry-supported CME programs and interact with industry representatives and may believe they are not influenced by these interactions. Also, these physicians may enjoy the interactions with industry representatives and derive educational value for themselves and benefits for their patients (such as drug samples) from industry-supported CME programs.

In contrast, physicians in practice fewer years, who were more likely to prefer non-industry-supported courses, may be more skeptical of industry-sponsored CME programs. These attitudes may partly be a result of recent efforts by medical schools, residency programs, professional societies, and organizations,\textsuperscript{9,10} such as No Free Lunch,\textsuperscript{11} to raise awareness of potential conflicts of interest in physician–industry interactions.

Responses also varied significantly according to type of course attended. More physician respondents who attended non-industry-supported courses than industry-supported courses believed that industry bias exists. Several reasons may account for this finding. Some participants of non-industry-supported CME programs may consciously avoid industry-supported programs, and some respondents likely took into account the type of course they were attending and answered the questions to justify their choice.

A majority of the survey respondents—including nearly one-half the participants of non-industry-supported courses—believed CME courses should accept industry support if doing so reduces the overall cost of the course. Although it has been reported that gifts, regardless of the size, may instill in the physician recipient a sense of obligation to reciprocate,\textsuperscript{1,12} it is unknown whether industry-defrayed costs of attending CME programs engender a similar obligation.

Regardless of physician preferences and attitudes, financial support of CME programs by industry is likely to continue. How should concerns about industry support of CME be addressed? First, CME organizers should explicitly acknowledge that the primary function of CME is to improve the welfare of patients by enhancing the knowledge and skills of physicians,\textsuperscript{6} not to facilitate physician–industry interactions. Several professional organizations, including the American Medical Association,\textsuperscript{10} the ACCME,\textsuperscript{13} and the American College of Physicians,\textsuperscript{9} have made recommendations regarding industry support of CME programs.

<table>
<thead>
<tr>
<th>Characteristic of Respondents</th>
<th>Response by Percentage</th>
<th>$P$ Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>All respondents (n = 1120)</td>
<td>Yes: 35.9</td>
<td>53.0</td>
</tr>
<tr>
<td>Sex (n = 1006)</td>
<td>Female (n = 224)</td>
<td>41.5</td>
</tr>
<tr>
<td></td>
<td>Male (n = 782)</td>
<td>33.1</td>
</tr>
<tr>
<td>Years in practice (n = 1024)</td>
<td>0-10 (n = 202)</td>
<td>43.6</td>
</tr>
<tr>
<td></td>
<td>11-20 (n = 287)</td>
<td>33.5</td>
</tr>
<tr>
<td></td>
<td>21-30 (n = 287)†</td>
<td>41.8</td>
</tr>
<tr>
<td></td>
<td>≥31 (n = 248)</td>
<td>24.2</td>
</tr>
<tr>
<td>Type of course attended (n = 1120)</td>
<td>Not supported by industry (n = 457)</td>
<td>52.7</td>
</tr>
<tr>
<td></td>
<td>Supported by industry (n = 663)†</td>
<td>24.3</td>
</tr>
</tbody>
</table>

*Pearson chi-squared test.
†Significant predictor of response in the multivariate model.
Taken together, these recommendations can be summarized as follows:

- Industry support must be completely unrestricted;
- All CME faculty conflicts of interest must be declared before the program begins;
- The industry sponsor should have no role in the planning or evaluation of program content;
- The topics should be presented without bias, particularly if the products of the industry sponsor are discussed; and
- Support should not be given to participants but to the program organizers to reduce registration fees.

Adherence to these recommendations should prevent most inappropriate industry influence or bias in CME programs.

This study has several limitations. Although the number of respondents was large and the response rate excellent, the survey itself was brief, which precluded an in-depth examination of physician attitudes regarding industry support of CME programs. In addition, participants of courses organized and conducted by Mayo Clinic College of Medicine School of CME, Rochester, Minn, were the only group surveyed. However, physicians attend these courses from nearly every state in the United States and a number of different countries. Nevertheless, results may not be generalizable to all physicians (such as individuals who do not attend courses organized and conducted by Mayo Clinic or physicians who attend noninternal medicine courses). Also, the survey assessed participant preferences and attitudes regarding industry support of CME programs, not the actual effects of industry support on programs.

ACKNOWLEDGMENTS

The authors thank Kelley M. Sandvik for her assistance in gathering data. Editing, proofreading, and reference verification were provided by the Section of Scientific Publications, Mayo Clinic.

References

Risedronate-induced Hepatitis

To the Editor:

Risedronate, a bisphosphonate, has been approved by the U.S. Food and Drug Administration for treatment and prevention of osteoporosis, as well as for treatment of Paget's disease. This case describes an 81-year-old woman who developed considerable hepatitis after receiving long-term risedronate therapy.

The patient began risedronate therapy in 2001 after bone mineral densitometry revealed postmenopausal osteoporosis, with normal liver function test results (aspartate aminotransferase and alanine aminotransferase). There was no history of liver or renal disease and no use of illicit substances. Ethanol use was minimal, with no more than 3 drinks per week. She was converted from a once-daily (5 mg) to once-weekly (35 mg) preparation of risedronate in November 2003.

Mild elevation of aspartate aminotransferase and alanine aminotransferase (less than twice normal) was noted in May 2004 and worsened to greater than twice normal by November 2004. She was asymptomatic. The patient was also receiving fluvastatin, metoprolol, acetaminophen (1 g 3 times daily), aspirin, and calcium carbonate.

Fluvastatin was discontinued on November 17, 2004. As depicted in Figure 1, the patient's transaminase levels continued to increase. Metoprolol and acetaminophen were withdrawn on January 6, 2005. Further testing, including bilirubin (total and direct), iron studies, ceruloplasmin, serum copper, urine protein electrophoresis, viral hepatitis serologies, alpha-fetoprotein, carbohydrate antigen 19-9, alpha-1-antitrypsin, anti-smooth muscle antibody, and antimitochondrial antibody, showed normal results. A small M-spike was noted on serum protein electrophoresis (0.47 g/dL) with negative serum and urine protein immunoelectrophoreses, and anti-nuclear antibody was positive with normal sedimentation rate. Alkaline phosphatase was mildly elevated to 233 µL (55-142 µL); fractionation revealed liver predominance. Hepatobiliary imaging with computed tomography noted no hepatic pathology. Liver biopsy was undertaken, revealing portal and lobular hepatitis, moderately active, with granulomatous features and mild steatosis, suggestive of a drug effect. Trichrome stain was negative for fibrosis. Iron stain showed minimal hemosiderosis, mainly within Kupffer cells.

Risedronate was discontinued on January 7, 2005. Subsequent improvement of transaminases was not noted until more than 4 weeks later, and complete resolution of liver function required approximately 12 months. The duration of transaminase elevation suggests that risedronate was responsible, given its long terminal half-life of 220 hours.1

Bisphosphonate hepatotoxicity is rare, with several case reports previously reported with alendronate.2-5 The underlying mechanism is unknown.3 A review of the literature (PubMed and Medline) revealed no prior reports of risedronate-induced liver injury. This case suggests that bisphosphonate hepatitis may be a class effect and may be prolonged in patients receiving extended-release preparations, because prior cases associated with daily alendronate use resolved by 3 months. Periodic liver function test monitoring should be considered for patients receiving risedronate.

Michael B. Phillips, MD
Mayo Clinic
Jacksonville, Florida
doi:10.1016/j.amjmed.2006.04.032

References


Requests for reprints should be addressed to Michael B. Phillips, MD, Department of Internal Medicine, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL 32259.

E-mail address: mbphillips@myway.com
Figure 1  Transaminase values with risedronate.
ADAPT Trial Data

To the Editor:

We were surprised to see data from the Alzheimer’s Disease Anti-inflammatory Prevention Trial (ADAPT) prominently featured in the recent article by Salpeter et al entitled “Meta-analysis: Cardiovascular Events Associated with Nonsteroidal Anti-inflammatory Drugs.”1 The data did not come from an ADAPT publication or from unpublished data provided by us for the purpose of publication. We therefore are unable to vouch for the accuracy of the ADAPT data as reported in the aforementioned publication because we know neither the source of the data nor the cutoff date for the counts summarized.

The authors state that they “sent 27 correspondences and received 8 responses; unpublished information was received for 5 studies.” The implication is that those requests produced data for the 5 studies listed as “unpublished.” In fact, we received an e-mail inquiry from Dr. Salpeter in January 2005 but responded that we could not grant permission to use ADAPT data in a meta-analysis until they had been published. We note the absence of the word “received” with respect to ADAPT data in the comments column of the table listing the studies included. The reference provided for ADAPT is a National Institutes of Health press release, which does not contain the data published.

The appropriation and publication of our data by Salpeter et al might well have jeopardized our own efforts at publication2 if the journal to which we had submitted the data rejected our manuscript because of The American Journal of Medicine publication. We also now must give account to our investigators and sponsors because of this publication, and will likely encounter anger and irritation from our study participants and IRBs, to whom we promised advance notice of our results before publication.

We believe that authors of meta-analyses have responsibilities to reveal their sources of unpublished data for readers. We believe further that journals have a responsibility to refuse publication of their meta-analyses involving unpublished data absent written assurance from the authors that the data are being used with consent of the investigators who collected them. Such action by journals would be consistent with journal policies to refuse publication of single-study results obtained without proper consent by individuals.

Officers of ADAPT:

John Breitner, MD, MPH
Chair of ADAPT
VA Puget Sound Health Care System and Department of Psychiatry and Behavioral Sciences University of Washington
Seattle, Wash

Denis Evans, MD
Chair, ADAPT Steering Committee
Rush Institute for Healthy Aging Rush University Chicago, Ill

Constantine Lyketsos, MD, MPH
Field Site Director
Department of Psychiatry and Behavioral Sciences Johns Hopkins University School of Medicine Baltimore, MD

Barbara Martin, PhD
Deputy Director, ADAPT Coordinating Center

Curtis Meinert, PhD
Director, ADAPT Coordinating Center Center for Clinical Trials Department of Epidemiology Johns Hopkins University Bloomberg School of Public Health Baltimore, MD

doi:10.1016/j.amjmed.2006.09.022

References
The Reply:

We are responding to the letter from Dr. Breitner and colleagues concerning our meta-analysis on cardiovascular events associated with nonsteroidal anti-inflammatory drugs (NSAIDs). The investigators of the Alzheimer’s Disease Anti-inflammatory Prevention Trial (ADAPT) contend that the data we used from their trial did not come from information provided by them, and therefore they cannot vouch for the accuracy of the data used. They are confused by our statement that unpublished information was used, because they cannot recall providing us with the data.

We would like to clarify that the ADAPT investigators did, in fact, provide one of us (E.J.T.) with the unpublished data for an invited commentary in January 2005, shortly after their preliminary results were publicized. The information was reported to be accurate as of December 2, 2004. In a separate communication, Dr. Salpeter wrote to Dr. Breitner requesting information on cardiovascular events in the trial, to be included in a meta-analysis. He responded that they were rushing to get their article published at that time. In February 2005, the ADAPT steering committee reported to the Federal Drug Administration Joint Advisory Committee that they would provide their full data in a publication planned for the near future. Dr. Breitner informed us that the New England Journal of Medicine had invited the ADAPT article for a February 2005 special issue publication. We assumed that their trial results would be published long before our meta-analysis finally appeared in July 2006.

We apologize for not notifying the ADAPT group of our intent to include the data that we had received from them in our analysis. This was an oversight on our part and was purely unintentional. We had never been told that we could not use their data until after it was published. It is standard practice when performing a meta-analysis to obtain as much published and unpublished information as possible, in order to limit the potential for publication bias and to increase statistical power. We reported that unpublished information was used for their trial and included a reference to the press release that announced the preliminary results. We have performed several meta-analyses to date, all with unpublished information; we have never been asked to provide approval from trial investigators to use their data and have not seen that requirement mentioned in meta-analysis guidelines.

Dr. Breitner and colleagues wonder if our publication of ADAPT data could affect their ability to publish their full results. We would like to reiterate that we had no intention of jeopardizing their chances for publication. And it is now apparent that our meta-analysis, which reported cardiovascular events from their trial, did not prevent them from getting their results accepted for publication in the Public Library of Science. In fact, in the context of how much impact the ADAPT trial had in shaping NSAID public policy and the cessation of other trials, we feel that the investigators have an essential responsibility to the medical community and the public of reporting their complete data in a timely fashion.

Shelley R. Salpeter, MD
Stanford University School of Medicine
Stanford, Calif
Santa Clara Valley Medical Center
San Jose, Calif

Eric J. Topol, MD
Scripps Research Institute
La Jolla, Calif

doi:10.1016/j.ajm.2006.10.007
LETTER

The Reply:

The American Journal of Medicine greatly regrets the 2 letters printed above. Clearly, a serious misunderstanding evolved between the 2 groups of outstanding and honorable investigators, resulting in this exchange of letters. In order to prevent such unhappiness in the future, The American Journal of Medicine will add the following guideline to our instructions to authors: “When manuscripts are submitted that report unpublished data, the editors of AJM will require that the authors of the submitted manuscript confirm that they have received written permission to use this data from the individual or group of investigators who generated the data.”

Joseph S. Alpert, MD
Editor-in-Chief

The American Journal of Medicine
Robert S. and Irene P. Flinn Professor of Medicine
Special Assistant to the Dean
University of Arizona College of Medicine
Tucson, Ariz

doi:10.1016/j.amjmed.2006.10.006
To the Editor:

We read with interest the recent article by Sohail et al. concerning the management of Staphylococcus prostatic valve endocarditis (PVE). This infection causes difficult management issues for patients and vexing therapeutic challenges for their health care providers. The study was a retrospective review, which definitely contributed to its limitation. In addition, the most recent guidelines for treatment of PVE were followed in less than 17% of cases. Triple antibiotic therapy (6 weeks of beta lactam or vancomycin plus rifampin and 2 weeks of aminoglycoside) has been recommended by the American Heart Association and European Society of Cardiology for the treatment of PVE caused by S. aureus. Recent studies showed that patients receiving combination therapy were 5.9 times more likely to be culture-negative than those receiving monotherapy and that those treated for more than 14 days were more likely to be culture-negative than those treated for 14 days or less. Unfortunately in this study, less than 50% of patients in the medically treated group were given dual therapy, and less than 17% received appropriate triple antibiotic treatment. In addition, methicillin-resistant S. aureus is increasingly more common, whereas all organisms were methicillin-sensitive in this study. Further studies are necessary to establish the effectiveness of triple antibiotic therapy versus surgical therapy in S. aureus PVE.

Ali Hassoun, MD
Alabama Infectious Diseases Center
Huntsville

doi:10.1016/j.amjmed.2006.02.021

References
The Reply:

We appreciate the insightful comments of Dr. Hassoun. His foremost concern is the limited use of triple antibiotic therapy in our cohort and its potential impact on the outcome. The principle reason why a combination of a beta-lactam or vancomycin plus rifampin for 6 weeks and gentamicin for 2 weeks was not uniformly administered was due to the period of study. Our cohort was seen between 1980 and 2000, and guidelines from the American Heart Association that provided an initial recommendation for triple antibiotic combination were not published until December 1995.

Gentamicin in combination with a beta-lactam agent or vancomycin was used in 39 (71%) of our study patients. However, it was not continued for more than 5 days in 11 (20%) patients due to serum creatinine elevation.

It is noteworthy that the recommendation for triple combination therapy in *S. aureus* prosthetic valve endocarditis (PVE) is based on the efficacy of this treatment strategy for coagulase-negative staphylococcal PVE and treatment results of experimental endocarditis and infected devices. Although there are data that suggest combination therapy use in *S. aureus* PVE can sterilize the blood stream more quickly than does monotherapy, its impact on mortality is unproven. All but 3 patients in this report received combination antibiotic therapy with only 2 agents, except in 3 cases where triple combination therapy was administered.

In conclusion, triple combination therapy is currently the recommended regimen for patients with *S. aureus* PVE and is consensus-based. Its impact on mortality has yet to be examined in prospectively conducted clinical trials.

Muhammad R. Sohail, MD
Walter R. Wilson, MD

References
Orthostatic Intolerance in Chronic Fatigue Syndrome

To the Editor:

The main outcome measure in Jones and colleagues’ study of the prevalence of orthostatic intolerance in subjects with chronic fatigue syndrome (CFS)\(^1\) was a 45-minute head-up tilt test, performed after exclusion of subjects with other medical conditions and subjects being treated with medications used to treat orthostatic intolerance. These exclusions were methodologically necessary to ensure that the observed rate of orthostatic intolerance was associated with CFS and not with the co-morbid medical conditions, and that medication use did not partially treat (and therefore obscure detection of) the underlying circulatory abnormalities.

The exclusions came at a steep methodologic cost to the representativeness of the sample. After the exclusion of 41 subjects, and refusal to participate by 23 more, the tilt portion of the study evaluated 10 subjects, just 14% of the original group of 74. The sample was entirely too small to allow firm conclusions to be drawn about the prevalence of orthostatic intolerance overall in those with CFS, or about the relative prevalence of postural tachycardia or neurally mediated hypotension in this group.

While not the main focus of the article, Table 7 summarizing the literature on orthostatic intolerance in those with CFS is incomplete. In addition to the small study by De Lorenzo and colleagues of 5 patients with postural tachycardia (reference 10 in the Table), the authors may not have known of a larger study by the same group. That study enrolled 78 subjects with CFS, 22 of whom (28%) developed hypotension during tilt, versus 0 of 38 controls.\(^2\)

We would also like to clarify the data abstracted in Table 7 from our randomized trial of fludrocortisone (reference 13). Among 171 with CFS who underwent tilt testing, the rate of neurally mediated hypotension was 62%; a further 4% met criteria for postural tachycardia syndrome alone. The combined prevalence of orthostatic intolerance was 66%, not 62% as reported in Table 7. Those who did not develop hypotension were excluded from the treatment portion of the randomized trial, but none of the 171 eligible subjects was excluded from the data on the prevalence of orthostatic abnormalities. We therefore are not sure what the authors meant by the notation “77% excluded.” Table 7 also includes some minor typographical errors: reference 20 appears twice (it should be reference 21 the first time, then reference 8 the second time), and reference 34 in Table 7 should be reference 37.

Any debate about the precise prevalence of postural tachycardia syndrome or neurally mediated hypotension in CFS has the potential to neglect a critical point. In our studies of orthostatic intolerance, now totaling over 220 subjects with CFS, quiet upright posture has been a strong and consistent physiologic stressor in over 95%. Even when not accompanied by hemodynamic changes, orthostatic stress typically has been associated with a provocation or exacerbation of characteristic CFS symptoms. Others have shown that these symptoms and hemodynamic abnormalities with orthostatic stress can be reversed upon application of external lower-body compression (reference 37). Whether orthostatic disorders are primary or secondary, this evidence suggests that they play an important contributing role in the phenomenology of the illness for a substantial proportion of affected individuals. In contrast to Jones and colleagues, we would agree with Freeman that a better understanding of the mechanisms of the disordered response to orthostatic stress would go a long way toward improving our understanding of the pathophysiology and treatment of CFS symptoms.\(^3\)

Peter C. Rowe, MD
Department of Pediatrics
Johns Hopkins University School of Medicine
Baltimore, Md

Katherine E. Lucas, PhD
Department of Pediatrics
Johns Hopkins University School of Medicine
Department of Epidemiology
Johns Hopkins University Bloomberg School of Public Health
Baltimore, Md

doi:10.1016/j.amjmed.2006.02.033

References
LETTER

Asian Indians and Coronary Artery Disease Risk

To the Editor:

Being a physician from an ethnic minority myself, I read with great interest the recent article by Larosa and Brown, “Cardiovascular risk factors in minorities.”1 Asian Indians, with an estimated population of 1 million in the United States, including 35,000 physicians, deserve special mention due to a number of unique characteristics. Asian Indians (Indians, Pakistani, Bangladeshi, Sri Lankans) have substantially higher rates, early onset, and high mortality rate from coronary artery disease.2 According to World Health Organization estimates, by 2010, Asian Indians will represent 60% of the world’s cardiac patients. Expatriate Indians in their newly adopted countries have 3-5 times more chance of developing coronary artery disease than native population or other immigrant groups.3 Urban immigration within India itself is associated with increased coronary artery disease risk.4 Despite a high rate of coronary artery disease among Asian Indians, prevalence of smoking, hypertension, and obesity remains low in comparison with developed countries, suggesting the role of altered diet and lifestyle factors. The literature in the past decade points to a number of mechanisms for this increased risk.

- Although total cholesterol and low-density lipoprotein cholesterol (LDL-C) are not elevated, in comparison with Caucasians, levels of Lp(a) are high.5
- Percentage of favorable large high-density lipoprotein (HDL) particles are low,6 and smaller size HDL particles are more; HDL-2b particle concentration is low despite having normal total HDL measurements (>40 mg/dL), both studies pointing toward a problem with reverse cholesterol transport.7
- Greater sensitivity for saturated fat consumption, as Indian immigrants in Great Britain have less total fat and saturated fat consumption (38.8% and 13.7% of energy consumed) when compared with British natives (42.2% and 18.5%) but coronary death rates are 40% higher.8 Total and saturated fat and cholesterol consumed by urban Indians are much lower than in developed countries, however, the serum cholesterol of urban Indians is not proportionally lower.9
- Greater visceral fat and abdominal adiposity, even with body mass index in the normal range, indicating insulin resistance.3
- Smaller size of the coronary arteries when compared with Caucasians, even after correcting for body surface area.10

Considering the rising incidence and prevalence of coronary artery disease among migrating Indians both urban and transcontinental, it is tempting to speculate a lower threshold than affluent societies for susceptibility to coronary artery disease.

Sastry Prayaga, MD
SUNY Upstate Medical University
Syracuse, NY
doi:10.1016/j.amjmed.2006.02.042

References

Obstructive Sleep Apnea in Minorities

To the Editor:

We read with great interest the recent review by LaRosa and Brown1 dealing with the analysis of cardiovascular disease risk factors in minority groups. Our comments are focused on the role of obstructive sleep apnea as one other risk factor for cardiovascular disease, also frequently affecting minorities. Obstructive sleep apnea is characterized by recurrent episodes of cessation of respiratory airflow during sleep secondary to upper airway collapse on inspiration. Recently obstructive sleep apnea was identified as an independent risk factor for fatal and nonfatal cardiovascular events.2 Established risk factors for obstructive sleep apnea include: male gender, increasing age, overweight and obesity, and craniofacial and upper airway abnormalities. Smoking, menopause, alcohol use before sleep, nighttime nasal congestion, and ethnicity also are suggested to be determinants on the incidence and severity of obstructive sleep apnea. This primary sleep disorder is common and affects approximately 2% and 4% of middle-aged men and women in the general population,3 although the prevalence of obstructive sleep apnea has been mainly estimated in white populations. Several newer studies have demonstrated a similar prevalence of obstructive sleep apnea in subjects from different ethnic groups.4-6 However, the importance of established risk factors for obstructive sleep apnea is different among ethnic groups. A higher prevalence of obesity is related to an increased prevalence of obstructive sleep apnea among American Indians and Hispanics.7 In contrast, obesity seems not related to increased prevalence of obstructive sleep apnea in Asian populations, but it is explained by craniofacial structure peculiarities. Compared with whites, younger African-Americans present a higher prevalence of and more severe obstructive sleep apnea.8 An increased severity of obstructive sleep apnea among elderly African-Americans compared with white populations has been reported.9 In summary, obstructive sleep apnea is a well-recognized cardiovascular risk factor affecting all ethnic groups equally, in terms of prevalence of disease, but a wide range of ethnic differences influence the role of the established risk factors for obstructive sleep apnea in its contribution to disease presentation and severity for each ethnic group.

Miguel A. Arias, MD, PhD
Department of Cardiology
Complejo Hospitalario de Jaén
Jaén, Spain

Alberto Alonso-Fernández, MD, PhD
Department of Neumology
Hospital Universitario Son Dureta
Palma de Mallorca, Spain

Francisco García-Río, MD, PhD
Department of Neumology
Hospital Universitario La Paz
Madrid, Spain

doi:10.1016/j.amjmed.2006.01.026

References
The comments of Arias et al about sleep apnea should remind us to consider the increased risk of coronary disease in such patients. Although this condition does not rise to the level of the major coronary risk factors, its antecedents and concomitants are shared with other coronary risk factors, and physicians should be alerted, as Dr. Arias suggests, to look for evidence of vascular disease in such patients.

Dr. Prayaga correctly points out the unique susceptibility of South Asian populations to develop coronary atherosclerosis when exposed to Western diets. In fact, in the first iterations of our article, we devoted a separate section to this topic, making many of the same important points that Dr. Prayaga presents. Unfortunately, the manuscript was judged to be too lengthy and that section dropped by the wayside. We are grateful to Dr. Prayaga for resurrecting the topic and agree with her comments. In fact, the heightened susceptibility of populations in developing countries, as they become “westernized,” to metabolic syndrome and atherosclerosis, is worthy of its own, separate review.

John C. LaRosa, MD
Clinton D. Brown, MD
SUNY Downstate Medical Center
Brooklyn, NY

doi:10.1016/j.amjmed.2006.05.036
Outcomes and Physician Specialty Among Patients with Heart Failure: Association Is Not Causation

To the Editor:

The article by Foody and colleagues addresses an intricate issue regarding outcomes among elderly patients hospitalized with heart failure depending on the specialty of their attending physician. The authors conclude that their study demonstrates an excess risk of 30-day mortality when older patients with heart failure are cared for by family physicians and that a cardiology consultation did not reduce this risk.

We think that, according to the principles of scientific reasoning, a single study does not allow anyone to demonstrate the truthfulness of a hypothesis or an association. A single study, when properly designed and conducted, can only provide some clarification, a fact that is based on the logic of causation. A theory or a hypothesis is demonstrated only when the study proves to have validity and when all criteria for causation are fulfilled, one of which is consistency, that is, when different studies, carried out with different individuals in different circumstances bring consistent results.

The study by Foody and colleagues makes it difficult for the reader to answer several questions closely related to the assessment of its internal validity. First, patients treated by cardiologists were younger, had fewer comorbid conditions, and were more likely to be admitted to large urban teaching hospitals than patients cared for by other physicians. To take this into account, mortality was adjusted for different variables, but with the exception of age, it seems that it was not adjusted for important potential confounders, such as comorbidity (we do not think that “clinical variables” accounted for comorbidity in the adjustments), illness severity, and treatment compliance.

Second, according to the adjusted relative risks of 30-day mortality that were estimated, the risk of patients treated by cardiologists was lower only with regard to the risk of those treated by family physicians (relative risk = 1.31) because, strictly speaking, it is the only estimate with a 95% confidence interval that does not overlap 1.0 (95% confidence interval, 1.16-1.49).

Third, the excess risk of 30-day mortality of patients cared for by family physicians compared with those treated by cardiologists is of an order of magnitude between 16% and 49%, not much higher, a result attained with reasonable precision given the high number of patients included in the risk estimation.

On the basis of these 3 facts, in addition to the limitations stated by the authors, we think that this study only brings about some data that are consistent with an excess risk of 30-day mortality when older patients with heart failure are cared for by family physicians and, therefore, does not demonstrate it yet. Decision making based on preliminary results may be harmful for patients.

Carlos Campillo, MD, MPH, PhD
Balearic Health Service
Palma de Mallorca, Spain

Jordi Forteza-Rey, MD, PhD
Department of Internal Medicine
Son Dureta University Hospital
Palma de Mallorca, Spain

doi:10.1016/j.amjmed.2006.03.043

References
LETTER

The Reply:

Drs. Campillo and Forteza-Rey raise important issues regarding the interpretation of any observational study—association does not infer causation. As clearly stated throughout our manuscript, observational data are inherently limited and never serve to prove causation. However, based on analysis of the largest sample of older patients hospitalized with heart failure available utilizing a robust risk-adjustment approach, we continue to believe that our study “demonstrates an excess risk of 30-day mortality when older patients with heart failure are cared for by family physicians and that cardiology consultation did not reduce this risk.” In fact, what we have shown through rigorous assessment of a large, robust dataset is that patients cared for by family physicians have a statistically higher 30-day mortality, estimated at 31% higher than nonfamily physicians.

While causation cannot be inferred, our article highlights that quality of care is different for heart failure patients being cared for by family practitioners compared with internists or cardiologists, and that these differences are associated with differences in outcomes.

We believe that given the specialty-related disparities in care and outcomes, further studies are warranted so that further insights can be gained regarding the relationship between physician specialty and outcomes. The overriding objective of this line of research is to identify barriers to high quality care so that all patients with heart failure may reap the benefits of evidence-based care and achieve improved clinical outcomes.

While Drs. Campillo and Forteza-Rey believe that “decision-making based on preliminary results may be harmful for patients,” we believe that decision-making ignoring preliminary results of well-done observational studies may be equally harmful.

JoAnne Micale Foody, MD, FACC, FAHA
Department of Internal Medicine
Section of Cardiovascular Medicine
Yale University School of Medicine
New Haven, Conn
doi:10.1016/j.amjmed.2006.07.040

References

LETTERS

Reduce the Risk of Reinfarction and Related Complications after Myocardial Infarction

To the Editor:

The high prevalence of reinfarction as the cause of death following myocardial infarction (MI)\(^1\) has not been well addressed by current guidelines.\(^2\) Because 71% of fatal reinfarction cases in post-MI patients were undiagnosed before autopsy in a recent clinical trial,\(^1\) more sensitive detection strategy for reinfarction is needed. Monitoring of serial creatine kinase MB (CKMB) in post-MI patients has been recommended in textbook and less than 50% daily decline has been considered a threshold for detecting reinfarction or extended infarction.\(^3\) Troponin I measurement is likely comparable or better than creatine kinase MB (CKMB) for detecting reinfarction.\(^4\) Post-MI troponin I levels usually decrease over 50% in 24 hours or over 30% in 12 hours (correlated with the decrease of CKMB) after peaking in patients without severe renal insufficiency. Slower decrease is likely associated with reinfarction and often becomes faster after adjusting treatment such as increasing beta blockade and renin-angiotensin-aldosterone blockade.

Reinfarction is associated with heart failure. Troponin I and brain natriuretic peptide (BNP) are independent predictors of cardiac dysfunction and mortality.\(^5\) Serial monitoring of BNP in consecutive patients admitted for acute MI showed that post-MI BNP level of 200 pg/mL (or 100 pg/mL + baseline) appeared to be a threshold, below which there was minimal risk of complications during 4-week follow-up, and BNP fluctuations were mostly limited to the first week.\(^6\) This threshold of BNP correlates well with the conclusion of a large cohort study on preoperative evaluation that BNP level of 189 pg/mL was the best cut-off point for predicting short-term cardiac events.\(^7\) In asymptomatic or minimally symptomatic patients, BNP has been shown to be a better predictor for cardiac events and mortality than left ventricular ejection fraction (LVEF), New York Heart Association functional class, Goldman index or any other common test or clinical index.\(^7,8\) Monitoring of serial BNP has been recommended for acute coronary syndrome.\(^9\) BNP levels above 80 pg/mL were associated with high risk of worsening heart failure and death in 2-year follow-up. Lowering BNP to below 80 pg/mL was associated with significant risk reduction.\(^9\) BNP increment is likely a sensitive and specific test for cardiac ischemia.\(^10\) This suggests that stress tests should be avoided in patients with rising BNP, since it is associated with high risk of ischemia. Serial BNP is also likely to be the best test in risk stratifying heart failure patients.\(^8\) In 2 randomized controlled trials,\(^11,12\) BNP-guided therapy reduced cardiovascular events by half or more (including death, hospital admission, and heart failure decompensation).

Based on the above evidence, the following low risk approach may be considered as a supplement to current guidelines for reducing post-MI complications while awaiting further research:

1. Consider serial monitoring troponin I and BNP (eg, every 12 hours or daily) and adjusting treatment accordingly to maximize troponin I decrease and minimize BNP level until the following short-term (within 1 week post-MI) goals are reached: Elevated troponin I should decrease over 30% every 12 hours or over 50% every 24 hours in patients without severe renal insufficiency. BNP level should be decreasing or stable and below 100 pg/mL + baseline or below 200 pg/mL with unknown baseline. If this goal is not achievable, stabilize BNP to the lowest level achievable while balancing the risks of hypotension, electrolyte abnormalities and other complications.

2. Any rising BNP, or slower than expected decrease of troponin I, is associated with high risk of ischemia and reinfarction. Consider close monitoring and avoiding stress test in these patients.

3. The longer-term goals of management may include stabilizing BNP below 80 pg/mL or as low as tolerated while balancing the risks of other complications.

Charles Shang, MD
Department of Medicine, Cambridge Health Alliance,
Harvard Medical School
Everett, Mass

References


LETTER

The Reply:

We are gratified that Drs. Shang and Gu agreed that the high prevalence of reinfarction is not well recognized. However, we are concerned that the authors use our article to support a clinical practice for which there is no evidence. There is a general consensus that multiple biomarkers, including brain natriuretic peptide (BNP), high-sensitivity troponin T, and high-sensitivity C-reactive protein, have prognostic value.

However, proposing that troponins and BNP should be monitored in a serial fashion until certain prespecified levels are reached is a hypothesis that is clearly unrelated to the message of our article. The clinical recommendations, including the avoidance of stress testing in selected patients and the use of a target BNP, in long-term management are speculative.

The 2 important messages that can be derived from our autopsy study are that recurrent infarction represents the most likely cause of death in patients after acute myocardial infarction and that signs and symptoms of recurrent infarction in this patient population are difficult to interpret. For these reasons, the clinician should have a low threshold for suspecting and initiating diagnostic measures designed to detect recurrent ischemia as the cause of disease progression in this patient population.

Increased troponin levels and BNP levels are important components of the diagnostic armamentarium in these patients and may assist in the detection of ischemia and reduced left ventricle function if the clinical picture requires clarification. Although Drs Shang and Gu suggest an attractive algorithm for the use of these biochemical markers as prognostic indicators, we consider it premature to make such recommendations based on current documentation.

Stein Ørn, MD
Kenneth Dickstein, MD, PhD
University of Bergen
Stavanger University Hospital
Norway

LETTER

Progression of Chronic Kidney Disease: Can It Be Prevented or Arrested?

To the Editor:

The review article by Jaber and Madias was informative and interesting. The concept of renoprotective effect of angiotensin-converting enzyme (ACE) inhibitors and type 2 angiotensin receptor blockers (ARBs), as mentioned by the authors, is widely believed and has been adapted in guidelines for the management of hypertension.

However, this concept has been challenged in a recent large meta-analysis. Comparison of ACE inhibitors or ARBs with other antihypertensive drugs yielded only a small relative risk reduction of 0.71 (95% confidence interval, 0.49-1.04) for doubling of creatinine and end-stage renal disease (relative risk 0.87, 95% confidence interval, 0.75-0.99).

When blood pressure differences were reduced substantially by antihypertensive treatment in control groups, there was no evidence of a significant salutary effect of ACE inhibitors or ARBs on renal outcomes in patients with diabetes. Small benefits of the ACE inhibitors or ARBs seen in nondiabetic renal disease are explained by small study bias because small negative studies are more likely not to be published, and those small studies that have been published are likely to be of lower quality than large trials and more prone to bias.

In placebo-controlled trials, the renoprotective effects of ACE inhibitors or ARBs that are seen are probably largely secondary to blood pressure reduction rather than an effect on kidneys per se. This is supported by a post hoc analysis of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial, the largest hypertensive trial ever conducted in which patients were randomized to receive chlorthalidone, amlodipine, or lisinopril for a mean of 4.9 years. Neither amlodipine nor lisinopril was superior to chlorthalidone in reducing the rate of development of end-stage renal disease or 50% reduction in glomerular filtration rate.

The current prevalence of end-stage renal disease in the United States is greater than 400,000 and is projected to increase to 660,000 by the year 2010; therefore, more studies are needed in the field of renoprotection.

Ashok K. Malani, MD, FRCS
Hussam Ammar, MD
Heartlands Regional Medical Center
St Joseph, Mo

Mohamed H. Yassin
Maryland General Hospital
Baltimore, Md

doi:10.1016/j.amjmed.2006.01.028

References

The Reply:

We thank Malani and colleagues for their comments challenging the renoprotective efficacy of inhibitors of the renin-angiotensin system (RAS) on the strength of 2 recently published studies.1,2 As emphasized in our article, aggressive blood pressure control is an established renoprotective modality. Yet, many patients with chronic kidney disease (CKD) do not attain specified targets of blood pressure control. Hypertension of CKD is commonly resistant and requires a multitude of antihypertensive medication. More so, a redesign of systems of care might be needed before most patients with CKD attain blood pressure targets. Our article also emphasized the strong evidence in support of the renoprotective efficacy of inhibitors of the RAS in diabetic and nondiabetic, proteinuric CKD, independent of their blood pressure-lowering properties. In our judgment, the strength of the available evidence justifies the wide adoption by several organizations of RAS inhibition as a renoprotective modality independent of blood pressure control.

In regard to the recent article on the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT),1 we would argue that its post hoc analysis, the lack of baseline and follow-up proteinuria data, the undefined nature of the CKD of participants, and the relatively limited size of the diabetic subgroup with moderate-to-severe CKD (defined as glomerular filtration rate < 60 mL/min/1.73 m²) preclude firm conclusions. As for the large meta-analysis,2 it was largely influenced by the inclusion of the mammoth ALLHAT with all its shortcomings. Whereas large studies can provide more precise estimates of effects and associations, large sample size alone does not guarantee external validity. Indeed, ALLHAT’s original design centered on cardiovascular outcomes, and we consider objectional the inclusion of its secondary renal events in the meta-analysis.2

Although meta-analyses are now widely used to provide evidence in support of clinical strategies, it is instructive to reflect on a seminal article by LeLorier et al.3 examining the extent of discrepancies between meta-analyses and subsequent large randomized controlled trials. The article concluded that outcomes of large trials were not predicted accurately 35% of the time by the previously published meta-analyses, the overall positive predictive value of the meta-analyses being 68%. Mega-trials can be fraught with problems, and there is no substitute for clear and hard reasoning in designing a meta-analysis or a mega-trial.4 Therefore, special emphasis should be placed on exploring sources of heterogeneity among the trials included in a meta-analysis rather than in formulating definitive conclusions.

In conclusion, we recognize the urgent need of continued investigation at the experimental and clinical levels for identifying novel approaches aimed at delaying/arresting progression or even attaining regression of CKD.

Bertrand L. Jaber, MD
Nicolaos E. Madias, MD
Department of Medicine
Tufts University School of Medicine
Division of Nephrology
Caritas St Elizabeth’s Medical Center
Boston, Mass
doi:10.1016/j.amjmed.2006.06.039

References