1. Medicare prescription drug coverage: A very long wait for a very modest benefit • ARTICLE
   Pages 325-329
   James E. Dalen and Douglas J. Hartz

2. The prevalence and impact of alcohol problems in major depression: A systematic review • REVIEW ARTICLE
   Pages 330-341
   Lynn E. Sullivan, David A. Fiellin and Patrick G. O’Connor

3. Office management of heart failure in the elderly • REVIEW ARTICLE
   Pages 342-348
   Michael W. Rich

4. A clinical and therapeutic approach to thyrotoxicosis with thyroid-stimulating hormone suppression only • REVIEW ARTICLE
   Pages 349-361
   Giovanni Papi, Elizabeth N. Pearce, Lewis E. Braverman, Corrado Betterle and Elio Roti

5. What triggered this sudden eruption? • DISCUSSION
   Pages 362-363
   Lisa Scatena and Parwathi “Uma” Paniker

6. Persistent lower abdominal and groin pain: What is the diagnosis? • DISCUSSION
   Pages 364-367
   R. Scott Stephens, Cynthia Brown and Charles M. Wiener

7. A treatment option for some failing hearts • DISCUSSION
   Pages 368-370
   Julia H. Indik
| 8. | The impact of low health literacy on the medical costs of Medicare managed care enrollees • ARTICLE  
   Pages 371-377  
   David H. Howard, Julie Gazmararian and Ruth M. Parker |
| 9. | Community-acquired pneumonia as the initial manifestation of serious underlying diseases • ARTICLE  
   Pages 378-383  
   Miquel Falguera, Mariela Martín, Agustín Ruiz-González, Ricard Pifarré and Mercè García |
| 10. | Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia • ARTICLE  
    Pages 384-392  
| 11. | Prolonging the return visit interval in primary care • ARTICLE  
    Pages 393-399  
    Gordon Schectman, Gary Barnas, Prakash Laud, Laura Cantwell, Monica Horton and Edwin J. Zarling |
| 12. | Adjustment for do-not-resuscitate orders reverses the apparent in-hospital mortality advantage for minorities • ARTICLE  
    Pages 400-408  
    Naomi Bardach, Shoujun Zhao, Steven Pantilat and S. Claiborne Johnston |
| 13. | Iatrogenic events resulting in intensive care admission: Frequency, cause, and disclosure to patients and institutions • ARTICLE  
    Pages 409-413  
    Lisa Soleymani Lehmann, Ann Louise Puopolo, Shimon Shaykevich and Troyen A. Brennan |
| 14. | *Helicobacter pylori* eradication in the management of patients with idiopathic thrombocytopenic purpura • ARTICLE  
    Pages 414-419  
    Roberto Stasi, Zaccaria Rossi, Elisa Stipa, Sergio Amadori, Adrian C. Newland and Drew Provan |
15. Infectious agents and human immune diseases: Lessons from *Helicobacter pylori* • EDITORIAL
   *Pages 420-421*
   Giovanni Emilia, Mario Luppi and Giuseppe Torelli

16. The effect of self monitoring of blood glucose concentrations on glycated hemoglobin levels in diabetic patients not taking insulin: A blinded, randomized trial • ARTICLE
   *Pages 422-425*
   Mayer B. Davidson, Maria Castellanos, Don Kain and Petra Duran

17. Intravenous cyclophosphamide in refractory polyneuropathy associated with IgM monoclonal gammopathy: An uncontrolled open trial • SHORT COMMUNICATION
   *Pages 426-430*
   Mohamed A. Hamidou, Cristina Belizna, Sandrine Wiertlewsky, Marie Audrain, Charlotte Biron, Jean-Yves Grolleau and Jean-Marie Mussini

18. Dispensing of proton pump inhibitor medication is independently associated with subsequent asthma emergency hospital utilization • ARTICLE
   *Pages 431-434*
   Michael Schatz, Peter T. Chen, Eric Macy and Robert S. Zeiger

19. 52 precepts that medical trainees and physicians should consider regularly • DISCUSSION
   *Pages 435-438*
   Scott M. Wright, David B. Hellmann and Roy C. Ziegelstein

20. Medical professionalism and the generation gap • DISCUSSION
   *Pages 439-442*
   Lawrence G. Smith

21. Severe hypertriglyceridemia in a patient with lupus • CORRESPONDENCE
   *Pages 443-444*
   Dirk J. Blom and A. David Marais

22. Absence of effect of folic acid flour fortification on anticonvulsant drug levels • CORRESPONDENCE
   *Pages 444-445*
   Joel G. Ray, Loralie J. Langman, Muhammad M. Mamdani and David E.C. Cole
23. **Dysautonomia in Gulf War syndrome and in fibromyalgia** • CORRESPONDENCE
   *Page 446*
   Manuel Martinez-Lavin and Antonio G. Hermosillo

24. **Letter regarding: A 70-year-old African-American man with a history of fever, chills, and malaise** • CORRESPONDENCE
   *Pages 446-447*
   Victor J. Castro

25. **Literature and medical practice** • DISCUSSION
   *Page 448*
   Helle Mathiasen
EDITORIAL

Medicare prescription drug coverage: A very long wait for a very modest benefit

James E. Dalen, MD, MPH, Douglas J. Hartz, BS, BPA

Arizona Health Sciences Center, University of Arizona.

Nearly every western nation except the United States has health insurance for all its citizens, and most have some form of coverage for prescription drugs for its senior citizens. The struggle for universal health insurance, including some coverage for drugs for all Americans, has been, and will continue to be, a long and hazardous journey.1

In 2006, we will reach an important milestone in that journey: because of the passage of the Medicare Prescription Drug, Improvement and Modernization Act in 2003 (MMA), the 41 million Americans who are eligible for Medicare will have health insurance that will include a modest prescription drug benefit.2 The MMA, which was strongly supported by drug manufacturers, is a voluntary program that will be administered by private companies. This act has 2 features that have evoked considerable criticism. The first is that there is a very formidable gap in coverage. Recipients pay for the first $250, then they have a co-pay of 25% for the next $2250 of drug expenditures per year, and then there is no coverage until the senior has spent $5100. This gap, which has been termed “the doughnut hole”,2 was inserted in an attempt to keep the costs of drug coverage less than $400 billion over a 10-year period.2

The second feature of this act to receive widespread criticism is that it forbids Medicare from negotiating for discounts from drug manufacturers2 even though Federal legislation passed in 1990 requires drug manufacturers to give major discounts to government agencies, including the Veterans Administration, Medicaid, and the Department of Defense.3 The average discount to federal agencies is about 40% less than the average wholesale price,4 the price that the drug manufacturers suggest that wholesalers charge retail pharmacies.

Prescription drugs for chronic conditions

In 1997, 90% of seniors had at least one chronic condition requiring prescription drugs.5 Thirty percent take more than 3 medications and 15% have 5 or more chronic conditions requiring prescription drugs.6

Steinberg et al7 studied drug use by a large group of Medicare beneficiaries in 1998. Eighty-two percent were taking one or more medications for a chronic condition. The most frequent chronic conditions requiring prescription medications in this population are shown in Table 1.

The prescription drugs most frequently prescribed for Medicare patients in 2000 as reported by the Kaiser Family Foundation8 are shown in Table 2.

Eight of these 10 drugs are used to treat 1 of the chronic conditions shown in Table 1. Six are used to treat cardiovascular disease: Furosemide, HCTZ, Lipitor, Norvasc, Plavix, and Toprol. Fosamax is used to treat osteoporosis and Celebrex is for arthritis. Note that only 2 of these 10 drugs are generic (Furosemide and HCTZ); the other 8 are brand-name drugs. In the general population, about one-half of prescriptions are for brand-name drugs,9 whereas in the Medicare population nearly two-thirds are brand-name agents.10

It is very likely that an increasing percentage of the Medicare population will receive brand-name medications for these chronic conditions in the future. The number of new drugs used to treat heart disease and hypertension will continue to accelerate in the upcoming years. New guidelines for the treatment of hypertension will lead to more seniors receiving drug therapy for hypertension.11 New guidelines for the use of statins to lower LDL will greatly
increase their use in the elderly.12 Tseng et al13 reported that 66% of seniors in a Medicare-Choice population had hypercholesterolemia, and thus are candidates for long-term statin therapy. The introduction of new (and more expensive) platelet active agents14 and new oral anticoagulants15 will replace or be added to the use of aspirin and generic anticoagulants in seniors with cardiovascular disease. New guidelines for the prevention and treatment of congestive heart failure will increase the use of prescription drugs such as ACE inhibitors in the elderly.16

Other new, effective brand name drugs will increase the cost of treating chronic conditions in the Medicare population. Increased awareness of osteoporosis and available effective therapy has caused a sharp increase in treatment with brand name medications.17 The increased use of expensive COX-2 selective nonsteroidal anti-inflammatory drugs (NSAIDS) rather than inexpensive nonselective NSAIDs will increase the cost of treating arthritis in the elderly.18

### Costs of prescription medications

The 4 decades since Medicare was enacted have seen a remarkable increase in the expenditures for prescription medications. Medicare beneficiaries, who make up 12% of the population, account for one-third of total U.S. drug expenditures.7 By 2003, drug expenditures for the Medicare population were estimated to be $955 billion by the Congressional Budget Office (CBO),24 compared with total expenditures by Medicare of $277 billion. It is estimated that total expenditures for prescription drugs for the Medicare population in 2006, the year that the Medicare Modernization Act (MMA) becomes effective, will be $102 billion.6

There are multiple factors that have caused the marked increase in expenditures for prescription drugs by the Medicare population over the past several decades. With the aging of our population, the number of seniors with chronic conditions requiring prescription drugs is increasing. Research has led to the availability of an increased number of effective drugs for the treatment of these conditions. The number of medications per senior is increasing; from 1997 to 2001, there was a 23% increase in the number of prescriptions per senior.25 These effective new drugs are brand-name agents, covered by patent and therefore much more expensive than generic drugs. Brand-name drugs are priced at 2 to 3 times the cost of the drugs that they replace.9

Another major cause of the increase in drug expenditures in the United States is that the prices of brand-name drugs are much higher in the United States than in the rest of the world. Most other industrialized countries limit drug prices by a variety of mechanisms such as the use of formularies, reference pricing, or price controls. These techniques are highly effective; prices of brand-name drugs in other industrialized nations are 34–59% lower than U.S. prices.19

An additional factor causing increased drug expenditures for seniors is the fact that prices for brand-name drugs are accelerating. Moeller et al20 reported that the cost of prescription drugs increased 26% greater than the rate of inflation between 1997 and 2001. Families USA21 reported that the average wholesale price of the 30 brand-name drugs most prescribed for seniors had increased by another 22% from 2001 to 2004.21 In the 3 months after passage of the Medicare Prescription Bill Act in December 2003, the cost of brand-name drugs increased by 3.4%, compared with an inflation rate of 1.2%.22 A spokesperson for the pharmaceutical industry said these price hikes reflect increased spending for research.21 However, some suggest that these increases may have been a preparation for the Medicare Drug Discount Program that began in June 2004.22

As the number and cost of prescription drugs has accelerated, the cost of drugs per individual has increased. In 2003, the average expenditure per Medicare beneficiary (paid by the senior or by a third party) was $2322/year.1 The Congressional Budget Office estimates that the average cost of drugs per beneficiary in 2006, the first year of MMA, will be $3155.23

Drug costs for those seniors with multiple chronic medical conditions are much greater than the averages for the total senior population. In 1995, 5% of seniors had drug costs greater than $4000;7 this increased to 16% in 2003.19 In 1998, 10% of seniors had drug expenses more than $6000.24

### Who pays for seniors’ prescription drugs?

In 2001, 64% of Medicare beneficiaries had some form of third party coverage for drugs25 as shown in Table 3. However, the coverage is far from complete.

---

**Table 1** Chronic conditions requiring prescription drugs in a medicare population (％)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular disease (including hypertension)</td>
<td>55</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>23</td>
</tr>
<tr>
<td>Arthritis</td>
<td>16</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>14</td>
</tr>
<tr>
<td>Acid-peptic disease</td>
<td>14</td>
</tr>
<tr>
<td>Depression</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 2** Most frequently prescribed drugs for seniors

<table>
<thead>
<tr>
<th>Drug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celebrex</td>
</tr>
<tr>
<td>Fosamax</td>
</tr>
<tr>
<td>Furosemide</td>
</tr>
<tr>
<td>Hydrochlorothiazide (HCTZ)</td>
</tr>
<tr>
<td>Lipitor</td>
</tr>
<tr>
<td>Norvasc</td>
</tr>
<tr>
<td>Plavix</td>
</tr>
<tr>
<td>Premarin</td>
</tr>
<tr>
<td>Toprol</td>
</tr>
<tr>
<td>Xalatan ophthalmic solution</td>
</tr>
</tbody>
</table>
Employer-related coverage for retirees and for those seniors still working tends to be the most generous coverage. However, as drug costs have accelerated, many companies have increased the employee’s co-pays or totally eliminated drug benefits.

Two-thirds of Medicare HMOs (Medicare + Choice) offer some drug benefits. However, nearly 90% of HMOs cap the amount of drug benefits: caps average $600 to $800 per year. As with employer-related drug benefits, co-pays continue to increase, especially for brand name drugs. More than half of HMOs restrict drug coverage to generic drugs. Because most drugs prescribed for seniors are brand-name agents, this causes a major increase in out-of-pocket spending by these seniors.

Medicaid provides drug coverage, but only 12% of seniors are enrolled in Medicaid even though nearly half of all seniors have incomes less than 200% of the federal poverty level, and 10% have incomes less than the federal poverty level. Medigap coverage for drugs is available, but it is expensive: $1000 to $3500 per year. Premiums continue to increase as drug prices increase.

The 36% of seniors without drug coverage are older and sicker than those with drug benefits. Those without insurance pay 15–30% more for each prescription than third-party payers pay and at least 40% more than they would pay in other countries.

The acceleration of the number and the prices of brand-name drugs prescribed for seniors and the progressive erosion of supplemental drug benefits has had the inevitable result of a marked increase in out-of-pocket spending for prescription drugs by seniors.

**Consequences of the increased out-of-pocket costs for prescription drugs**

The average senior’s out-of-pocket spending for prescription drugs increased from $644 in 2000 to $1147 in 2004. The seniors who are most vulnerable to the increasing costs of drugs are those with multiple chronic conditions who are poor and have no drug coverage. Thirty percent of seniors have 3 or more chronic conditions requiring prescription drugs. Seniors who are poor but not on Medicaid may spend 50% of their income on health care, prescription drugs, and co-pays.

Multiple studies have shown that when out-of-pocket expenses for drugs increase because of increasing co-pays or loss of drug benefits, patients respond by skipping or discontinuing medications. In a study of California seniors without drug insurance, 29% skipped or failed to fill prescriptions. Of particular concern is that patients who lacked drug benefits skipped or discontinued medications essential for the treatment of chronic medical conditions such as diabetes, congestive heart failure, and hypertension.

Tseng et al studied a group of Medicare HMO patients whose drug expenditures exceeded an annual cap in 2002. Twenty-four percent decreased their drug use by skipping doses or stopping a medication. As in the case of the California patients studied by the Kaiser Foundation, these patients skipped or discontinued drugs used to treat hypertension, hyperlipidemia, congestive heart failure, and coronary disease.

Rector and Venus studied a Medicare HMO with a disproportionate number of patients with chronic conditions who were prescribed a mean of 4.7 prescription drugs. Nearly one-third “stinted” medications because of cost; that is, they skipped or reduced doses or discontinued or failed to fill prescriptions. Stinting was related to income: 38% with an income less than $12 000 per year stinted, compared with 17% with an income greater than $48 000. When out-of-pocket drug expense was less than $50 per month, 16% of patients stinted; when it exceeded $300 per month, 47% stinted.

These studies make it very clear that when out-of-pocket costs of prescription drugs increase, compliance decreases, even for drugs essential for the treatment of chronic conditions. Failure to treat chronic medical disorders leads to poorer health and increased hospitalizations and increased physician visits, and thus results in increased medical costs.

A dramatic illustration of the effect of increasing out-of-pocket drug expenses on compliance with drug therapy and its consequences was reported by Tamblyn et al. In Quebec, before 1996, prescriptions were free for the elderly if they were poor and, if not poor, they paid $2 per prescription. In 1996, the law was changed: the co-pay became $25 per prescription. Drug use was examined before and after the new law became effective. After the new law, 9% of the elderly discontinued essential drugs. The incidence of adverse events (defined as first acute hospital admission or death) in those who discontinued essential drugs doubled: from 5.8 to 12.6 per 10 000 person-months. In addition, there was a significant increase in emergency room visits by those who discontinued essential drugs.

Christian-Herman et al studied the effects of a large Medicare HMO’s decision to restrict drug coverage to generic drugs in one of its groups. As expected, the costs to the HMO for drugs in the generic-only group decreased, and the out-of-pocket costs for drugs for its

---

**Table 3 Type of drug coverage for medicare beneficiaries 2001 (%)**

<table>
<thead>
<tr>
<th>Type of Coverage</th>
<th>2001 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employer sponsored</td>
<td>29</td>
</tr>
<tr>
<td>Medicare HMO</td>
<td>15</td>
</tr>
<tr>
<td>Medicaid</td>
<td>11</td>
</tr>
<tr>
<td>Medigap</td>
<td>7</td>
</tr>
<tr>
<td>Other Public</td>
<td>2</td>
</tr>
<tr>
<td>No coverage</td>
<td>36</td>
</tr>
</tbody>
</table>

---

Christian-Herman et al studied the effects of a large Medicare HMO’s decision to restrict drug coverage to generic drugs in one of its groups. As expected, the costs to the HMO for drugs in the generic-only group decreased, and the out-of-pocket costs for drugs for its
members increased. However, the number of hospital days per member in the generic-only group increased when compared with a similar group that continued to have coverage for brand name drugs.

Because Medicare pays for hospitalization and for physician visits and emergency room visits, decreased compliance with drug therapy for chronic conditions increases total Medicare costs. It is clear that compliance with outpatient drug therapy is highly dependent on the out-of-pocket costs for drugs. Revision of the Medicare Prescription Act to eliminate the major gap in coverage for prescription drugs would increase compliance for drug therapy of chronic conditions. Increased compliance with drug therapy would significantly decrease total Medicare costs.

The 2003 Medicare Prescription Act\(^5\) as it stands, with its huge gap in coverage for prescription drugs, will be of marginal benefit to seniors with multiple medical conditions. The increasing use and the increasing cost of effective brand-name drugs will continue to cause seniors to skip or discontinue essential drugs, resulting in unnecessary hospitalizations, emergency room visits, and physician visits. Medicare will bear the costs of these preventable healthcare expenditures.

Anderson et al\(^6\) have shown how Medicare can eliminate the huge gap in drug coverage (“the hole in the doughnut”) without additional expenditures. They show that by allowing Medicare to negotiate with drug manufacturers as Medicaid, the Veterans Administration, and the Department of Defense are allowed to do, a discount as modest as 20% off the average wholesale price would save enough money to eliminate the entire gap in coverage. Because Medicare determines how much they will pay physicians, hospitals, and nursing homes, it is not unreasonable that Medicare should determine how much they will pay for prescription drugs.

Broadening Medicare prescription drug coverage would improve compliance with outpatient drug therapy and would thereby decrease total Medicare costs. Most important, broadening Medicare prescription drug coverage would improve the health of our senior population.

References


The prevalence and impact of alcohol problems in major depression: A systematic review

Lynn E. Sullivan, MD, David A. Fiellin, MD, Patrick G. O’Connor, MD, MPH

Department of Internal Medicine, Yale University School of Medicine.

ABSTRACT: Major depression and alcohol problems are common in primary care, yet little is known about the prevalence of alcohol problems in patients with depression or alcohol’s effect on depression outcomes. We strove to answer the following questions: How common are alcohol problems in patients with depression? Does alcohol affect the course of depression, response to antidepressant therapy, risk of suicide/death, social functioning and health care utilization? In which alcohol categories and treatment settings have patients with depression and alcohol problems been evaluated? English language studies from MEDLINE, PsychINFO, and Cochrane Controlled Trial Registry were reviewed. Studies were selected using predefined criteria if they reported on the prevalence or effects of alcohol problems in depression. Thirty-five studies met criteria and revealed a median prevalence of current or lifetime alcohol problems in depression of 16% (range 5–67%) and 30% (range 10–60%), respectively. This compares with 7% for current and 16–24% for lifetime alcohol problems in the general population. There is evidence that antidepressants improve depression outcomes in persons with alcohol dependence. Alcohol problems are associated with worse outcomes with respect to depression course, suicide/death risk, social functioning, and health care utilization. The majority of the studies, 34 of 35 (97%), evaluated alcohol abuse and dependence, and 25 of 35 (71%) were conducted in psychiatric inpatients. We conclude that alcohol problems are more common in depression than in the general population, are associated with adverse clinical and health care utilization outcomes, and that antidepressants can be effective in the presence of alcohol dependence. In addition, the literature focuses almost exclusively on patients with alcohol abuse or dependence in psychiatric inpatient settings, and excludes patients with less severe alcohol problems and primary care outpatient settings.

© 2005 Elsevier Inc. All rights reserved.

KEYWORDS: Depression; Depressive disorder; Alcohol drinking; Alcoholism; Comorbidity; Prevalence; Treatment outcomes; Primary health care

Major depression affects approximately 9.9 million Americans\(^1\) or roughly 5% of the population and is the leading cause of disability in the United States.\(^2\) The prevalence of depression ranges from 5% to 9% in pri-

Dr. Sullivan is supported by the National Institute on Drug Abuse Physician Scientist Award (NIDA # K12 DA00167), and Dr. Fiellin is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar. Funding for this project was made possible through grants from the Robert Wood Johnson Foundation’s Program of Research Integrating Substance Use Issues into Mainstream Healthcare (PRISM), the National Institute on Drug Abuse and the National Institute on Alcohol Abuse and Alcoholism. Requests for reprints should be addressed to Lynn E. Sullivan, MD, Yale University School of Medicine, 333 Cedar St., P.O. Box 208025, New Haven, CT 06520-8025.

E-mail address: lynn.sullivan@yale.edu.

0002-9343/$ -see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjmed.2005.01.007
Patients with depression frequently have alcohol problems. One household survey reported a prevalence of 16% for any alcohol diagnosis in depressed patients. These conditions have been found to coexist in diverse age, sex, and racial groups. Alcohol problems in depressed patients present diagnostic and management challenges and may adversely affect the course of depression and its response to standard therapies. Therefore, this systematic review was undertaken to 1) determine the prevalence of alcohol problems in patients with depression, 2) evaluate alcohol’s effects on health-related outcomes, and 3) determine the alcohol categories and treatment settings addressed in these studies. In addition, we assessed the methodologic quality of the literature.

This review addresses the following clinical questions: How common are alcohol problems in patients with depression? Will a patient’s depression course and response to antidepressant therapy be affected by alcohol problems? Do alcohol problems affect the risk of suicide/death in patients with depression? How do alcohol problems impact the social functioning and health care utilization of depressed patients? In which alcohol categories and treatment settings have patients with depression and alcohol problems been evaluated?

### Methods

#### Search strategy

We identified English language studies by searching the following databases: MEDLINE (1980 to September 2002), PsychINFO (1984 to September 2002), and the Cochrane Central Register of Controlled Trials (through 4th quarter of year 2002). Specific medical subject headings and text words in MEDLINE are listed in Table 1.

Publications were limited to clinical trials, consensus development conferences, editorials, guidelines, journal articles, meta-analyses, multicenter studies, reviews, twin studies, and validation studies.

Using a similar search strategy, we reviewed the Cochrane Database of Systematic Reviews using alcohol and depression in the title or as keywords and accessed the Cochrane Depression, Anxiety, and Neurosis Group and Cochrane Drugs and Alcohol Group.

#### Inclusion criteria

Studies were included if they met the following 5 criteria: 1) adult sample (≥19 years old for MEDLINE, ≥18 years old for PsychINFO); 2) current major depression (not “depressive symptoms” or dysthymia) according to standard or operationally defined criteria; 3) current (past year) or lifetime alcohol problem or use; 4) no polysubstance abuse; and 5) data offered on the prevalence of alcohol problems or use in depression or their relationship to depression-related outcomes.

Two reviewers independently applied these criteria to all abstracts. One hundred abstracts were randomly selected to determine a simple kappa statistic of 0.66, indicating fair to good agreement between the reviewers. When inclusion criteria were unclear from the abstract, the article was reviewed. In the 43/1579 (3%) cases of discordant readings, final inclusion/exclusion was determined by consensus be-

### Table 1 Specific Medical Subject Headings (MeSH) terms, main terms, and text words in MEDLINE and PsychINFO used to identify articles on alcohol use in major depression

<table>
<thead>
<tr>
<th>Concept</th>
<th>Terms</th>
<th>Text words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEDLINE Alcohol</strong></td>
<td>Alcoholism, alcohol drinking, ethanol, alcohol-related disorders, drinking behaviors</td>
<td>Alcohol, alcohol abuse, alcohol dependence, alcohol use disorders, problem drinking, hazardous drinking, heavy drinking, alcoholic, alcohol intake, alcohol consumption, harmful drinking, at-risk drinking</td>
</tr>
<tr>
<td>Depression</td>
<td>Depressive disorder, anti-depressive agents, depression</td>
<td>Major depression, unipolar depression, depressed patients, primary depression</td>
</tr>
<tr>
<td>Prevalence Outcomes</td>
<td>Prevalence, comorbidity</td>
<td>Treatment response, medication adherence, psychosocial functioning</td>
</tr>
<tr>
<td><strong>PsychINFO Alcohol</strong></td>
<td>Alcoholism, drinking behaviors, alcohol drinking patterns, ethanol, alcohol abuse</td>
<td>Alcohol, alcohol abuse, alcohol dependence, alcohol use disorders, problem drinking, hazardous drinking, heavy drinking, alcoholic, alcohol intake, alcohol consumption, harmful drinking, at-risk drinking</td>
</tr>
<tr>
<td>Depression</td>
<td>Major depression, anti-depressant drugs</td>
<td>Depression (depress$)</td>
</tr>
<tr>
<td>Prevalence Outcomes</td>
<td>Comorbidity</td>
<td>Prevalen$, comorbid$</td>
</tr>
<tr>
<td></td>
<td>Treatment outcomes, patient compliance, quality of life</td>
<td>Morbidity, mortality, treatment response, medication adherence, psychosocial functioning</td>
</tr>
</tbody>
</table>
between the 2 reviewers upon review of the manuscript. Selected bibliographies were reviewed for additional manuscripts.

Data extraction

The following data were extracted by 1 of 2 reviewers: treatment setting, demographics, alcohol problem or use category (Table 2), diagnostic tools used for depression, prevalence of depression, prevalence of alcohol problems or use, prevalence of alcohol problems or use in depression, and descriptions of interventions and outcomes. We categorized these studies of patients with current depression according to either current or lifetime alcohol problems or use, or both. Because of concordance between the diagnosis of alcoholism according to Research Diagnostic Criteria and Feighner criteria with the definitions of alcohol abuse and dependence found in the Diagnostic and Statistical Manual,24 we combined studies addressing alcoholism,25-28 alcohol abuse, and dependence. Similarly, research has demonstrated agreement between International Statistical Classification of Disease and Related Health Problems and Diagnostic and Statistical Manual criteria for alcohol dependence,29 so we combined studies using these criteria. When reported, quantitative data including appropriate statistical test results for outcomes were extracted.

Quality of evidence

Randomized clinical trials

Using a standard quality checklist,30 2 reviewers independently evaluated the quality of the randomized clinical trials. This 5-point checklist addresses the quality of randomization, blinding, and the handling of patients who withdraw or drop out of the study. According to this scoring system, articles with a score of 0–2 are considered poor quality and articles with a score of 3–5 are considered

| Table 2 Categories and definitions for alcohol problems |
|-------------|-------------------------------------------------|---------------------------------|
| Category     | Definition                                      | Organization or classification system |
| At-risk drinking | Men > 14 drinks per week or >4 drinks per occasion | National Institute on Alcoholism and Alcohol Abuse (NIAAA) |
|             | Women > 7 drinks per week or >3 drinks per occasion |                                             |
| Hazardous drinking | At risk for adverse consequences from alcohol | World Health Organization (WHO) |
| Harmful drinking  | Alcohol is causing physical or psychological harm | WHO |
| Alcohol abuse   | One or more of the following events in a year: | American Psychiatric Association (APA) |
|                 | Recurrent use resulting in failure to fulfill major role obligations | |
|                 | Recurrent use in hazardous situations | |
|                 | Recurrent alcohol-related legal problems (eg, driving under the influence) | |
|                 | Continued use despite social or interpersonal problems caused or exacerbated by alcohol | |
| Alcohol dependence | Three or more of the following events in a year: | APA |
|                 | Tolerance | |
|                 | Increased amounts to achieve effect | |
|                 | Diminished effect from same amount | |
|                 | Withdrawal | |
|                 | Substantial time spent obtaining, using, or recovering from alcohol’s effects | |
|                 | Important activities given up or reduced because of alcohol | |
|                 | Drinking more or longer than intended | |
|                 | Persistent desire or unsuccessful efforts to cut down or control alcohol use | |
|                 | Use continued despite being aware of having a physical or psychological problem caused or exacerbated by alcohol | |
| Alcoholism       | According to specific classification system | Schedule of Affective Disorders and Schizophrenia (SADS) |
|                 | | Research Diagnostic Criteria (RDC) |
|                 | | International Statistical Classification of Disease and Related Health Problems (ICD) criteria |
|                 | | Feighner criteria |
Results

Literature search results

The search identified 2027 abstracts. After excluding duplicates (n = 448), 1579 abstracts remained (900 from MEDLINE and 679 from PsychINFO). Review of the Cochrane databases produced no additional articles. Of the 1579 studies, 151 (10%) were excluded because they did not focus on adult subjects, 1259 (80%) because the subjects did not meet criteria for current depression, 113 (7%) because they did not assess alcohol problems or use, 21 (1%) because of the presence of polysubstance use, and one (<1%) because it did not address prevalence or outcomes. Four additional manuscripts19,35–37 were identified through review of bibliographies, resulting in the identification of 38 articles.

Among these 38 articles, 2838,39 were excluded because more complete data were subsequently published20,40. The 1 article that addressed alcohol use41 was excluded, leaving 35 articles. The diversity of research designs, heterogeneous definitions of alcohol problems, and diversity of outcomes precluded meta-analysis.42 Of the 35 studies, 20 (57%) addressed the prevalence of alcohol problems in depression, 6 (17%) addressed their impact on depression, and 9 (26%) addressed both.

Description of studies

Twenty-eight (80%) studies defined depression using the Diagnostic and Statistical Manual.12,43–45 Others used the Research Diagnostic Criteria46 (n = 2, 6%), the Schedule of Affective Disorders and Schizophrenia47,48 and the Research Diagnostic Criteria46 (n = 1, 3%), the International Statistical Classification of Disease and Related Health Problems49 (n = 2, 6%), the Feighner criteria11 (n = 1, 3%), or clinician diagnosis (n = 1; 3%). Twenty-four of 35 (68%) studies defined alcohol problems using Diagnostic and Statistical Manual criteria, with the remainder using the Research Diagnostic Criteria (n = 2, 6%), the International Statistical Classification of Disease and Related Health Problems (n = 2, 6%), the Feighner criteria (n = 1, 3%), quantity/frequency measures (n = 1, 3%), or no operational definition (n = 5, 14%).

Recruitment strategies were described as hospital admissions/outpatient visits (n = 16, 46%), sampling (n = 7, 20%), study recruitment (n = 5, 14%), and volunteers (n = 5, 14%). Two studies (6%) did not specify recruitment procedures. Eighteen studies (51%) reported subject age with the mean age ranging from 32 to 63 years. Twenty-two studies (63%) reported sex, with the majority of the subjects being women.

Prevalence studies

How common are alcohol problems in patients with depression?

Of the 29 prevalence studies, 15 (52%) examined current alcohol problems in depression while 10 (34%) addressed lifetime alcohol problems, and 4 (4%) addressed both (Table 3). The median prevalence from the 19 studies of current alcohol problems was 16% (range 5–67%). The median prevalence of the 14 studies of lifetime alcohol problems was 30% (range 10–60%).

Three studies, including 2 national surveys, addressed the question of whether alcohol problems are more common in patients with depression than in the general population. In a study of general medicine outpatients, Sherbourne et al50 found higher rates of current and lifetime alcohol abuse or dependence in patients with depression, 6% and 19% respectively, than in those with hypertension, diabetes, or heart disease, 3% to 4% and 14% to 16%. Grant and Harford20 found higher rates of current and lifetime alcohol problems in the patients with depression, 21% and 40% respectively, compared with those without depression, 7% and 16%, respectively, and the general population, 7% and 18%, respectively. Similarly, Regier et al19 found a prevalence of lifetime alcohol problems to be 16% in the patients with depression compared with 13% in the general population.

We conclude that both current and lifetime alcohol problems are common in patients with depression and more common than in the general population.

Outcome studies

Of the 15 studies of depression outcomes, 10 (67%) addressed current alcohol problems and 5 (33%) addressed lifetime alcohol problems in persons with depression. Three (20%) were randomized clinical trials, while 12 (80%) used an observational design. The studies categorized alcohol problems as at-risk drinking (n = 1, 7%), and alcohol abuse and dependence (n = 14, 93%).
### Table 3  Studies reporting prevalence estimates

<table>
<thead>
<tr>
<th>Source/Year</th>
<th>Treatment Setting</th>
<th>Age (mean ± SD, range, years)</th>
<th>Sex (% Male)</th>
<th>Category of Alcohol Problem</th>
<th>Prevalence of Major Depression (n/N, %)</th>
<th>Prevalence of Alcohol Problems (n/N, %)</th>
<th>Prevalence of Alcohol Problems in Major Depression (n/N, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hirschfeld et al, 1989</td>
<td>Multicenter NIMH study: psychiatric inpatient/outpatient</td>
<td>Major Depression without alcohol: 42 ± 16; Major Depression with alcohol: 36 ± 13</td>
<td>42</td>
<td>Current abuse/dependence</td>
<td>368/1000 - 37%</td>
<td>79/368 - 21%</td>
<td>79/368 - 21%</td>
</tr>
<tr>
<td>Loo et al, 1990</td>
<td>Multicenter study: psychiatric inpatient/outpatient antidepressant trial</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>816/1231 - 66%</td>
<td>277/1231 - 23%</td>
<td>139/816 - 17%</td>
</tr>
<tr>
<td>Sanderson et al, 1990</td>
<td>Psychiatric outpatient clinic</td>
<td>36 ± 12</td>
<td>44</td>
<td>Current abuse/dependence</td>
<td>197/576 - 34%</td>
<td>Not stated</td>
<td>16/197 - 8%</td>
</tr>
<tr>
<td>Speer et al, 1992</td>
<td>Mental health institute residential program</td>
<td>63</td>
<td>36</td>
<td>Current abuse/dependence</td>
<td>55/128 - 43%</td>
<td>29/128 - 25%</td>
<td>3/55 - 5%</td>
</tr>
<tr>
<td>Sherbourne et al, 1993</td>
<td>Medical outcomes study: community outpatient practice, hospital-based clinic, HMO, general medical providers/mental health specialists</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>775/2195 - 35%</td>
<td>Not stated</td>
<td>6%</td>
</tr>
<tr>
<td>Mulder et al, 1994</td>
<td>Recruited patients, unspecified setting</td>
<td>Not stated</td>
<td>49</td>
<td>Current abuse/dependence</td>
<td>109/109 - 100%</td>
<td>35/109 - 32%</td>
<td>35/109 - 32%</td>
</tr>
<tr>
<td>Grant and Harford, 1995</td>
<td>Community household survey</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>1427/42 862 - 3%</td>
<td>7%</td>
<td>21%</td>
</tr>
<tr>
<td>Salloum et al, 1995</td>
<td>Inpatient/outpatient at academic psychiatric facility</td>
<td>&lt;40 51%, 40-60 29%, 60+ 20%</td>
<td>33</td>
<td>Current abuse/dependence</td>
<td>2660/8139 - 33%</td>
<td>993/8139 - 12%</td>
<td>239/2660 - 9%</td>
</tr>
<tr>
<td>Fischer and Goethe, 1998</td>
<td>Psychiatric inpatient hospital</td>
<td>Median: 39</td>
<td>38</td>
<td>Current abuse/dependence</td>
<td>293/293 - 100%</td>
<td>Not stated</td>
<td>91/293 - 31%</td>
</tr>
<tr>
<td>Lin et al, 1998</td>
<td>V.A. psychiatric inpatient hospital</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>14/49 - 29%</td>
<td>6%</td>
<td>7%</td>
</tr>
<tr>
<td>Winokur et al, 1998</td>
<td>NIMH Collaborative Depression Study</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>678/678 - 100%</td>
<td>15%</td>
<td>102/678 - 15%</td>
</tr>
<tr>
<td>Parikh et al, 1999</td>
<td>Community survey</td>
<td>35</td>
<td>33</td>
<td>Current abuse/dependence</td>
<td>333/8116 - 4%</td>
<td>Not stated</td>
<td>52/333 - 16%</td>
</tr>
<tr>
<td>Lyketsos et al, 1999</td>
<td>Primary care/pyschiatric outpatient practices</td>
<td>50-58</td>
<td>29</td>
<td>Current abuse</td>
<td>768/768 - 100%</td>
<td>Not stated</td>
<td>153/768 - 20%</td>
</tr>
<tr>
<td>Fortney et al, 1999</td>
<td>Primary care/pyschiatric inpatient/outpatient V.A. hospital</td>
<td>Major depression without alcohol: 52; Major depression with alcohol: 44</td>
<td>96</td>
<td>Current abuse/dependence</td>
<td>10 036/67 878 - 15%</td>
<td>50 988/67 878 - 75%</td>
<td>6034/10 036 - 60%</td>
</tr>
<tr>
<td>Kessing, 1999</td>
<td>Psychiatric inpatient</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current abuse/dependence</td>
<td>17 477/20 350 - 86%</td>
<td>Not stated</td>
<td>Median - 5% (range 2% for first hospitl admission for major depression to 17% for 10th hospital admission)</td>
</tr>
<tr>
<td>Bartels et al, 2002</td>
<td>Primary care outpatient hospital/V.A./community health center—Primary Care Research in Substance Abuse and Mental Health for the Elderly (PRISME) Study</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Current at-risk drinking</td>
<td>689/2240 - 31%</td>
<td>620/2240 - 28%</td>
<td>52/689 - 8%</td>
</tr>
<tr>
<td>Source/year</td>
<td>Treatment setting</td>
<td>Age (mean ± SD, range, years)</td>
<td>Sex (% male)</td>
<td>Category of alcohol problem</td>
<td>Prevalence of major depression (n/N, %)</td>
<td>Prevalence of alcohol problems (n/N, %)</td>
<td>Prevalence of alcohol problems in major depression, (n/N, %)</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td>-----------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Melartin et al, 2002</td>
<td>Psychiatric inpatient, primary care outpatient</td>
<td>40 ± 11</td>
<td>27</td>
<td>Current abuse/dependence</td>
<td>269/542 = 50%</td>
<td>Not stated</td>
<td>66/269 = 25% Inpatients: 39% Outpatients: 22%</td>
</tr>
<tr>
<td>Regier et al, 1990</td>
<td>NIMH Epidemiological Catchment Area (ECA) Survey of households, residents of long-term mental hospitals, nursing homes, penal institutions</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>6-month: 3% 1-month: 2%</td>
<td>14% (N = 20,291) 17%</td>
<td></td>
</tr>
<tr>
<td>Brady et al, 1991</td>
<td>Psychiatric inpatient hospital</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>23/100 = 23%</td>
<td>58/58 = 100%</td>
<td>68% 16/58 = 28% 50–60%</td>
</tr>
<tr>
<td>Cook et al, 1991</td>
<td>Primary care V.A. inpatient</td>
<td>Major depression without alcohol: 62 ± 9 (55–84); Major depression with alcohol: 58 ± 5 (55–67)</td>
<td>41</td>
<td>Lifetime abuse/dependence</td>
<td>47/263 = 18%</td>
<td>Not stated</td>
<td>22/47 = 47%</td>
</tr>
<tr>
<td>Mueser et al, 1992</td>
<td>Psychiatric inpatient</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>59/218 = 27%</td>
<td>40/218 = 18%</td>
<td>24/59 = 41%</td>
</tr>
<tr>
<td>Leibenluft et al, 1993</td>
<td>NIMH outpatient research subjects or recruited locally</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>588/955 = 62%</td>
<td>Not stated</td>
<td>176/588 = 30%</td>
</tr>
<tr>
<td>Mueller et al, 1994</td>
<td>NIMH Collaborative Depression Study—“psychiatric facilities”</td>
<td>Major depression without alcohol: 40 ± 15; Major depression with alcohol: 37 ± 13</td>
<td>40</td>
<td>Lifetime abuse/dependence</td>
<td>100/100 = 100%</td>
<td>Not stated</td>
<td>Not stated 17/100 = 17% 30%</td>
</tr>
<tr>
<td>Malone et al, 1995</td>
<td>Psychiatric inpatient hospital</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>396/396 = 100%</td>
<td>Not stated</td>
<td>Not stated 37/375 = 10%</td>
</tr>
<tr>
<td>Fava et al, 1996</td>
<td>Psychiatric outpatient treatment program</td>
<td>Men: 39 Women: 38</td>
<td>Not stated</td>
<td>Life time abuse/dependence</td>
<td>375/375 = 100%</td>
<td>Not stated</td>
<td>Not stated 37/375 = 10%</td>
</tr>
<tr>
<td>Abraham and Fava, 1999</td>
<td>Psychiatric hospital outpatient</td>
<td>Non-alcohol-dependent: 32 Alcohol dependent: 32</td>
<td>41 Alcohol dependent: 48</td>
<td>Lifetime dependence</td>
<td>180/180 = 100%</td>
<td>148/180 = 82%</td>
<td>Overall dependence: 30%</td>
</tr>
<tr>
<td>Rae et al, 2002</td>
<td>Outpatient psychiatric clinic</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Lifetime abuse/dependence</td>
<td>479/1300 = 37%</td>
<td>Not stated</td>
<td>29/479 = 6% 181/479 = 38%</td>
</tr>
<tr>
<td>Zimmerman et al, 2002</td>
<td>Psychiatric outpatient practice</td>
<td>39 ± 12</td>
<td>32</td>
<td>Current abuse/dependence</td>
<td>Lifetime abuse/dependence</td>
<td>121/126 = 96%</td>
<td>Use in nondependent patients: 121/126 = 96% Use in dependent patients: 47/54 = 87%</td>
</tr>
<tr>
<td>Source/year</td>
<td>Patients with major depression (n)</td>
<td>Treatment setting</td>
<td>Category of alcohol problem</td>
<td>Study design</td>
<td>Jadad score/quality index score</td>
<td>Outcome in presence of alcohol problems</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
<td>----------------------------</td>
<td>--------------</td>
<td>---------------------------------</td>
<td>----------------------------------------</td>
<td></td>
</tr>
</tbody>
</table>
| Cook et al,26 1991 | 58 | Primary care V.A. inpatient | Lifetime alcoholism | Cohort score = 15 | Decreased proportion receiving electroconvulsive therapy during hospital admission (13% vs. 43%, \( P = 0.027 \))
|             |     |                                |                |               | Increased proportion receiving tricyclic antidepressant medications alone during hospital admission (81% vs. 40%, \( P = 0.005 \))
|             |     |                                |                |               | Decreased full acute major depression treatment response (25% vs. 41%, \( P = \) reported as nonsignificant)
|             |     |                                |                |               | Increased proportion receiving tricyclic antidepressant medications at time of hospital discharge (88% vs. 57%, \( P = 0.027 \))
|             |     |                                |                |               | Decreased proportion describing maximum recovery 1 month post-hospital discharge (25% vs. 55%, \( P = 0.05 \))
|             |     |                                |                |               | Increased proportion describing a lack of overall improvement in major depression at follow-up (43% vs. 12%, \( P = 0.014 \))
|             |     |                                |                |               | Increased proportion with hospital re-admission (25% vs. 0%, \( P = 0.004 \))

Mueller et al,28 1994 | 588 | Psychiatric inpatient/outpatient | Lifetime alcoholism | Cohort score = 15 | Decreased likelihood of recovery from major depression over 10 years (6/520 = 1.2% per week vs. 15/520 = 2.3% per week)*
|             |     |                                |                |               | Level of alcohol use had no effect on mean time (in weeks) to recovery from major depression (Wilcoxon \( t^2 = 0.14) or relapse to major depression
|             |     |                                |                |               | Level of alcohol use Mean weeks to recovery
|             |     |                                |                |               | Mild-moderate 43
|             |     |                                |                |               | Severe 53
|             |     |                                |                |               | Very severe 38

Labbate et al,52 1997 | 96 | Primary care inpatient | Lifetime abuse/dependence | Cohort score = 16 | No difference in number of major depressive episodes in patients with alcohol problems:
|             |     |                                |                |               | Number of episodes Odds ratio (95% CI)
|             |     |                                |                |               | 1 1.0
|             |     |                                |                |               | 2 0.8 (0.4–1.8)
|             |     |                                |                |               | 3 1.3 (0.7–2.6)

Melartin et al,51 2002 | 269 | Psychiatric inpatient/general medicine outpatient | Current abuse/dependence | Cross-sectional score = 15 | Increased use of alcohol in response to depressive symptoms and increased likelihood that alcohol relieved depressive symptoms (specific level of increase not reported)
|             |     |                                |                |               | Improved score on Hamilton Depression Scale (change = −6.4, \( P < 0.01 \)); Improved score on Beck Depression Inventory (change = −10.6, \( P < 0.005 \))

Leibenuft et al,54 1993 | 59 | Psychiatric outpatient | Lifetime dependence | Cross-sectional score = 4 | Increased use of alcohol in response to depressive symptoms and increased likelihood that alcohol relieved depressive symptoms (specific level of increase not reported)
|             |     |                                |                |               | Improved score on Hamilton Depression Scale (change = −9.1 ± 10.5 vs. Placebo group: −1.9 ± 16.6, \( P = NS \))
|             |     |                                |                |               | Improved score on Beck Depression Inventory (fluoxetine group: change = −7.6 ± 11.2 vs. Placebo group: 0.56 ± 14.9, \( P = 0.18 \))
|             |     |                                |                |               | Improved score on Hamilton Depression Scale (fluoxetine group: change = −6.5 ± 12.8 vs. Placebo group: 0.9 ± 12.1, \( P = 0.17 \))

Cornelius et al,55 1993 | 12 | Psychiatric inpatient | Current dependence | Case series: 8-week open-label of fluoxetine score = 13 | Increased use of alcohol in response to depressive symptoms and increased likelihood that alcohol relieved depressive symptoms (specific level of increase not reported)
|             |     |                                |                |               | Improved score on Hamilton Depression Scale (fluoxetine group: change = −9.1 ± 10.5 vs. Placebo group: −1.9 ± 16.6, \( P = NS \))
|             |     |                                |                |               | Improved score on Beck Depression Inventory (fluoxetine group: change = −7.6 ± 11.2 vs. Placebo group: 0.56 ± 14.9, \( P = 0.18 \))

Cornelius et al,56 1995 | 21 | Psychiatric inpatient | Current dependence | 12-week double-blind, placebo-controlled randomized clinical trial of fluoxetine score = 3 | Improved score on Hamilton Depression Scale (fluoxetine group: change = −6 ± 9.6 vs. Placebo group: −2 ± 13.3, \( P < 0.05 \))
|             |     |                                |                |               | Improved score on Beck Depression Inventory (fluoxetine group: change = −6.5 ± 12.8 vs. Placebo group: 0.9 ± 12.1, \( P = 0.17 \))

Cornelius et al,57 1997 | 51 | Psychiatric inpatient | Current dependence | 12-week double-blind, placebo-controlled randomized clinical trial of fluoxetine score = 3 | Improved score on Hamilton Depression Scale (fluoxetine group: change = −6 ± 9.6 vs. Placebo group: −2 ± 13.3, \( P < 0.05 \))
|             |     |                                |                |               | Improved score on Beck Depression Inventory (fluoxetine group: change = −6.5 ± 12.8 vs. Placebo group: 0.9 ± 12.1, \( P = 0.17 \))
Will the course of the patient’s depression and response to antidepressant therapy be affected by alcohol problems?

Two of 6 studies examining the effects of alcohol problems on depression course reported an increased risk of relapse and decreased likelihood of recovery (Table 4). Corneliussen et al found that patients with lifetime alcohol problems and current depression had more hospital readmissions for depression (25% vs. 0%, \( P = 0.004 \)), a lower rate of recovery from depression 1 month after hospital discharge (25% vs. 55%, \( P = 0.05 \)), and were more likely to describe a lack of overall improvement in depression at follow-up (43% vs. 12%, \( P = 0.014 \)) than patients with depression alone. \(^{26}\) Similarly, Mueller et al found that lifetime alcohol problems decreased the likelihood of recovery from depression from 2.3% to 1.2% over 10 years. \(^{28}\) In contrast, 3 studies found no association between alcohol problems and relapse or recurrent episodes of depression. \(^{25,51,52}\) It is interesting to note that Leibnluft et al found that patients with depression and alcohol dependence reported increased alcohol use in response to depressive symptoms and increased likelihood of alcohol relieving these symptoms than those without alcohol dependence. \(^{53}\)

Overall, the studies examining the efficacy of antidepressants in depression and alcohol problems reported a decrease in depressive symptoms, despite the patients having concomitant alcohol problems. One open-label study of fluoxetine in patients with depression and alcohol dependence found statistically significant reductions on standardized depression scales such as the Hamilton Depression Scale (−6.4 \( P \leq 0.01 \)) and the Beck Depression Inventory (−10.6, \( P < 0.005 \)). \(^{35}\) The first of 2 subsequent placebo-controlled randomized clinical trials performed found a nonstatistically significant improvement in depression scores in the treatment group. \(^{55}\) The second larger study found a statistically significant intergroup difference between the fluoxetine group and the placebo group on the Hamilton Depression Scale (−6 ± 9.6 vs. −2 ± 13.3, \( P = 0.05 \)) but a nonstatistically significant improvement on the Beck Depression Inventory (−6.5 ± 12.8 vs. 0.9 ± 12.1, \( P = 0.17 \)). \(^{56}\) In a follow-up case series of 31 of these patients, the degree of depressive symptoms in the fluoxetine group changed minimally during the year following completion of the trial, providing evidence that the beneficial effects of fluoxetine persisted even in persons with alcohol dependence. \(^{57}\)

An unblinded study of patients with depression who were receiving antidepressant treatment compared the outcomes of those with and without a diagnosis of lifetime alcohol dependence and found similar rates of remission of depressive symptoms in both groups. \(^{56}\) However, when the level of alcohol use was included in the analysis, subjects with greater alcohol use (≥28 drinks/week) had significantly smaller changes in their depression scores (−8.3 vs. −14.3, \( P = 0.007 \)) than those with lower levels of alcohol use (≤28 drinks/week). \(^{56}\)
We conclude that the data on alcohol’s effect on depression course are equivocal but that there is evidence that antidepressants can be effective in patients with alcohol problems.

Do alcohol problems affect the risk of suicide/death in patients with depression?

Four observational studies examined the risk of suicide/death as a depression outcome (Table 5)\(^{26,27,58,59}\). While all 4 studies concluded that a current or lifetime alcohol problem in patients with depression was associated with an increased risk of severe suicidal symptoms or acts, only 2 studies had statistically significant results \((P < 0.05)\).\(^{27,58}\)

We conclude that alcohol problems may increase the risk of suicide/death in patients with depression.

<table>
<thead>
<tr>
<th>Source/year</th>
<th>Patients with major depression (n)</th>
<th>Treatment setting</th>
<th>Category of alcohol problem</th>
<th>Study design Jadad score/quality index score</th>
<th>Outcomes in presence of alcohol problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corneliussen et al, 1995</td>
<td>5732</td>
<td>Psychiatric outpatient</td>
<td>Current dependence</td>
<td>Cross-sectional score = 12</td>
<td>Increased suicidal symptoms (scored as 0 = symptom not present, 1 = mild, 2 = moderate, 3 = severe; depressed nonalcoholics had mean score = 1.6 ± 1.2 vs. depressed nonalcoholics had mean score = 1 ± 0.97, (P &lt; 0.001))</td>
</tr>
<tr>
<td>Kessing et al, 1999</td>
<td>20,868</td>
<td>Psychiatric inpatient</td>
<td>Current alcoholism</td>
<td>Cohort score = 15</td>
<td>Increased risk of suicidal act (hazard ratio [95% CI] = 1.4 (1.2–1.7), (P &lt; 0.001)) or suicide (hazard ratio [95% CI] = 1.8 (1.1–3.1), (P &lt; 0.05))</td>
</tr>
<tr>
<td>Bartels et al, 2002</td>
<td>689</td>
<td>Primary care outpatient</td>
<td>Current at-risk drinking</td>
<td>Cross-sectional score = 11</td>
<td>Increased suicidal ideation in presence of alcohol (15% vs. 12%) and decreased death ideation (27% vs. 34%)</td>
</tr>
<tr>
<td>Cook et al, 1991</td>
<td>58</td>
<td>Primary care inpatient</td>
<td>Lifetime alcoholism</td>
<td>Cohort score = 15</td>
<td>Increased all-cause mortality 25% vs. 12%*</td>
</tr>
</tbody>
</table>

*No indicator of significance reported.

What is the impact of alcohol problems on social functioning and health care utilization for patients with depression?

Hirschfeld et al. found that depressed patients with a current alcohol problem had impaired relationships with their spouses at index, 6 months, and 12 months (all \(P < 0.01\)) but not after 24 months (Table 6).\(^{25}\) Similarly, Cook et al examined the effect of alcohol problems on interpersonal relations in persons with depression and found a higher rate of divorce (38% vs. 4%, \(P = 0.003\)) and living alone (46% vs. 15%, \(P = 0.02\)).\(^{26}\) Two studies examined the effect of alcohol problems on health care utilization for depression.\(^{26,60}\) In a study of 10,036 depressed persons with and without alcohol dependence, those with current alcohol dependence had more hospital days (199 vs. 100, \(P < 0.002\)) and outpatient visits (200 vs. 141, \(P = 0.002\)) over 4 years.\(^{60}\) In contrast, Cook and colleagues observed fewer

<table>
<thead>
<tr>
<th>Source/year</th>
<th>Patients with major depression (N)</th>
<th>Treatment setting</th>
<th>Category of alcohol problem</th>
<th>Study design Jadad score/quality index score</th>
<th>Outcome in presence of alcohol problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hirschfeld et al, 1989</td>
<td>368</td>
<td>Psychiatric inpatient/outpatient</td>
<td>Current alcoholism</td>
<td>Cohort score = 14</td>
<td>Impairment in spousal relationship assessed using the Longitudinal Interval Follow-up Evaluation (LIFE) scale with significant differences at index, 6 months, and 12 months (all (P &lt; 0.01)), but not at 18 or 24 months. Impaired social functioning at index (NS), 6 months, 12 months, 18 months, and 24 months (all (P &lt; 0.05))</td>
</tr>
<tr>
<td>Cook et al, 1991</td>
<td>58</td>
<td>Primary care inpatient</td>
<td>Lifetime alcoholism</td>
<td>Cohort score = 15</td>
<td>Higher rate of divorce at follow-up (38% vs. 4%, (P = 0.003)) and living alone (46% vs. 15%, (P = 0.02)), decreased hospital days (49 ± 53 vs. 22 ± 12, (P = 0.004))</td>
</tr>
<tr>
<td>Fortney et al, 1999</td>
<td>10,036</td>
<td>Psychiatric and primary care inpatient/outpatient</td>
<td>Current dependence</td>
<td>Cohort score = 18</td>
<td>Increased hospital days (199 vs. 100, (P &lt; 0.002)), increased outpatient visits (200 vs. 141, (P = 0.002)) (used retransformed expected values of adjusted log means)</td>
</tr>
</tbody>
</table>
hospital days (22 ± 12 vs. 49 ± 53, P = 0.004) over a 4-year period in a study of 58 patients with lifetime alcohol problems and current depression.26 In addition to notable differences in sample size, the former study examined current alcohol problems while the latter examined lifetime alcohol problems.

We conclude that alcohol problems have deleterious effects on social functioning and current, but not lifetime, alcohol problems increase health care utilization.

Alcohol categories and treatment settings

In which alcohol categories and treatment settings have patients with depression and alcohol problems been evaluated?

The studies we identified examined a limited number of categories of alcohol problems including at-risk drinking (n = 1, 3%), alcohol abuse (n = 4, 11%), alcohol dependence (n = 9, 26%), combined alcohol abuse/dependence (n = 17, 49%), and alcoholism (n = 4, 11%). Thirty-four of 35 (97%) studies focused on the more severe forms of alcohol problems, namely alcoholism, alcohol abuse, and alcohol dependence. With regard to treatment setting, 25 of 35 studies (71%) were conducted in psychiatric settings; 16 (46%) examined inpatients while 11 (31%) examined outpatients. Three (9%) studies were based in primary care, 4 (11%) were in both psychiatric and primary care settings, and 3 (9%) did not specify.

We conclude that few studies have been conducted evaluating low levels of alcohol use in primary care settings.

Quality of evidence

Randomized clinical trials

Of the 3 randomized clinical trials, 2 (67%) were assigned a quality score of 3 of a possible 5 points, indicating excellent quality,55,56 while 1 (33%) received a score of 1, indicating poor quality.57,61 None of the studies had an appropriate description of randomization or of the number of reasons for patient withdrawal. Two (67%) provided an appropriate description of double-blinding.55,56 The 2 reviewers had 100% agreement in their ratings.

Observational studies.

The mean quality score (of a maximum of 31) for the 7 cohort studies was 16.1 (range 14–20), 10.5 (range 4–15) for the 4 cross-sectional studies, and 13.5 (range 12–15) for the 2 case series. The cohort studies also scored higher than the other types of studies on reporting, bias, and confounding. The evaluation of external validity was high for all but 1 of the studies.53 None of the 12 observational studies addressed power. The 2 reviewers had 100% agreement in their ratings.

Discussion

Our review reveals that alcohol problems are prevalent in patients with depression, with median rates of 16% for current alcohol problems and 30% for lifetime alcohol problems. The studies that compared these rates with those of nondepressed patients or of the general population found higher rates of current and lifetime alcohol problems in depressed patients. The identified literature supports the efficacy of antidepressant therapy in patients with depression and alcohol problems. It also provides evidence of the association between alcohol problems and adverse depression outcomes. Specifically, alcohol problems in depression are associated with a worse depression course, an increased risk for relapse to and decreased likelihood of recovery from depression, increased suicide/death risk, worsening social functioning, and increased health care utilization.

Of note, the literature we reviewed focuses almost exclusively on alcohol abuse and dependence. Studies of lower levels of alcohol use such as at-risk, hazardous, or harmful drinking for which primary care physicians are more likely to provide care are rare. Further, most studies were conducted in psychiatric settings, limiting their relevance to primary care. Given that the majority of patients with depression or alcohol problems seek treatment in a primary care setting,60-63 it is important to determine whether similar results will be found in this setting and across the spectrum of alcohol problems.

Prevalence estimates for alcohol abuse and dependence in the general population range between 7% for current and 16–24% for lifetime alcohol problems.22,62 Our review of alcohol abuse and dependence in patients with depression revealed higher prevalence estimates than those found in the general population. Prevalence estimates for less severe alcohol problems in primary care- and population-based studies range between 0.3% to 29%.14,18,63 Unfortunately it is not possible to compare these prevalence estimates with those found in patients with depression, as our review reveals that these surveys have not yet been reported.

There are limitations to this review. Relatively few studies examined alcohol problems in depression, and the methodological quality of these studies was variable. The clinical trials often did not provide complete information about study design or information on subject follow-up. Many studies used a range of recruitment strategies and had small sample sizes with short periods of observation, limiting their generalizability and the ability to assess the duration of effects. While 11 of the 12 observational studies achieved high scores on external validity, these studies performed less well in other areas. The studies used an array of diagnostic criteria and outcome measurements, making it challenging to arrive at comprehensive conclusions about prevalence and outcomes. The observational studies examining the effects of alcohol on depression outcomes presented equivocal findings. In contrast, antidepressant therapy was found to be effective in depressed patients with alcohol problems. A major
References


Office management of heart failure in the elderly

Michael W. Rich, MD

Cardiovascular Division, Washington University.

KEYWORDS: Heart failure; Elderly; Office management

ABSTRACT: Heart failure is one of the most common conditions affecting older patients seen by clinicians in routine office practice. This article reviews the clinical features, diagnosis, and management of heart failure in elderly patients evaluated in the ambulatory care setting, and provides concise, practical information relevant to all major aspects of care. Specific topics include the role of diagnostic testing, such as echocardiography and B-type natriuretic peptide; principles of nonpharmacological management, including patient education, diet, exercise, and daily weights; drug therapy for systolic heart failure as well as heart failure with preserved left ventricular systolic function; end-of-life issues; and when to refer the patient to a specialist. Although heart failure in the elderly differs in many important respects from heart failure occurring in middle-aged patients, the general approach to diagnosis and management is similar in younger and older patients.

© 2005 Elsevier Inc. All rights reserved.

Heart failure is the leading cause of hospitalization and the most costly medical illness among senior citizens. In addition, the clinical features and management of heart failure in older adults differ in many important respects from heart failure occurring in middle-aged individuals (Table 1).

Clinical features

Among all adults regardless of age, the most common symptoms of heart failure are exertional dyspnea, orthopnea, effort intolerance, and edema. However, elderly patients are more likely to have comorbid conditions, such as pulmonary, renal, hepatic, or thyroid disease; anemia; depression; venous insufficiency; obesity; or chronic deconditioning that render the “typical” symptoms of heart failure relatively nonspecific. In addition, atypical symptoms, including impaired cognition, confusion, irritability, nausea, diarrhea, and anorexia, become increasingly common presentations of heart failure in the elderly, and in some cases these symptoms may be the sole manifestation of worsening heart failure.

Pulmonary rales, elevated jugular venous pressure, dependent edema, and an S3 or S4 gallop rhythm are the most common physical findings in elderly patients with heart failure. However, as with symptoms, the sensitivity and specificity of these findings decrease with age. In particular, pulmonary rales may be due to chronic lung disease or atelectasis, and peripheral edema may be due to venous insufficiency or medications, especially calcium channel blockers.
Diagnosis

Accurate diagnosis of heart failure in elderly patients remains challenging, in part due to the nonspecific and atypical nature of presenting symptoms and signs and in part due to the high prevalence of comorbid conditions that may mimic the symptoms and signs of heart failure. Figure 1 provides an algorithm for evaluating elderly patients with suspected heart failure. If the clinical index of suspicion for heart failure is high (ie, documented prior history of heart failure, major risk factors, typical symptoms and signs), an empiric course of diuretic therapy is appropriate. Clinical improvement accompanied by documented weight loss provides additional evidence in support of the diagnosis.

In patients with an intermediate probability for heart failure—based upon history and physical examination—additional evaluation with a standard chest radiograph and a plasma B-type natriuretic peptide (BNP) level is warranted. The presence of cardiomegaly, pulmonary venous engorgement, interstitial edema, and pleural effusions establishes a diagnosis of heart failure; a normal chest radiograph makes heart failure unlikely. Alternatively, chest radiographs may support chronic lung disease, pulmonary vascular disease, or pneumonia as explanations for the patient’s symptoms. Recently, the plasma BNP level has been shown to be a useful adjunct in the diagnostic assessment of patients with unexplained shortness of breath. Because BNP levels increase with age, especially in women, BNP levels are somewhat less reliable in older patients than in younger patients with suspected heart failure. Nonetheless, a BNP level <100 pg/mL makes active heart failure unlikely. A BNP level >400 pg/mL strongly suggests the presence of heart failure, although acute pulmonary embolization and volume overload in the setting of renal failure can also produce BNP elevations in this range. BNP levels in the range of 100–400 pg/mL are nondiagnostic.

Patients with low probability for heart failure based on initial assessments (ie, no prior history of heart failure, no major risk factors, equivocal or atypical symptoms and

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Features distinguishing heart failure in the elderly from heart failure occurring during middle age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle age</td>
</tr>
<tr>
<td>Prevalence</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Sex</td>
<td>Men &gt; Women</td>
</tr>
<tr>
<td>Etiology</td>
<td>CAD</td>
</tr>
<tr>
<td>Clinical features</td>
<td>Typical</td>
</tr>
<tr>
<td>LVEF</td>
<td>Reduced</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>Few</td>
</tr>
<tr>
<td>RCTs</td>
<td>Many</td>
</tr>
<tr>
<td>Therapy</td>
<td>Evidence-based</td>
</tr>
<tr>
<td>Physician</td>
<td>Cardiologist</td>
</tr>
</tbody>
</table>

CAD = coronary artery disease; LVEF = left ventricular ejection fraction; RCTs = randomized clinical trials.

Table 2 | Principal effects of aging on cardiovascular structure and function |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased vascular stiffness</td>
</tr>
<tr>
<td>Increased myocardial stiffness</td>
</tr>
<tr>
<td>Decreased β-adrenergic responsiveness</td>
</tr>
<tr>
<td>Impaired mitochondrial ATP production</td>
</tr>
<tr>
<td>Decreased baroreceptor responsiveness</td>
</tr>
<tr>
<td>Impaired sinus node function</td>
</tr>
<tr>
<td>Impaired endothelial function</td>
</tr>
<tr>
<td>Net effect: marked reduction in cardiovascular reserve</td>
</tr>
</tbody>
</table>

ATP = adenosine triphosphate.

Figure 1 Office evaluation of elderly patients with suspected heart failure. CAD = coronary artery disease; HF = heart failure; LVEF = left ventricular ejection fraction; PLVSF = preserved left ventricular systolic function.
signs) should undergo appropriate evaluation to establish an alternative diagnosis. However, if the diagnosis remains in doubt after initial assessment, a chest radiograph and BNP level should be obtained to screen for heart failure.

All patients with definite or probable heart failure should undergo an echocardiogram to assess left ventricular (LV) systolic and diastolic function, as well as to evaluate for wall motion abnormalities, chamber size and wall thickness, and the presence of significant valvular or pericardial disorders. In addition, exercise or pharmacologic stress testing with echocardiographic or radionuclide imaging should be considered in patients with moderate or high likelihood for having CAD, especially those who are suitable candidates for revascularization in the event that stress testing reveals severe ischemia.

Management

Optimal management of heart failure in patients of all ages involves control of risk factors, patient education and self-management, and the judicious use of medications. The goals of therapy are to relieve symptoms, improve quality of life, reduce the need for hospitalization, and prolong functional survival.

General measures

Hypertension, diabetes, and dyslipidemia should be treated aggressively in accordance with established guidelines. Use of all tobacco products should be discontinued, and alcohol should be used in moderation (no more than 1–2 drinks/day) or not at all. Patients with severe CAD and significant symptomatic or silent ischemia confirmed by noninvasive testing should be considered for percutaneous or surgical coronary revascularization. Patients with other treatable conditions that may exacerbate heart failure severity should be managed appropriately (eg, anemia, thyroid disease, sleep disorders, depression). The use of nonsteroidal anti-inflammatory drugs (NSAIDs), which promote sodium and water retention and antagonize the effects of most heart failure medications, should be avoided or minimized.

Ideally, all patients with heart failure should receive individualized instruction about the causes, symptoms, and management of this chronic condition. Several patient-oriented educational brochures and online resources are available (see Appendix). In particular, patients (preferably in conjunction with family members or significant others) should be instructed to limit sodium intake to no more than 2 000 mg/day (no added salt and avoidance of high sodium foods; Table 3) and total fluid intake to no more than about 64 ounces/day. (Note: Patients should be warned that the oft-cited recommendation to drink 8–10 glasses of water every day does not apply to persons with heart failure!)

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Selected foods with high sodium content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canned vegetables and soups (unless labeled low sodium)</td>
<td></td>
</tr>
<tr>
<td>Tomato juice and tomato sauces</td>
<td></td>
</tr>
<tr>
<td>Prepared meats and lunch meats</td>
<td></td>
</tr>
<tr>
<td>Bacon, sausage, hot dogs</td>
<td></td>
</tr>
<tr>
<td>Anchovies, sardines, regular canned tuna and salmon</td>
<td></td>
</tr>
<tr>
<td>Frozen entrées and pot pies</td>
<td></td>
</tr>
<tr>
<td>Most canned meats and other canned entrées</td>
<td></td>
</tr>
<tr>
<td>Most “fast” foods</td>
<td></td>
</tr>
<tr>
<td>Macaroni and cheese (boxed)</td>
<td></td>
</tr>
<tr>
<td>Stuffing mix</td>
<td></td>
</tr>
<tr>
<td>Pickles, sauerkraut, olives</td>
<td></td>
</tr>
<tr>
<td>Chips, pretzels (salted), other snack foods</td>
<td></td>
</tr>
<tr>
<td>Microwave popcorn</td>
<td></td>
</tr>
<tr>
<td>Most sauces (eg, ketchup, barbecue, cocktail, soy, teriyaki, Worcestershire)</td>
<td></td>
</tr>
<tr>
<td>Many seasonings (eg, meat tenderizer, monosodium glutamate (MSG), taco seasoning)</td>
<td></td>
</tr>
</tbody>
</table>
All heart failure patients should obtain a bathroom scale (cost: <$15) and monitor and record their daily weights. Weights should be measured at the same time each day, generally in the morning after voiding but before eating. A “dry weight” should be established (based on the home scale, not the office scale), and patients should be counseled to maintain their weight within narrow limits (eg, ±2 lbs. in patients with mild to moderate heart failure, ±1 lb. in patients with severe heart failure). With proper instruction, many elderly patients can self-adjust their diuretic dosage to maintain their weights within the prescribed range.

Regular exercise helps preserve functional capacity in heart failure patients, and older heart failure patients should be advised to engage in low-intensity aerobic exercise (eg, walking, stationary cycling) at least 3–5 times per week (preferably daily), using symptoms as a guide to gauge exercise intensity and duration. Strengthening exercises using light weights or low-resistance bands are also desirable if available. Patients should be cautioned to avoid strenuous exercise and to stop exercising if they feel faint or light-headed, or if they experience chest discomfort, excessive sweating, or undue shortness of breath.

Finally, patients should be provided with explicit instructions about how and when to contact their physician, for example, weight gain of 5 or more pounds above dry weight, increased shortness of breath or edema. In many cases, aggressive management of these “warning signs” can stave off a visit to the Emergency Room or admission to the hospital.

### Drug therapy of systolic heart failure

**Figure 2** summarizes current recommendations for the treatment of systolic heart failure, defined as an LV ejection fraction of less than 40%. Angiotensin-converting enzyme (ACE) inhibitors are indicated in all patients with LV systolic dysfunction, whether symptomatic or asymptomatic, and **Table 4** lists starting and target dosages for ACE inhibitors currently approved for the treatment of heart failure. In general, elderly heart failure patients should be started on the lowest available dose, and the dose should be gradually increased at 2- to 4-week intervals as tolerated to achieve the target dosage proven to be effective based on the results of prospective randomized clinical trials. Blood pressure, renal function, and serum potassium levels should be monitored closely during dose titration and periodically thereafter. Angiotensin receptor blockers (ARBs) are suitable alternatives in patients unable to tolerate ACE inhibitors due to cough, but ACE inhibitors remain the preferred first line agents. In addition, it is important to recognize that the incidence rates of renal dysfunction and hyperkalemia are similar with ACE inhibitors and ARBs.

The combination of hydralazine and isosorbide dinitrate has recently been shown to reduce mortality in African-Americans with moderate to severe systolic heart failure who were already receiving ACE inhibitors and beta-blockers. Therefore, addition of these agents should be considered in African-Americans who remain symptomatic de-

### Table 4 Angiotensin-converting enzyme inhibitors for systolic heart failure

<table>
<thead>
<tr>
<th>Agent</th>
<th>Starting dose</th>
<th>Target dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Captopril</td>
<td>6.25 mg TID</td>
<td>50 mg TID</td>
</tr>
<tr>
<td>Enalapril</td>
<td>2.5 mg BID</td>
<td>10–20 mg BID</td>
</tr>
<tr>
<td>Lisinopril</td>
<td>2.5–5 mg qd</td>
<td>20–40 mg qd</td>
</tr>
<tr>
<td>Ramipril</td>
<td>1.25–2.5 mg qd</td>
<td>10 mg qd</td>
</tr>
<tr>
<td>Quinapril</td>
<td>10 mg BID</td>
<td>40 mg BID</td>
</tr>
<tr>
<td>Fosinopril</td>
<td>5–10 mg qd</td>
<td>40 mg qd</td>
</tr>
</tbody>
</table>

*Agents approved by the Food and Drug Administration (FDA) for the treatment of heart failure in the United States

### Table 5 Beta-blockers for systolic heart failure

<table>
<thead>
<tr>
<th>Agent</th>
<th>Starting dose</th>
<th>Target dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisoprolol</td>
<td>1.25 mg qd</td>
<td>10 mg qd</td>
</tr>
<tr>
<td>Carvedilol</td>
<td>3.125 mg BID</td>
<td>25–50 mg BID</td>
</tr>
<tr>
<td>Metoprolol tartrate</td>
<td>6.25 mg BID</td>
<td>75 mg BID</td>
</tr>
<tr>
<td>Metoprolol succinate CR/XL</td>
<td>12.5–25 mg qd</td>
<td>200 mg qd</td>
</tr>
</tbody>
</table>

**Figure 3** Management of heart failure with preserved left ventricular systolic function. ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; LVEF = left ventricular ejection fraction.
Diuretics are the most effective agents for maintaining optimal fluid balance, and most heart failure patients will require maintenance diuretic therapy, usually with a loop diuretic (furosemide, bumetanide, torsemide). The diuretic dosage should be titrated to maintain the patient at the predefined “dry weight”. As noted above, many elderly patients can effectively manage their diuretic dosage with proper instruction.

Table 6  Indications for referral to a cardiologist or geriatrician

<table>
<thead>
<tr>
<th>Indications for referral to a cardiologist or geriatrician</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant Ischemic symptoms or severe ischemia on stress testing</td>
</tr>
<tr>
<td>Significant valvular heart disease by echo (esp. severe aortic stenosis)</td>
</tr>
<tr>
<td>Persistent moderate to severe symptoms despite standard therapy</td>
</tr>
<tr>
<td>Severe left ventricular systolic dysfunction (ejection fraction &lt;30%)</td>
</tr>
<tr>
<td>Documented or suspected life-threatening arrhythmias</td>
</tr>
<tr>
<td>Frailty or progressive decrease in physical function</td>
</tr>
<tr>
<td>Recurrent falls not likely to be of cardiac origin</td>
</tr>
<tr>
<td>Age over 80 with multiple comorbid conditions</td>
</tr>
<tr>
<td>Inability of caregiver to provide effective care</td>
</tr>
</tbody>
</table>

Despite conventional therapy. These drugs are also an acceptable alternative in patients unable to take ACE inhibitors or ARBs (eg, due to advanced renal insufficiency or hyperkalemia). The starting dose of hydralazine is 25 mg TID and the dose should be titrated to 75 mg TID as tolerated. The starting dose of isosorbide dinitrate is 10 mg TID and the dose should be titrated to 40 mg TID over a period of several weeks.

In the past, beta-blockers were widely viewed as being contraindicated in patients with heart failure. However, numerous prospective randomized trials have now conclusively demonstrated that long-term beta-blocker therapy improves LV function and reduces hospitalizations and mortality in patients with mild to severe symptomatic heart failure. As a result, beta-blockers should now be considered standard therapy in patients with systolic heart failure. Table 5 summarizes initial and target doses of beta-blockers approved for treatment of heart failure in the United States. As with ACE inhibitors, therapy should be initiated at a very low dose, and the dosage should be titrated at 2- to 4-week intervals to achieve the target dose based on the results of published clinical trials. Recent data suggest that carvedilol may be associated with greater benefits than metoprolol, and many heart failure specialists now consider carvedilol to be the preferred beta-blocker.

When initiating beta-blocker therapy, patients should be advised that they may experience a transient worsening of heart failure symptoms. These symptoms usually resolve with continued therapy but may recur with subsequent dosage increases. Nonetheless, with careful monitoring, most heart failure patients, including the elderly, are able to tolerate at least moderate doses of beta-blockers.

Digoxin reduces symptoms and hospitalizations in heart failure patients of all ages but has no effect on mortality. Therefore, use of digoxin should be reserved for use in patients who remain significantly symptomatic despite ACE inhibitors, beta-blockers, and diuretics. Recent data indicate that the optimal therapeutic range for digoxin is a serum level of 0.5–0.8 ng/mL. For most elderly patients with preserved renal function, a daily digoxin dose of 0.125 mg is sufficient; for patients with significant renal impairment (est. creatinine clearance <50 cc/min), a lower dose may be appropriate.

Spironolactone is an aldosterone antagonist that reduces mortality and heart failure hospitalizations in patients with advanced heart failure (New York Heart Association class III–IV symptoms and LV ejection fraction <30%). Similarly, eplerenone has been shown to improve outcomes in patients with heart failure and an ejection fraction <40% following acute myocardial infarction. The dose of spironolactone is 12.5–25 mg once daily, and the dose of eplerenone is 25–50 mg once daily. Both drugs may cause significant hyperkalemia, especially in elderly patients with renal insufficiency, and these agents are contraindicated in patients with a serum creatinine level ≥2.5 mg/dL or estimated creatinine clearance <30 cc/min. In addition, spironolactone may cause painful gynecomastia in up to 10% of patients.

Management of heart failure with preserved LV systolic function

At the time of this writing, no drugs have been approved in the United States for treatment of heart failure with preserved LV systolic function. In addition, although the ARB candesartan reduces heart failure hospitalizations in this population, it has no effect on mortality. Pending the results of ongoing trials, treatment of heart failure with preserved LV systolic function should focus on effective blood pressure control in accordance with established treatment guidelines and maintenance of euvolemia through dietary sodium restriction and diuretic therapy (Figure 3). It is important to avoid overdiuresis, however, as this may result in worsening renal function and prerenal azotemia. In patients with CAD, ischemia should be treated aggressively, and coronary revascularization should be considered if appropriate. Patients with heart failure and preserved LV systolic function are at increased risk for atrial fibrillation; when symptomatic, anti-arrhythmic therapy to restore and maintain sinus rhythm should be considered.

End-of-life issues

Elderly heart failure patients have a median survival of less than 3 years, and patients with advanced heart failure have a 1-year mortality rate of 25–50%. The prognosis is thus worse than for most forms of cancer, and heart failure patients should be encouraged to develop an advance directive and appoint a
durable power of attorney. Hospice care should be offered to patients with terminal heart failure who are not suitable candidates for or do not wish to consider more aggressive therapeutic options.

Indications for referral

Table 6 lists major indications for referral to a cardiologist or geriatrician. Additional reasons for cardiology consultation might include recent onset of atrial fibrillation or flutter, uncertainty about the need for diagnostic cardiac catheterization or other invasive interventions, and assistance with managing valvular heart disease or other complex cardiac conditions. The primary indications for referral to a geriatrician are for evaluation of cognitive dysfunction, frailty, or other “geriatric syndromes” (eg, falls, incontinence, weight loss, depression).

Prevention

Primary prevention of heart failure can be achieved through aggressive management of established coronary risk factors, including hypertension, dyslipidemia, diabetes, tobacco use, obesity, and physical inactivity. Table 7 demonstrates the efficacy of anti-hypertensive therapy in reducing the risk of heart failure in elderly patients. Similarly, treatment of other risk factors reduces the risk of CAD, thereby reducing the incidence of heart failure.

Appendix

Resources

For health professionals

- www.acc.org
- www.americanheart.org
- www.hfsa.org
- www.nhlbi.nih.gov

Full text, executive summary, pocket addition, and PDA download versions of the ACC/AHA heart failure guidelines; to access, click on Clinical Statements/Guidelines, then Search Statements by Topic A–Z, then “H” for heart failure, then scroll down.

To order brochures, click on Publications and Resources, then Get the Facts, then Conditions, Procedures, and Care, then Living with Congestive Heart Failure.

Official website of the Heart Failure Society of America.

Access to guidelines for management of hypertension, hyperlipidemia, and obesity.

For patients

- www.abouthf.org
- www.americanheart.org
- www.heartfailure.org
- www.nhlbi.nih.gov

Teaching materials sponsored by the Heart Failure Society of America.

American Heart Association website; click on Disease and Conditions, then Congestive Heart Failure.

Teaching materials in English and Spanish.

For informational materials, click on A-Z Diseases and Conditions Index, then on Heart Failure.

References


Table 7  Effect of antihypertensive therapy on incident heart failure in older adults

<table>
<thead>
<tr>
<th>Trial</th>
<th>Year</th>
<th>n</th>
<th>Age, yrs</th>
<th>Risk reduction for heart failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWPHE</td>
<td>1985</td>
<td>840</td>
<td>&gt;60</td>
<td>22%</td>
</tr>
<tr>
<td>Coope</td>
<td>1986</td>
<td>884</td>
<td>60–79</td>
<td>32%</td>
</tr>
<tr>
<td>STOP-HTN</td>
<td>1991</td>
<td>1627</td>
<td>70–84</td>
<td>51%</td>
</tr>
<tr>
<td>SHEP</td>
<td>1991</td>
<td>4736</td>
<td>≥60</td>
<td>55%</td>
</tr>
<tr>
<td>STONE</td>
<td>1996</td>
<td>1632</td>
<td>60–79</td>
<td>68%</td>
</tr>
<tr>
<td>Syst-Eur</td>
<td>1997</td>
<td>4695</td>
<td>≥60</td>
<td>36%</td>
</tr>
<tr>
<td>Syst-China</td>
<td>1998</td>
<td>2394</td>
<td>≥60</td>
<td>38%</td>
</tr>
</tbody>
</table>

EWPHE = European Working Party on Hypertension in the Elderly; SHEP = Systolic Hypertension in the Elderly Program; STONE = Shanghai Trial of Nifedipine in the Elderly; STOP-HTN = Swedish Trial in Old Patients with Hypertension; Syst-China = Systolic Hypertension in China Trial; Syst-Eur = Systolic Hypertension in Europe Trial.


REVIEW

A clinical and therapeutic approach to thyrotoxicosis with thyroid-stimulating hormone suppression only

Giovanni Papi, MD, Elizabeth N. Pearce, MD, MSc, Lewis E. Braverman, MD, Corrado Betterle, MD, Elio Roti, MD

Department of Internal Medicine (GP), Endocrinology Unit, ASL Modena; Section of Endocrinology, Diabetes, and Nutrition (ENP, LEB), Boston Medical Center, Boston University School of Medicine; Department of Medical and Surgical Sciences (CB), Chair of Endocrinology, University of Padova; and Institute of Endocrinology (ER), University of Milan.

ABSTRACT: Subclinical hyperthyroidism is defined as normal serum free thyroxine (T4) and triiodothyronine (T3) concentrations and persistently suppressed thyroid stimulating hormone (TSH) concentrations. The most common cause of subclinical hyperthyroidism is the use of suppressive doses of L-thyroxine for treatment of hypothyroidism or, less commonly, diffuse nontoxic goiter or thyroid carcinoma (exogenous subclinical hyperthyroidism). Endogenous subclinical hyperthyroidism may be caused by a variety of thyroid disorders that result in overproduction and release of thyroid hormones from the gland with normal/high 24-hour thyroid radiiodine uptake or by inflammation in the thyroid resulting in release of excess thyroid hormones and low 24-hour thyroid radiiodine uptake. Several groups have investigated whether persistent endogenous or exogenous subclinical hyperthyroidism, like overt hyperthyroidism, causes symptoms, adverse effects on the cardiovascular and skeletal systems, and increased mortality, whether endogenous subclinical hyperthyroidism evolves to overt thyrotoxicosis, and whether or not it should be treated. The present report reviews the most important and recent studies of subclinical hyperthyroidism and attempts to draw conclusions based upon the literature and the authors' experience.

The entity known as “subclinical hyperthyroidism” is characterized by a persistently suppressed serum thyroid stimulating hormone (TSH) associated with normal circulating free thyroid hormone concentrations. Increased serum free thyroxine (T4) or free triiodothyronine (T3) and suppressed TSH concentrations define overt thyrotoxicosis.

Several studies have investigated whether persistent endogenous (thyroid source) or exogenous (administered thyroid hormone) subclinical hyperthyroidism, like the overt form, causes symptoms, has effects on the cardiovascular and skeletal systems, or increases mortality. Others investigated whether endogenous subclinical thyrotoxicosis evolves to overt disease, and whether or not it should be treated. Pertinent articles from 1999 to 2004, identified by Medline searches, were included in this review. Important older studies were included, particularly where more recent data were unavailable.
Causes of subclinical hyperthyroidism and differential diagnosis

Many conditions may cause subclinical hyperthyroidism; all forms of overt hyperthyroidism may be preceded by a period of subclinical disease (Table 1). A useful classification distinguishes between endogenous subclinical hyperthyroidism with a high/normal 24-hour thyroid radioiodine uptake, suggesting a hyperfunctioning thyroid gland, and subclinical hyperthyroidism with a low/absent radioiodine uptake due to inflammatory changes in the thyroid (thyroiditis). The most common causes of endogenous subclinical hyperthyroidism are Graves’ disease and toxic nodular goiter. Exogenous subclinical hyperthyroidism is due to the administration of TSH suppressive doses of L-thyroxine and results in a low or absent thyroid radioiodine uptake. In cases where it is difficult to distinguish between endogenous and exogenous causes (ie, surreptitious or inadvertent thyroid hormone ingestion) or between subclinical or overt hyperthyroidism and nonthyroidal illness, a serum thyroglobulin may be useful. It is logical that the serum thyroglobulin concentration will be elevated in the setting of endogenous thyroid hormone overproduction, decreased or undetectable in the setting of exogenous thyroid hormone ingestion, and normal in the setting of nonthyroidal illness.

True subclinical hyperthyroidism must be distinguished from nonthyroidal illnesses. In patients with acute nonthyroidal illness, the serum T3 concentration is always decreased. The serum T4 and free T4 values also may be low in those patients who are most severely ill, and the serum TSH is often decreased. In this “sick euthyroid syndrome,” the peripheral conversion of T4 to T3 is decreased due to impaired outer ring 5’-deiodinase activity, and the secretion of TSH and thyrotropin releasing hormone are often reduced. In order to confirm the diagnosis of subclinical hyperthyroidism, and to exclude the presence of nonthyroidal illness or laboratory errors, repeat measurement of the serum TSH over the course of many weeks is recommended (Figure). In addition, it is important to obtain a detailed medication history in patients with low serum TSH values because the chronic use of corticosteroids, nandrolone, somatostatin, octreotide, lanreotide, dopamine, and bro-mocriptine decreases TSH secretion but does not affect serum T4 and T3 concentrations.

Subacute painful thyroiditis, painless lymphocytic thyroiditis (usually postpartum), and drug-induced thyroiditis may be triphasic disorders—thyrotoxicosis, hypothyroidism, and euthyroidism. The initial phase of thyroiditis may present as transient subclinical hyperthyroidism.

In normal pregnancy during the first trimester, when human chorionic gonadotropin (a weak TSH-agonist) levels are highest, serum free T4 levels are high-normal or slightly elevated, resulting in a decreased serum TSH concentration. Mild gestational thyrotoxicosis does not affect fertility, premature delivery, miscarriage, or neurologic development of the newborn and young child and, therefore, rarely requires treatment. Pregnancy should be excluded in evaluating young women with low serum TSH values.

The laboratory definition of subclinical hyperthyroidism

The introduction of sensitive TSH assays in clinical practice has overcome some difficulties in unequivocally defining subclinical hyperthyroidism. The U.S. Preventive Services Task Forces has defined subclinical thyrotoxicosis as a serum TSH concentration below the lower limit of the reference range with normal free T4 and free T3 values. However, subclinical hyperthyroidism has been defined variably in different settings, with cutoff serum TSH values ranging from 0.1 to 0.5. This variability makes it difficult to directly compare different studies.

<table>
<thead>
<tr>
<th>Table 1 Causes of subclinical hyperthyroidism defined by the 24-hour thyroid radioiodine uptake</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Common causes</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Rare causes</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Radiation thyroiditis</strong></td>
</tr>
<tr>
<td><strong>Factitious thyrotoxicosis</strong></td>
</tr>
<tr>
<td><strong>Struma ovarii</strong></td>
</tr>
<tr>
<td><strong>Metastatic differentiated thyroid cancer</strong></td>
</tr>
</tbody>
</table>

*Radioactive iodine imaging is contraindicated in pregnant women. Gestational hyperthyroidism is caused by overproduction and secretion of thyroid hormone as is found in causes of high/normal radioactive iodine uptake thyrotoxicosis.
Epidemiology

The prevalence of subclinical hyperthyroidism in various geographic areas depends on dietary iodine intake. Low iodine intake induces a compensatory growth of the thyroid gland that, over time, may develop into a diffuse goiter, followed by the development of hyperplastic nodules and, finally, areas of autonomously functioning tissue, which may cause subclinical or overt hyperthyroidism.

Belfiore et al. investigated autonomously functioning thyroid nodules in two samples from northeastern Sicily. In patients from an iodine-deficient area, the prevalence of autonomously functioning thyroid nodules (4.4%) and toxic nodules (1.3%) was significantly higher than in subjects living in an iodine-sufficient area. The Pescopagano survey reported a similar prevalence of autonomously functioning thyroid nodules (4.7%) in another iodine-deficient Italian sample; thyroid functional autonomy occurred in both nodular and diffuse goiter, but its relative frequency was significantly greater in the former group (21.2% vs. 6.4%; P <0.0001). In this survey the prevalence of subclinical hyperthyroidism progressively increased from 0.7% in children to 15.4% in subjects >75 years old. The prevalence of subclinical hyperthyroidism was higher than that reported in iodine-sufficient areas.

Prevalence studies have also been carried out in samples living exclusively in iodine-sufficient areas. The Whickham survey reported a 3% prevalence of low but measurable serum TSH values (0.05–0.5 mU/L). There was a 3.9% prevalence of serum TSH values <0.1 mU/L among 2575 ambulatory >60-year-old subjects in the Framingham Heart Study Cohort. In the Colorado Thyroid Disease Prevalence Study, based on 25 862 participants, a 2.1% prevalence of serum TSH values 0.01 to 0.3 mU/L was reported, with a prevalence of 0.9% among individuals not taking thyroid medication. In the U.S. National Health and Nutrition Examination Survey (1988–1994), an overall 0.7% prevalence of subclinical hyperthyroidism (defined as serum TSH <0.1 mU/L and total T4 <169.9 nmol/L) was found in the United States adult population, with some racial differences: in the white, non-Hispanic population, subclinical hyperthyroidism prevalence was higher (0.8%) than in the black, non-Hispanic (0.6%) and Mexican-American (0.5%) populations. Finally, the SWAN Study measured serum TSH concentrations in 3242 42–52-year-old, pre- and early perimenopausal U.S. women; 3.2% had serum TSH values <0.5 mU/L; African-American women had significantly lower mean serum TSH concentrations than white, Hispanic, and Chinese women.

Consequences of persistent subclinical hyperthyroidism

Effects on the cardiovascular system

Several adverse effects of subclinical hyperthyroidism on the cardiovascular system have been reported; the most important are resting tachycardia, atrial arrhythmias, increased cardiac muscle mass, and diastolic dysfunction. Studies have been conducted in patients with both exogenous and endogenous subclinical hyperthyroidism (Table 2).

In a cohort of subjects not taking thyroid hormone, the prevalence of atrial fibrillation in patients with subclinical hyperthyroidism was 12.7%, compared with 2.3% in euthyroid subjects. The relative risk for atrial fibrillation in subjects with subclinical hyperthyroidism was 5.2. This study confirmed an earlier report. Most recently, a retrospective review of charts of women with atrial fibrillation demonstrated that the type of atrial fibrillation present (persistent vs. paroxysmal) did not differ significantly based on thyroid status.

Some, but not all, case control studies have noted impaired diastolic function characterized by delayed relaxation in patients with exogenous subclinical hypothyroidism.
<table>
<thead>
<tr>
<th>Authors and references</th>
<th>Type of subclinical hyperthyroidism</th>
<th>Study design</th>
<th>Thyroid stimulating hormone (TSH) level of study group (mU/L)</th>
<th>Subjects (n)</th>
<th>Effects reported in subclinical hyperthyroid subjects</th>
</tr>
</thead>
</table>
| Tseng et al, 1989<sup>19</sup> | Exogenous and endogenous | Case control | Mean 0.26 ± 0.12; normal free thyroxine (T4) index and total triiodothyronine (T3) ≤0.1 | 15 34 | -Shorter isovolumetric contraction time  
-Shorter pre-ejection period  
-Lower pre-ejection period-left ventricular ejection time ratios  
-Increased risk for atrial fibrillation  |
| Sawin et al, 1994<sup>10</sup> | Exogenous and endogenous | Case cohort | Mean ≤0.1 | 61 1576 | -Increased left ventricular end-systolic volume  
-Impaired left ventricular diastolic filling  
-Reduced exercise capacity  
-Effects reversed after 4 months beta blocker treatment  |
| Biondi et al, 1996<sup>16</sup> | Exogenous Case control; Uncontrolled clinical trial | Mean ≤0.05 | 10 age-, sex-, habitus-, and lifestyle matched | -Increased left ventricular end-systolic volume  
-Impaired left ventricular diastolic filling  
-Reduced exercise capacity  
-Effects reversed after 4 months beta blocker treatment  |
| Shapiro et al, 1997<sup>23</sup> | Exogenous | Case control | Mean ≤0.01 | 17 17 age- and sex-matched | -No interventricular septum or left ventricular posterior wall thickness  
-No difference in end-diastolic dimension or diastolic flow velocities  
-Increased left ventricular mass index  |
| Biondi et al, 2000<sup>24</sup> | Endogenous | Case control | Mean 0.15 ± 0.1 | 23 23 age-, sex-, and lifestyle matched | -Increased peak of aortic flow velocity  
-Prolonged diastolic relaxation time  
-Increased interventricular and left ventricular posterior wall septum thickness-Increased ventricular fractional shortening time  
-Increased mean velocity of circumferential fiber shortening  |
| Mercuro et al, 2000<sup>21</sup> | Exogenous Uncontrolled clinical trial | Mean ≤0.1; normal free T3 ≤0.1 | 19 0 | -Increased end-diastolic dimension  
-Increased left ventricular mass index  
-Reduced peak overload, peak oxygen uptake, and anaerobic threshold during exercise  
-Increased interventricular septum and left ventricular posterior wall thickness  |
<p>| Auer et al, 2001&lt;sup&gt;17&lt;/sup&gt; | Endogenous | Case control | Mean &lt;0.4; normal free T3 and free T4 | 613 22,300 | -Increased prevalence of atrial fibrillation  |</p>
<table>
<thead>
<tr>
<th>Authors and references</th>
<th>Type of subclinical hyperthyroidism</th>
<th>Study design</th>
<th>Thyroid stimulating hormone (TSH) level of study group (mU/L)</th>
<th>Subjects (n)</th>
<th>Effects reported in subclinical hyperthyroid subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petretta et al, 2001</td>
<td>Endogenous</td>
<td>Case control</td>
<td>&lt;0.3; normal free T3 and free T4</td>
<td>30</td>
<td>Decreased isovolumetric relaxation time - No difference in left ventricular mass - Decreased cardiac output after radioactive iodine treatment - Increased systemic vascular resistance after radioactive iodine treatment</td>
</tr>
<tr>
<td>Faber et al, 2001</td>
<td>Endogenous</td>
<td>Uncontrolled clinical trial</td>
<td>0.006–0.09; normal free T3 and free T4</td>
<td>6</td>
<td>Decreased cardiac output after radioactive iodine treatment - Increased systemic vascular resistance after radioactive iodine treatment</td>
</tr>
<tr>
<td>Cetinarslan et al, 2003</td>
<td>Endogenous</td>
<td>Case control</td>
<td>Mean 0.1 ± 0.1</td>
<td>36</td>
<td>Shorter minimal electrocardiographic p wave duration - Greater difference between maximal and minimal p wave duration</td>
</tr>
<tr>
<td>Donatelli et al, 2003</td>
<td>Unspecified</td>
<td>Retrospective chart review</td>
<td>≤0.1; normal free T3 and free T4</td>
<td>19</td>
<td>Decreased left ventricular end-diastolic volume - No difference in prevalence of persistent vs. paroxysmal atrial fibrillation</td>
</tr>
<tr>
<td>Sgarbi et al, 2003</td>
<td>Endogenous</td>
<td>Clinical trial</td>
<td>≤0.07; normal free T3 and free T4</td>
<td>10</td>
<td>After 6 months methimazole: - Decreased atrial and ventricular premature beats - Decreased left ventricular mass index - Decreased left ventricular septum and posterior wall thickness</td>
</tr>
<tr>
<td>Völzke et al, 2004</td>
<td>Endogenous</td>
<td>Case control</td>
<td>0.1–0.3 or &lt;0.1 with normal free T3 and free T4</td>
<td>300</td>
<td>Increased carotid intima media thickness</td>
</tr>
<tr>
<td>Gullu et al, 2004</td>
<td>Exogenous</td>
<td>Case control; Uncontrolled clinical trial</td>
<td>0.1–0.4</td>
<td>12</td>
<td>Increased left ventricular mass index - Increased isovolumetric relaxation time - Decreased left ventricular mid-systolic diameter - Decreased early and late diastolic flow velocities - Decreased exercise capacity - Improved diastolic function and exercise capacity after 3 months atenolol treatment</td>
</tr>
</tbody>
</table>
based on echocardiographic parameters. Some\textsuperscript{21,22} but not all\textsuperscript{19,20} of these have reported increased left ventricular mass, particularly increased posterior wall and interventricular septum thickness. Cardiac changes in subclinical hyperthyroidism appear to be associated with decreased exercise capacity. When Mercuro et al\textsuperscript{21} reduced subjects’ L-thyroxine dose to the minimal amount required to maintain the serum TSH concentration at 0.1 mU/L or lower, echocardiographic and ergometabolic parameters normalized. Two recent echocardiographic case control studies\textsuperscript{24,25} have similarly demonstrated decreased diastolic function in patients with endogenous subclinical hyperthyroidism. Left ventricular mass was significantly increased in only one of those two studies.\textsuperscript{24} All of these studies are limited by small sample sizes, and some are further limited by the use of control subjects unmatched for body mass index and usual physical activity.

Intima media thickness has been used as a proxy measure for atherosclerotic risk. Ultrasonography has been used to evaluate carotid arteries in subjects over age 45 years.\textsuperscript{26} Intima media thickness was significantly increased in subclinically hyperthyroid individuals compared with euthyroid subjects, and intima media thickness was inversely correlated with serum TSH concentration after adjustment for other known risk factors for increased intima media thickness such as age, blood pressure, and the presence of diabetes. The underlying mechanism explaining this finding is not clear.

The cardiovascular effects of treatment for subclinical hyperthyroidism have been assessed in a few small trials. Two uncontrolled trials have reported improvements in diastolic function and exercise capacity in patients with exogenous subclinical hyperthyroidism after 3–4 months of treatment with beta blockers.\textsuperscript{20,22} Another trial studied 10 patients with endogenous subclinical hyperthyroidism, before and after 6 months of methimazole treatment.\textsuperscript{27} After achieving euthyroidism, patients had significant decreases in heart rate, total number of beats during 24 hours, number of atrial and ventricular premature beats, and left ventricular mass index. Finally, in an uncontrolled trial, women with subclinical hyperthyroidism from toxic nodular goiter were compared before and after radioactive iodine treatment (post-treatment serum TSH ranged from 0.28 to 1.12 mU/L, suggesting that euthyroidism may not have been achieved in all subjects).\textsuperscript{28} Following treatment, there were significant decreases in mean heart rate (11%) and cardiac output (19%), and a significant 30% increase in systemic vascular resistance.

Effects on the skeletal system

Overt hyperthyroidism is associated with increased bone resorption, low bone mineral density, and fractures in postmenopausal women.\textsuperscript{29–34} Whether subclinical hyperthyroidism of endogenous or exogenous origin causes low bone mineral density or increases fracture rate is still controversial.\textsuperscript{35–39} Table 3 lists some of the most important studies on the effects of exogenous and endogenous subclinical hyperthyroidism on the skeleton; we have focused on studies using fracture as an endpoint, meta-analyses, and studies published within the past 5 years.

Exogenous subclinical hyperthyroidism and the skeleton

The results of numerous small studies examining the effects of exogenous subclinical hyperthyroidism on bone density have been inconsistent. A 1994 meta-analysis pooled data from 13 cross-sectional studies in which bone density was measured in both premenopausal and postmenopausal women with suppressed serum TSH concentrations due to long-term L-thyroxine treatment, and in control subjects.\textsuperscript{40} TSH-suppressive therapy caused 0.17–0.46% bone loss per year in premenopausal women, a value not significantly different from controls; in contrast, bone loss averaged 0.77–1.39% per year in postmenopausal women, a value significantly higher than in controls. These results were substantially confirmed by a 1996 meta-analysis, which evaluated the effects of suppressive L-thyroxine therapy on bone mineral density from 41 published cross-sectional studies, including 1250 individuals.\textsuperscript{41} Subjects who had a past history of overt hyperthyroidism were excluded. In this analysis, exogenous subclinical hyperthyroidism was associated with significant bone loss at all skeletal sites in postmenopausal women, but not in premenopausal women.

Since these meta-analyses were published, two additional small case control studies have found no effect of exogenous subclinical hyperthyroidism on bone density in premenopausal women.\textsuperscript{42,43} Another case control study similarly found no bone density effect of exogenous subclinical hyperthyroidism in women.\textsuperscript{44} However, this study was limited by a control group that was not matched for age or menopausal status.

There are currently only limited data examining the effects of exogenous subclinical hyperthyroidism on bone density in men. One study measured bone mineral density in 17 men given long-term L-thyroxine TSH-suppressive therapy.\textsuperscript{45} Free T4 was increased in 41% of the sample, consistent with overt rather than subclinical hyperthyroidism. There was no significant difference in bone mineral density between the L-thyroxine-treated men and age- and weight-matched controls. Another study evaluated bone density in men treated with long-term suppressive L-thyroxine doses for thyroid cancer and 32 men with newly diagnosed, treated Graves’ disease.\textsuperscript{46} Z scores in both groups were lower than expected but there were no significant differences in bone density between the two groups. However, some patients in both groups had elevated free T4 levels, and treated Graves’ disease patients do not constitute an ideal control group.

Four large case-cohort studies to date have used fracture as an endpoint in evaluating the effects of exogenous subclinical hyperthyroidism. The first did not find excess fracture risk in subjects taking suppressive doses of L-thyroxine compared with subjects who were euthyroid on L-thyrox-
Table 3  Skeletal complications reported in patients with endogenous and exogenous subclinical hyperthyroidism

<table>
<thead>
<tr>
<th>Authors and references</th>
<th>Type of subclinical hyperthyroidism</th>
<th>Study design</th>
<th>Thyroid stimulating hormone (TSH) level of study group (mU/L)</th>
<th>Subjects</th>
<th>Study</th>
<th>Control</th>
<th>Reported effects of subclinical hyperthyroidism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leese et al, 1992&lt;sup&gt;47&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>&lt;0.05</td>
<td>691 subjects with subclinical hyperthyroidism on L-thyroxine</td>
<td>448 subjects euthyroid on L-thyroxine</td>
<td>No excess fracture risk in patients on L-thyroxine, even with suppressed TSH</td>
<td></td>
</tr>
<tr>
<td>Foldes et al, 1993&lt;sup&gt;52&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Case control</td>
<td>≤0.15; normal free triiodothyronine (T3) and free thyroxine (T4)</td>
<td>37 women with subclinical hyperthyroidism</td>
<td>68 age-matched euthyroid women</td>
<td>Low bone density at femoral neck and radius in postmenopausal women, but not premenopausal women, with subclinical hyperthyroidism</td>
<td></td>
</tr>
<tr>
<td>Faber &amp; Galloe, 1994&lt;sup&gt;46&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Meta-analysis</td>
<td>Unspecified</td>
<td>758 women with subclinical hyperthyroidism (from 13 studies)</td>
<td>3141 euthyroid women (from 13 studies)</td>
<td>Loss of bone mineral density at the distal forearm, femoral neck or lumbar spine in postmenopausal women, but not in premenopausal women</td>
<td></td>
</tr>
<tr>
<td>Mudde et al, 1994&lt;sup&gt;56&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Clinical trial</td>
<td>≤0.1; normal free T3 and free T4</td>
<td>8 post-menopausal women with subclinical hyperthyroidism from nodular goiter</td>
<td>8 post-menopausal women with nodular goiter; euthyroid after methimazole treatment</td>
<td>Distal, but not proximal, forearm bone mineral density was higher in the treated than in the untreated subjects with subclinical hyperthyroidism after 24 months</td>
<td></td>
</tr>
<tr>
<td>Uzzan, 1996&lt;sup&gt;41&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Meta-analysis</td>
<td>Unspecified</td>
<td>Total of 1250 subjects from 41 studies</td>
<td></td>
<td>Decreased bone mineral density at all skeletal sites in postmenopausal women, but not in premenopausal women</td>
<td></td>
</tr>
<tr>
<td>Marcocci et al, 1997&lt;sup&gt;45&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>&lt;0.01; normal free T3; free T4 elevated in 41%</td>
<td>34 men with subclinical hyperthyroidism</td>
<td>34 age- and weight-matched euthyroid men</td>
<td>No difference in bone mineral density at lumbar spine, femoral neck, Ward’s triangle or trochanter between groups</td>
<td></td>
</tr>
<tr>
<td>Faber et al, 1998&lt;sup&gt;57&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Clinical trial</td>
<td>&lt;0.2; normal free T3 and free T4</td>
<td>16 post-menopausal women with subclinical hyperthyroidism from nodular goiter</td>
<td>12 post-menopausal women euthyroid after radioactive iodine treatment for nodular goiter</td>
<td>2% bone loss per year in postmenopausal women, prevented by treatment of subclinical hyperthyroidism</td>
<td></td>
</tr>
<tr>
<td>Nuzzo et al, 1998&lt;sup&gt;42&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Mean 0.11 ± 0.07</td>
<td>40 pre-menopausal women with subclinical hyperthyroidism</td>
<td>General population data</td>
<td></td>
<td>No bone loss at the lumbar spine and right femoral neck in premenopausal women</td>
<td></td>
</tr>
<tr>
<td>Gürlek &amp; Gedik, 1999&lt;sup&gt;53&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Case control</td>
<td>≤0.1; normal free T3 and free T4</td>
<td>15 pre-menopausal women with subclinical hyperthyroidism</td>
<td>15 age-, height-, weight-, and smoking status-matched euthyroid women</td>
<td>No differences in femoral neck, lumbar, or forearm bone mineral density</td>
<td></td>
</tr>
<tr>
<td>-No difference in markers of bone turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3  Skeletal complications reported in patients with endogenous and exogenous subclinical hyperthyroidism (continued)

<table>
<thead>
<tr>
<th>Authors and references</th>
<th>Type of subclinical hyperthyroidism</th>
<th>Study design</th>
<th>Thyroid stimulating hormone (TSH) level of study group (mU/L)</th>
<th>Subjects</th>
<th>Control</th>
<th>Reported effects of subclinical hyperthyroidism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kumeda et al, 2000&lt;sup&gt;55&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Case control</td>
<td>&lt;0.4</td>
<td>19 pre-menopausal women with subclinical hyperthyroidism</td>
<td>30 pre-menopausal euthyroid women</td>
<td>-Increased markers of bone turnover in women with persistently low TSH values on low-dose anti-thyroid drugs for Graves' disease compared to women with normalized TSH values</td>
</tr>
<tr>
<td>Bauer et al, 2001&lt;sup&gt;50&lt;/sup&gt;</td>
<td>Exogenous and endogenous</td>
<td>Case cohort</td>
<td>297 women with fracture</td>
<td>398 women without fracture</td>
<td>-3-fold increased risk for hip fracture and 4-fold for vertebral fracture in women with TSH ≤0.1 mU/L</td>
<td>-No difference in fracture risk for women with TSH 0.1–0.5 mU/L</td>
</tr>
<tr>
<td>Jodar et al, 2001&lt;sup&gt;46&lt;/sup&gt;</td>
<td>Exogenous and endogenous</td>
<td>Case control</td>
<td>Mean 0.2 ± 0.3</td>
<td>17 men on L-thyroxine</td>
<td>32 men with treated Graves’ disease</td>
<td>-Bone density lower than expected in both groups</td>
</tr>
<tr>
<td>Sheppard et al, 2002&lt;sup&gt;48&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>Unspecified</td>
<td>23 188 patients prescribed L-thyroxine</td>
<td>92 732 sex-, age-, primary care practice-, and duration in database-matched patients not on L-thyroxine</td>
<td>L-thyroxine treatment an independent predictor of fracture risk in men but not in women</td>
</tr>
<tr>
<td>Baldini et al, 2002&lt;sup&gt;44&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>0.1–0.7</td>
<td>43 pre- and post-menopausal women with subclinical hyperthyroidism</td>
<td>46 pre- and post-menopausal euthyroid women</td>
<td>-No difference in bone mineral density between groups</td>
</tr>
<tr>
<td>Van den Eeden et al, 2003&lt;sup&gt;49&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>Unspecified</td>
<td>501 post-menopausal women with hip fracture</td>
<td>533 age-matched women without fracture</td>
<td>L-thyroxine treatment is not a risk factor for hip fracture in women &gt; age 65 years</td>
</tr>
<tr>
<td>Ugur-Altun et al, 2003&lt;sup&gt;54&lt;/sup&gt;</td>
<td>Endogenous</td>
<td>Case control</td>
<td>Mean 0.04 ± 0.02</td>
<td>8 pre-menopausal women with subclinical hyperthyroidism</td>
<td>10 euthyroid women</td>
<td>-No difference in bone mineral density in premenopausal women</td>
</tr>
<tr>
<td>Larijani et al, 2004&lt;sup&gt;43&lt;/sup&gt;</td>
<td>Exogenous</td>
<td>Case control</td>
<td>&lt;0.1; normal free T3 and free T4</td>
<td>50 pre-menopausal women with subclinical hyperthyroidism</td>
<td>158 pre-menopausal euthyroid women</td>
<td>-No difference in bone mineral density</td>
</tr>
</tbody>
</table>
ine. Another reported a significant association between L-thyroxine use and hip fracture in men, but not in women. This analysis did not differentiate between replacement and suppressive L-thyroxine therapy, and serum TSH values were not obtained. A third cohort study also reported that L-thyroxine replacement therapy is not a risk factor for hip fracture in women older than 65 years of age. The fourth cohort study found that a history of subclinical hyperthyroidism was associated with an increased risk for hip fracture in a case-cohort study of women over age 65 years. This study included women with both endogenous and exogenous subclinical hyperthyroidism, although the analysis adjusted for thyroid hormone use. Women with serum TSH concentrations <0.1 mU/L had an adjusted threefold increased risk for hip fracture and a fourfold increased risk for vertebral fracture, compared with women who had normal serum TSH concentrations. There was no significant difference in fracture risk for women with serum TSH values 0.1–0.5 mU/L compared with euthyroid women. In postmenopausal women, administration of estrogen may prevent L-thyroxine-induced bone loss in women with suppressed serum TSH values.

Endogenous subclinical hyperthyroidism and the skeleton

Results of bone density studies in subjects with endogenous subclinical hyperthyroidism have been inconsistent. One case-control study demonstrated low bone density in postmenopausal, but not premenopausal, women with endogenous subclinical hyperthyroidism. Recent small case-control studies have shown no effect of endogenous subclinical hyperthyroidism on bone mineral density in premenopausal women. By contrast, another study evaluated premenopausal women with Graves’ disease who had normal free T3 and free T4 concentrations after at least 10 months of treatment with low-dose (<50 mg/day propylthiouracil or <5 mg/day methimazole) antithyroid medications. Nineteen had persistently low serum TSH values and had significantly higher concentrations of serum and urine bone turnover markers than women with normalized serum TSH values.

Two small clinical trials have addressed the bone effects of treatment of endogenous subclinical hyperthyroidism in postmenopausal women with nodular goiter. In one, 8 of 16 subjects were treated with methimazole. After 2 years, distal (but not proximal) forearm bone mineral density was improved in the treated group. Another study compared bone mineral density in 12 women who were euthyroid following radioactive iodine treatment with 16 untreated women with subclinical hyperthyroidism and found that the treatment prevented decreases in bone density.

Symptoms

In clinical practice, patients with subclinical hyperthyroidism commonly complain of thyrotoxic symptoms such as anxiety, heat intolerance, tremors, sweaty skin, insomnia, forgetfulness, or mood disorders. The question is whether patients with subclinical hyperthyroidism have symptoms of thyrotoxicosis or altered mood more frequently or to a greater degree than an appropriate control group.

One study assessed symptoms and signs of thyrotoxicosis and quality of life in 23 outpatients with subclinical hyperthyroidism (mean age 43 years, mean serum TSH 0.15 mU/L) and in 23 age-, sex-, and lifestyle-matched outpatient controls, using questionnaires. There was a significantly increased prevalence of palpitations, nervousness, tremor, heat intolerance, sweating in patients with subclinical hyperthyroidism, and quality of life was impaired.

The SWAN Study investigators reported a significantly increased prevalence of fearfulness and prolonged menses in pre- and perimenopausal women with serum TSH concentration <0.5 mU/mL, compared with euthyroid women. Low serum TSH values were not associated with changes in menstrual regularity, menopausal symptoms, or reproductive hormone concentrations.

Another study evaluated the influence of experimentally induced subclinical hyperthyroidism on cognitive function in a double-blind crossover design. Twenty-four young men were treated with 300 μg L-thyroxine or placebo for 3-week periods. Event-related brain potentials demonstrated early effects of subclinical hyperthyroidism on central information processing, with treated men requiring a greater effort to complete a visual search task.

Finally, four recent studies have examined the relationship between subclinical hyperthyroidism and dementia. Kalmijn et al prospectively investigated the relationship between thyroid status and the risk of dementia and Alzheimer’s disease among a random sample of 1843 subjects, aged ≥55 years. They demonstrated that individuals with subclinical hyperthyroidism at baseline (serum TSH <0.4 mU/L) had a more than threefold increased risk of dementia and of Alzheimer’s disease over a mean 2-year follow-up. The risk of dementia was particularly increased in patients with subclinical hyperthyroidism and elevated antithyroidperoxidase antibody titers. A similar association has also been reported by Dobert et al in a case-control study. Of 77 patients with dementia, 29% had serum TSH values ≤0.5 mU/L compared with 10% of controls, with a stronger association between vascular dementia and low serum TSH than between Alzheimer’s disease and low serum TSH. A second case-control study found that TSH levels, within the normal range, were associated with a twofold increase in risk factor for Alzheimer disease. However, van der Cammen et al did not find an association between serum TSH levels and the diagnosis of Alzheimer’s in a cross-sectional survey of 829 consecutive geriatric patients.

Progression to overt hyperthyroidism

Few studies have investigated the natural history of endogenous subclinical hyperthyroidism, in part because sensitive TSH assays are only relatively newly available. In one
series of 349 patients with autonomously functioning thyroid nodules, overt hyperthyroidism was present in 12.5% of patients <60 years old at baseline, and in 56.5% of older patients; it was more common in women than men (6:1 ratio), and in patients with nodules ≥3 cm in diameter. Of 159 untreated subjects who were euthyroid at baseline, 14 (9%) became overtly thyrotoxic within 1 to 6 years. In another series of 375 patients with untreated euthyroid autonomously functioning thyroid nodules, 67 developed overt hyperthyroidism over 18-36 months. In this study, the lack of a nocturnal TSH surge was highly predictive of the development of overt hyperthyroidism. It is likely that patients in all these series experienced a transitional period of subclinical hyperthyroidism.

After 11 months of follow-up, only one of a subset of 15 patients with subclinical hyperthyroidism (serum TSH <0.1 mU/L) from a larger cohort developed overt hyperthyroidism. This study is limited by the small sample, short duration of follow up, and by the fact that nonthyroidal illness syndrome was not excluded.

Mortality

A recent study assessed the long-term effects of subclinical hyperthyroidism in 1191 British individuals, age ≥60 years, not on thyroid medications. Serum TSH concentrations were measured at baseline, and mortality was assessed over 9733 person-years of follow-up. All-cause and cardiovascular mortality was significantly increased at 2, 3, 4, and 5 years, but not at 10 years, in subjects with baseline serum TSH concentrations <0.5 mU/L. No difference was found between subjects with TSH values <0.1 and those with TSH values of 0.1–0.5 mU/L. Baseline free T4 and free T3 values did not correlate with mortality.

In a 2-year prospective cohort of chronically ill patients aged 64–87 years, mortality rates for subjects with low but detectable baseline serum TSH values did not differ from those of euthyroid subjects. However, patients with completely suppressed baseline serum TSH values and normal peripheral thyroid hormone values had a significantly higher mortality rate (5 of 8) than those with normal baseline serum TSH values (18 of 64).

Therapeutic options for subclinical hyperthyroidism

An accurate diagnosis is the basis for successful treatment of endogenous subclinical hyperthyroidism. Four treatment options are available for endogenous subclinical hyperthyroidism: continued observation, anti-thyroid drugs, radioiodine therapy, and surgery. Because the systematic treatment of each disorder causing endogenous subclinical hyperthyroidism is beyond the scope of the present report, we summarize our diagnostic and therapeutic options for the management of subclinical hyperthyroidism in the Figure. Antithyroid drugs may be indicated when serum TSH suppression is caused by thyroid hyperfunction with a high or normal thyroid radioiodine uptake, but they are contraindicated when subclinical hyperthyroidism is due to thyroiditis where the radioactive iodine uptake is low. Radioactive iodine ablation and, less frequently, surgery may also be indicated in those patients with hyperfunctioning nodules or Graves’ disease.

Of particular concern is exogenous subclinical hyperthyroidism. Many patients worldwide are treated with thyroid hormone, either as therapy for hypothyroidism, as TSH-suppressive therapy for benign diffuse colloid goiter or thyroid cancer, or for other thyroid and nonthyroid disorders. In hypothyroid patients, the goal of L-thyroxine therapy is to maintain serum TSH concentrations within the normal range. Hypothyroid patients are at risk of overtreatment with L-thyroxine. In nontoxic goiter or thyroid cancer patients, suppression of the serum TSH concentration is often recommended.

Current opinions

The management of patients with endogenous subclinical hyperthyroidism remains controversial. The results of many reports have not provided definitive answers concerning the following important questions: Who needs screening? Who needs therapy? When should treatment be started? Published recommendations vary widely (Table 4).

American Thyroid Association members were recently surveyed about their approach to several hypothetical patients with subclinical hyperthyroidism. Most felt that further evaluation was warranted for patients with low serum TSH values. A majority of respondents favored observation for young women with subclinical hyperthyroidism, and reserved active therapy for postmenopausal women.

Conclusions

Subclinical hyperthyroidism is relatively prevalent in the general population. Some, but not all, studies suggest that it has adverse cardiac consequences. Some studies have shown adverse effects on bone density in postmenopausal women, but no studies to date have clearly demonstrated an increased risk for fracture in subclinical hyperthyroidism. Symptoms, if present, are subtle.
It seems clear that iatrogenic (exogenous) subclinical hyperthyroidism should be avoided except when clearly indicated for nontoxic colloid goiter or for thyroid cancer TSH suppression. When serum TSH values are below the normal range in patients who are not taking exogenous thyroid hormone, we would recommend a repeat measurement. If the repeated TSH value is again near or below the detectable limit, we would pursue other tests to determine the underlying etiology (Figure). If the repeated TSH measurement is below the normal range but readily detectable, we would suggest watchful waiting.

Although there have been no controlled trials of the efficacy of therapy for endogenous subclinical hyperthyroidism, small observational studies in patients with this disorder suggest that antithyroid therapy may be efficacious in improving bone metabolism and cardiac abnormalities. In the absence of definitive data, we believe that therapeutic decisions must be individualized, but that early treatment may be beneficial in patients with subclinical hyperthyroidism due to toxic multinodular goiter or Graves’ disease to prevent progression to overt hyperthyroidism. Patients with underlying cardiac risk factors and postmenopausal women with low bone density are particularly likely to benefit from therapy for endogenous subclinical hyperthyroidism.

Table 4 Published guidelines for screening and treatment of subclinical hyperthyroidism

<table>
<thead>
<tr>
<th>Author</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>American College of Physicians 1998&lt;sup&gt;80&lt;/sup&gt;</td>
<td>“It is reasonable to screen women older than 50 years of age for unsuspected but symptomatic thyroid disease . . . The treatment of patients found by screening to have persistent subclinical hyperthyroidism has not been studied . . . Patients who are found to have relatively specific symptoms and signs (such as goiter, nodule, eye findings of Graves’ disease, or tremor) should be referred to an endocrinologist for consideration of treatment. The management of patients without clinical findings is not clear because most of these persons remain healthy.”</td>
</tr>
<tr>
<td>American Thyroid Association 2000&lt;sup&gt;81&lt;/sup&gt;</td>
<td>“. . . recommends that adults be screened for thyroid dysfunction by measurement of the serum TSH [thyroid stimulating hormone] concentration, beginning at age 35 years and every 5 years thereafter. The indication for screening is particularly compelling in women.”</td>
</tr>
<tr>
<td>American Association of Clinical Endocrinologists 2002&lt;sup&gt;82&lt;/sup&gt;</td>
<td>“In patients with subclinical hyperthyroidism attributable to nodular thyroid disease, treatment seems warranted because of the high rate of conversion to clinical hyperthyroidism . . . All patients with subclinical hyperthyroidism should undergo periodic clinical and laboratory assessment to determine individual therapeutic options”</td>
</tr>
<tr>
<td>American Academy of Family Physicians 2003&lt;sup&gt;83&lt;/sup&gt;</td>
<td>“. . . recommends against the use of thyroid function tests for screening for thyroid disease in patients less than 60 years old and not neonates.”</td>
</tr>
<tr>
<td>American College of Obstetrics and Gynecology 2002&lt;sup&gt;84&lt;/sup&gt;</td>
<td>“It is appropriate to perform indicated testing of thyroid function in [pregnant] women with a personal history of thyroid disease or symptoms of thyroid disease. The performance of thyroid function tests in asymptomatic pregnant women who have a mildly enlarged thyroid is not warranted.”</td>
</tr>
<tr>
<td>Consensus panel 2004 (sponsored by the American Thyroid Association, Endocrine Society, and American Association of Clinical Endocrinologists)&lt;sup&gt;78&lt;/sup&gt;</td>
<td>“The consequences of subclinical thyroid disease (serum TSH 0.1–0.45 mU/L or 4.5–10.0 mU/L) are minimal and we recommend against routine treatment of patients with TSH levels in these ranges. There is insufficient evidence to support population-based screening. Aggressive case-finding is appropriate in pregnant women, women older than 60 years, and others at high risk for thyroid dysfunction.”</td>
</tr>
<tr>
<td>American Thyroid Association, Endocrine Society, and American Association of Clinical Endocrinologists&lt;sup&gt;85&lt;/sup&gt;</td>
<td>“. . . observe and monitor patients with partial TSH suppression (0.1–0.4 mU/L, but . . . treat patients with complete TSH suppression (&lt;0.1 mU/L).”</td>
</tr>
<tr>
<td>United States Preventive Services Task Force 2004&lt;sup&gt;86&lt;/sup&gt;</td>
<td>“. . . found fair evidence that the thyroid-stimulating hormone (TSH) test can detect subclinical thyroid disease in people without symptoms of thyroid dysfunction but poor evidence that treatment improves clinically important outcomes in adults with screen-detected thyroid disease . . . the evidence is insufficient to recommend for or against routine screening for thyroid disease in adults.”</td>
</tr>
</tbody>
</table>

References


70. Surks MI, Ortiz E, Daniels GH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA. 2004;291:228–238.
What triggered this sudden eruption?

Lisa Scatena, MD

University of Colorado, Departments of Dermatology and Internal Medicine

Parwathi “Uma” Paniker, MD, Images in Dermatology Editor

University of Colorado Health Sciences Center

Itchy, scaly papules and plaques abruptly emerged on the trunk and proximal extremities of a healthy 22-year-old man (Figure). The only illness in his recent medical history is an occurrence of acute pharyngitis two weeks earlier. After a rapid streptococcal test proved positive, the patient was initially treated with azithromycin. A course of levofloxacin followed.

A classic connection

The clinical presentation of this case is virtually pathognomonic for guttate psoriasis: a recent group A streptococcal pharyngitis is followed within two to three weeks by an eruption of small (1-15 mm), round, salmon-pink papules and plaques with silvery scale. Numerous droplet-like lesions are a hallmark of the disease and give it its name. In Latin, “gutta” means “drop”; think of the abbreviation “gtt.” used to denote drops (or guttae) when writing a prescription.

Guttate psoriasis generally occurs in children and young adults. It is equally prevalent in males and females, and no predilection for a particular race has been documented. A family history of psoriasis is often elicited in patients with the ailment. Reinforcing that observation is the finding that people carrying human leukocyte antigens (HLA) 13 and 17 appear to have a genetic predisposition to psoriasis, including guttate psoriasis.

Acute episodes are thought to be associated with T-cell stimulation by streptococcal superantigens. An upper respiratory infection precedes an outbreak in up to 66% of patients. In one study, antistreptolysin O (ASO) titers exceeding 200 units were found in 56% of participants with guttate psoriasis. Perianal streptococcal infections have also been associated with the disorder.

Clearing the lesions

In this patient, other diagnoses to consider include mycosis fungoides or cutaneous T-cell lymphoma, pityriasis lichenoides chronica, pityriasis rosea, secondary syphilis, and small plaque parapsoriasis. The lesions of guttate psoriasis rarely involve the palms and soles, helping to distinguish it from secondary syphilis. Certain medications, notably beta-blockers and lithium, can cause an eruption that looks like guttate psoriasis, and this possibility should be pursued when patients are taking such drugs.

The diagnosis is a clinical one and rarely requires biopsy. However, an evaluation for suspected cases of guttate psoriasis should include ASO titers, a bacterial throat or perianal culture, and a VDRL to rule out syphilis.

Guttate psoriasis is often self-limited and can resolve spontaneously within weeks to months of the initial eruption. Nonetheless, administration of an antibiotic with bactericidal activity against Group A streptococcus is recommended. For example, penicillin VK, 1 gm/d in divided doses for 10 to 14 days, or erythromycin, 1 gm/d in divided doses for 10 to 14 days, can be prescribed. Phototherapy with narrow-band or broad-band UVB can expedite the clearing of guttate lesions. Judicious use of natural sunlight can be helpful as well. Topical corticosteroids in classes III through VII also have a role in limited disease or in pruritic eruptions.

As in plaque psoriasis, patients with guttate psoriasis are susceptible to the Koebner phenomenon or isomor-
phic effect. That is, trauma to previously unaffected areas can give rise to new lesions. Advise patients to avoid rubbing, scratching, or otherwise injuring the skin to avoid secondary development of more papules and plaques. It is estimated that 33% to 68% of patients with a history of guttate psoriasis will later go on to develop chronic plaque psoriasis.

References


**Figure:** Many papules and plaques of various sizes cover this patient’s trunk and extend to the proximal extremities.
Persistent lower abdominal and groin pain: What is the diagnosis?

R. Scott Stephens, MD, Cynthia Brown, MD

Charles M. Wiener, MD, Diagnostic Dilemma Editor

Department of Medicine, Johns Hopkins School of Medicine.

Patient history

A 54-year-old man is admitted for persistent lower abdominal and groin pain that began after repair of bilateral inguinal hernias 7 months previously. He had already undergone multiple evaluations at other hospitals. These workups included computed tomography (CT), barium contrast studies, magnetic resonance imaging, colonoscopy, and esophagogastroduodenoscopy. Two months before his present admission, he required exploratory laparotomy for increasing pain. The discovery of necrotic omentum and pericholecystitis led to omentectomy and cholecystectomy. The pathology was reported as necrotic without a specific diagnosis.

The pain continued after the laparotomy. It is periumbilical, radiating into his groin and legs, and it worsens with movement, especially walking. The patient also complains of sporadic episodes of testicular pain, bowel urgency, diaphoresis, nausea, and vomiting. He denies diarrhea, hematemesis, melena, or rectal bleeding. His symptoms do not occur at rest and are not affected by eating. He has lost about 50 pounds over the preceding 6 months.

The patient has a history of hypertension that has recently worsened and coronary artery bypass grafting in 1997. He is also legally blind as a result of giant cell arteritis that was diagnosed in 1995. At that time, he was treated with prednisone for more than one year and also briefly with cyclophosphamide. When questioned, he reports further deterioration of his visual acuity, “flashes of light” in his visual fields, headache, and jaw claudication. Medications on admission include aspirin, hydrochlorothiazide, hydromorphine (prn), lansoprazole, metoprolol, prednisone (10 mg daily), and quinapril. The patient had a 20-pack-per-year history of tobacco use but quit at the time of his bypass surgery.

On admission

On physical examination, the patient appears comfortable. His blood pressure is 170/100 mm Hg, his heart rate is 88 bpm, and he is afebrile. The patient has 20/70 vision in his right eye and 20/400 in his left eye. Fundoscopic examination is unremarkable except for changes consistent with prior inflammation in his left eye. He has normal first and second heart sounds without murmurs, and an S4 is present. There are no carotid, renal, abdominal, or femoral bruits. His lungs are clear to auscultation. Bowel sounds are normal. Abdominal palpation demonstrates minimal diffuse tenderness without rebound or guarding. No masses are present, and the stool is negative for occult blood. During the examination, the patient develops Raynaud’s phenomenon in his right hand that persists for several minutes. His neurologic examination is intact except for his decreased vision. The patient has no rashes. Ambulation causes severe abdominal pain and vomiting.

Admission laboratory studies reveal anemia with a hematoctrit of 35.1% and a mean corpuscular volume of 97.8, a lactic acid of 2.5 mmol/L, an erythrocyte sedimentation rate of 72 mm/hour, a BUN of 17 mg/dL, and a creatinine of 0.8 mg/dL. The patient has no proteinuria or hematuria. Tests for antinuclear antibodies, anti-double-stranded-DNA antibodies, and antineutrophil cytoplasmic antibodies (AN-
CAAs) prove negative. Serologies for hepatitis B and C are negative, too.

An abdominal CT scan shows a 50% narrowing at the origin of the superior mesenteric artery (Figure 1). On angiogram, this artery and its branches appear smooth, beaded, and tapered, as do the splenic, hepatic, inferior mesenteric, left renal, and right iliac arteries. These results prompt treatment with intravenous corticosteroids for presumed vasculitis. Two weeks later, the patient has sudden severe abdominal pain. Emergent laparotomy discloses necrosis of the entire colon and terminal ileum. Microscopic examination of a mesenteric muscular artery indicates intraluminal thrombosis with areas of acute inflammation between the intima and media (Figure 2).

What is the diagnosis?

The findings on CT scan, angiography, and pathology are consistent with a medium-vessel vasculitis, and the patient’s constellation of symptoms fulfills the diagnostic criteria for polyarteritis nodosa (PAN). A review of pathology results from prior surgeries supports the diagnosis—multiple foci of inflammation, necrosis, and eosinophilic infiltrate involving muscular arteries were documented. In addition, the temporal artery biopsy 8 years earlier is consistent with arteritis.

PAN has been defined as a necrotizing inflammation of medium-sized or small arteries, which spares the smallest blood vessels and is not associated with glomerulonephritis. The disease is rare, with estimated incidences ranging from 4.6 to 77 per 1,000,000. It occurs in all races, and men and women are equally affected, with the majority of cases occurring in the 5th and 6th decades. While the cause remains unknown, a percentage of cases are associated with viral infection, particularly infection with hepatitis B virus (HBV). In the 1970s, 36% of patients with PAN were infected with hepatitis B; however, this number has now decreased to less than 10%, probably because the overall incidence of HBV infection has decreased. Other viruses have also been implicated, including human immunodeficiency virus, cytomegalovirus, parvovirus B19, human T-cell lymphotropic virus, and hepatitis C virus. It is suspected that in virus-associated PAN, deposition of soluble immune complexes consisting of viral antigens and specific antibodies leads to vascular lesions.

The diagnosis of PAN requires high clinical suspicion. Ill-defined symptoms such as fever, weight loss, arthritis, and arthralgias dominate at the time of presentation. Mono-neuritis multiplex is common, and skin nodules, purpuric
lesions, or livedo reticularis may be the first sign of disease in a sizable minority of patients. Symptoms consistent with temporal arteritis may also be seen; prompt temporal artery biopsy is vital. Abdominal pain is often present and may be due to either intestinal ischemia or vasculitis of the appendix or gallbladder. Orchitis, a classic symptom, is more common in HBV-associated PAN. Cardiac manifestations, most often congestive heart failure, can be seen, although angina and myocardial infarctions due to coronary vasculitis are believed to be relatively rare. Hypertension, which can be malignant, is frequently observed. Ischemic strokes and cerebral hemorrhages are seen in 10% of PAN patients. In 1990, the American College of Rheumatology set forth 10 criteria, the presence of 3 of which yielded an 82.2% sensitivity and 86.6% specificity for the diagnosis of PAN. These criteria are shown in the Table.

Proof of necrotizing vasculitis in small and medium-sized arteries is considered diagnostic. This can be accomplished with biopsy. Or, contrast angiography can be used to detect characteristic aneurysms. Up to 1 cm in diameter, these can be found within the renal, mesenteric, and hepatic vasculature. The beaded, tapered appearance noted in this patient, although suggestive of PAN, can be seen in other forms of necrotizing angiitis, such as Wegener’s granulomatosis and systemic lupus erythematosus. Aneurysms may regress as the artery heals.

Laboratory studies are often ambiguous. Patients generally have elevations in erythrocyte sedimentation rate and C-reactive protein. They are usually anemic, and leukocytosis with eosinophilia may be present. Tests for p-ANCAs and c-ANCAs are commonly negative, and complement levels are usually normal. CT examinations also tend to reveal vague findings, such as bowel wall thickening.

Immunosuppression is the cornerstone of treatment. First used in the 1950s, corticosteroids improved the 5-year survival rate from 10% to 50%, and they continue to be fundamental to any therapeutic regimen. For HBV-negative patients, methylprednisolone, 15 mg/kg, is used for the first 3 days, followed by prednisone, 1 mg/kg/day. The dosage is tapered as the clinical status improves. Cyclophosphamide is also used in severe cases. Traditionally used at a dosage of 2 mg/kg/day for one year, cyclophosphamide is increasingly being administered in a pulsed regimen. Although combination therapy with cyclophosphamide and corticosteroids yields a 5-year survival rate of 82%, the potential for severe side effects generally limits use of cyclophosphamide to 1 year. Relapses in completely recovered patients are rare.

Patients with HBV infection are treated differently because sustained immunosuppression could increase the risk of chronic hepatitis and cirrhosis. Corticosteroids are used first to diminish the inflammatory process, but antiviral treatment with lamivudine or interferon alfa is increasingly recognized as a vital part of treatment for these patients.

Surgery is necessary for appendicitis, acute cholecystitis, and ischemia, perforation, or necrosis of the bowel. However, abdominal complications due to mesenteric vasculitis are observed in up to 50% of patients with PAN, and surgical intervention significantly worsens the prognosis. Multiple perforation sites are frequently found. Concomitant immunosuppression inhibits wound healing and promotes infection. Nevertheless, aggressive medical therapy should be continued. Specifically, the corticosteroid dosage should not be decreased, and cyclophosphamide treatment should not be delayed.

Prognosis is estimated with the Five Factor Score. The absence of renal insufficiency, proteinuria (>1 g/day), cardiomyopathy, gastrointestinal tract or mesenteric involvement, and central nervous system involvement predict a 5-year mortality rate of 12%. The presence of 1 of these factors suggests a 5-year mortality rate of 25%, and the presence of 2 or more factors indicates a 5-year mortality rate of 46%.

Our patient was treated initially with IV methylprednisolone, 1 gm/day for 3 days, followed by IV methylprednisolone, 60 mg/day. Despite this, the patient developed an acute abdomen 2 weeks into therapy. Surgery uncovered frank necrosis of his entire colon and part of his terminal ileum. All surgical margins were positive for active ischemia and vasculitis. Within weeks, he again had severe abdominal pain, accompanied by gastrointestinal bleeding and lactic acidosis. Emergent surgery was scheduled, but the patient went into cardiac arrest. He was resuscitated and brought to the operating room where a laparotomy confirmed diffuse ischemia and necrosis of the small bowel. He was transferred to the surgical intensive care unit, where he soon died.

References


A treatment option for some failing hearts

Julia H. Indik, MD, PhD, ECG Image of the Month Editor

University of Arizona, Sarver Heart Center.

In recent years, the prognosis for patients with heart failure has improved substantially with advances in both pharmacologic treatments and the application of implantable devices. One new technology, cardiac resynchronization therapy (CRT), has emerged as an effective method for selected patients. Also known as biventricular pacing, it has been demonstrated in major trials and, when coupled with a defibrillator, to improve survival as well. For these reasons, it is important to know which patients are possible candidates for biventricular pacing and to understand its mode of action.

What it does

The goal of biventricular pacing is to achieve resynchronization of left ventricular wall motion. For example, wall motion is generally dyssynchronous in the setting of left bundle branch block (LBBB). When electrical conduction can only proceed along the right bundle, the ventricular septum is stimulated from right to left with very slow activation of the remainder of the left ventricle.

In particular, the lateral wall may be set in motion 100 msec or more after activation of the septum. As a result, lateral wall movement may occur so late in the cardiac cycle that it contributes little to mechanical systole. Placement of a pacing lead within a coronary sinus branch that lies along the epicardial lateral wall of the left ventricle allows for simultaneous activation of the lateral wall and septum, improving the efficiency of ventricular contraction.

When it’s used

CRT should be considered for patients with heart failure who have advanced left ventricular dysfunction with a QRS duration of at least 130msec, who meet the criteria of New York Heart Association functional class III or IV, and who are already receiving optimal medical therapy. In some cases, echocardiography using tissue Doppler imaging is needed to discern whether significant dyssynchrony exists.

For example, a biventricular ICD has been placed in a patient with LBBB. A baseline ECG displays a sinus rhythm with a prolonged PR interval of 240msec and a very wide QRS of 220msec (Figure 1). The typical hallmarks of LBBB are evident: the voltage in V1 is negative; the QRS duration is greater than 120msec; and leads I, V5 and V6 show a wide and notched R wave with no Q wave. A normal QRS complex should display a very small Q wave in these three leads.

An ECG following the procedure clearly illustrates the favorable effects of biventricular pacing (Figure 2). Note that there is a sinus rhythm of 80 bpm, and the PR interval has been shortened to 130 msec, which improves cardiac filling during diastole. The QRS complex, slightly narrower than baseline at 180msec, represents a fusion of simultaneous pacing from both the right and left ventricles. A look at data from lead V1 indicates that the paced QRS complex is now almost positive, and in fact, the presence of left ventricular pacing almost gives it the appearance of right bundle branch block.

Furthermore, the paced QRS complex in lead I is now negative, indicating that voltage is traveling away from the left side of the heart. This is the opposite of what would be
seen in standard pacing of only the right ventricle. In that situation, the paced QRS complex resembles the pattern produced by LBBB—it would be negative in lead V1 and have a large positive R wave in lead I. A chest x-ray clearly depicts the device and the positioning of the pacing wires (Figure 3).

Figure 1 The patient’s initial ECG provides characteristic evidence of left bundle branch block.

Figure 2 With biventricular pacing, the QRS is now negative in lead I and almost positive in V1. The PR interval has been shortened to 130 msec to improve ventricular filling during diastole.
**Figure 3** The black arrows highlight pacing wires in the right atrium and the right ventricle. A third pacing wire is in the coronary sinus where it enters a sinus branch that reaches the lateral wall of the left ventricle, as shown by the white arrows in the posterior-anterior view. In the lateral view, the third wire runs in a posterior position.

**References**

CLINICAL RESEARCH STUDY

The impact of low health literacy on the medical costs of Medicare managed care enrollees

David H. Howard, PhD,a Julie Gazmararian, MPH, PhD,b Ruth M. Parker, MDc

aDepartment of Health Policy and Management, 
bEmory Center on Health Outcomes and Quality, Rollins School of Public Health, and 
cDepartment of Medicine, Emory University School of Medicine, Emory University.

ABSTRACT

PURPOSE: To examine the impact of low health literacy on medical care use and costs.

METHODS: The study sample consisted of 3260 noninstitutionalized elderly persons enrolling in a Medicare managed care plan in 1997 in Cleveland, Ohio; Houston, Texas; South Florida; and Tampa, Florida. Health literacy—the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions—was measured using the Short Test of Functional Health Literacy in Adults. We used a 2-part regression model to examine the association between health literacy and medical costs, adjusting for age, sex, race/ethnicity, education, income, alcohol and tobacco consumption, and comorbid conditions. Results are presented as mean differences (with 95% confidence intervals [CI]) between the inadequate and adequate groups and, separately, the marginal and adequate groups.

RESULTS: When compared to those with adequate health literacy, emergency room costs were significantly higher ($108; 95% CI: $62 to $154; P <0.0001) among those with inadequate health literacy, while differences in total ($1551; 95% CI: −$166 to $3267; P = 0.08) and inpatient ($1543; 95% CI: −$89 to $3175; P = 0.06) costs were marginally significant. Total costs were higher in the marginal health literacy group, but the difference was not significant ($596; 95% CI: −$1437 to $2630; P = 0.57).

CONCLUSIONS: Persons with inadequate health literacy incur higher medical costs and use an inefficient mix of services.

© 2005 Elsevier Inc. All rights reserved.

KEYWORDS: Cost; Cost analysis; Health expenditures; Health knowledge; Patient education

Research on “health literacy”—the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions1,2—has gained momentum over the last 5 years.1,3 Health literacy was recently added to the nation’s Healthy People 2010 objectives2 and listed, along with self-management, as one of 20 priority areas in which quality improvement could transform health care in America.4 In addition, the Institute of Medicine recently released a report that calls attention to the negative consequences of low health literacy for patients and the health system.5

About one-half of the adult U.S. population has low functional literacy skills,6 meaning that they lack the literacy skills needed for full participation in American society.7 Although there are no reliable population-based estimates of health literacy, surveys in narrow subpopulations suggest that at least one-third of adults in the United States lack
health literacy skills as well. Studies have shown that patients with low levels of health literacy receive fewer preventive services, frequently fail to follow medication instructions, and have worse health outcomes. Moreover, patients with low health literacy skills are more likely to have higher utilization of health care services.

Despite the growing research base documenting the extent and impact of low health literacy, there has been a paucity of data examining the financial costs associated with health literacy. According to one widely cited study, low health literacy increases aggregate medical spending by $73 billion annually (1998 dollars). Lacking data containing both measures of health utilization and literacy, the authors imputed literacy levels to subjects in the Survey of Income and Program Participation based on demographic factors, employment status, and education levels. Because these variables are independently associated with health care spending, the estimate of the relationship between the imputed literacy measure and utilization is probably confounded.

Few other studies have addressed the subject of literacy and medical spending, in part due to the difficulty of obtaining suitable data sources. Weiss and Palmer examined the association between general literacy and charges in a sample of Arizona Medicaid beneficiaries. They found that charges were more than $10,000 higher among beneficiaries with low literacy. However, the small sample size (18 subjects had low literacy, 56 had higher literacy) limits the conclusions that can be drawn from the analysis because outliers may skew results.

Because of confounding, small sample sizes, and atypical samples, existing estimates of the relationship between health literacy and costs are not suitable for predicting the financial ramifications of health literacy interventions. Given the growing interest in programs to improve health literacy and make the health system easier to navigate for persons with low health literacy, however, it is important to gain an accurate understanding of its relationship to health care costs. The purpose of this study was to measure the relationship between health literacy and medical costs using one of the few datasets that contains reliable measures of both.

Methods

Study sample

Enrollment of study participants and data collection, which were conducted by the Prudential Center for Healthcare Research (now the Emory Center on Health Outcomes and Quality), have been described in detail previously. Individuals newly enrolling in the Medicare managed care plans of Prudential Healthcare in 4 locations (Cleveland, Ohio; Houston, Texas; South Florida, and Tampa, Florida) between December 1996 and August 1997 were eligible to participate. New members were contacted 3 months after enrollment, and those meeting the eligibility criteria were asked to complete an in-person survey. To be included in the study, members had to be comfortable speaking either English or Spanish, live in the community, and possess adequate visual and cognitive function. Persons with severe cognitive impairment (2% of those originally contacted) were excluded. Spanish-speaking patients were interviewed in Spanish. Of the 7471 enrollees who were originally contacted, 3247 refused to participate.

Baseline survey

Health literacy and selected demographic and health characteristics of the population were obtained from the baseline in-person survey. Health literacy was measured using the Short Test of Functional Health Literacy in Adults. The test uses actual materials that patients might encounter in the health care setting and consists of 2 parts. The reading comprehension section is a 36-item test that measures patients’ ability to understand instructions for preparation for an upper gastrointestinal tract radiograph series and the patient “Rights and Responsibilities” section of a Medicaid application. The numeracy section is a 4-item test measuring patients’ ability to comprehend directions for taking medicines, monitoring blood glucose, and keeping clinic appointments. Based on their responses, subjects were classified as having either “inadequate,” “marginal,” or “adequate” health literacy. The survey also included questions about respondents’ age, sex, race/ethnicity, income, education, tobacco and alcohol consumption, and self-reported chronic conditions. Prior studies have demonstrated high levels of agreement between patient self-reports of major chronic conditions and medical records.

Administrative data

We obtained cost and utilization data from Prudential Healthcare’s administrative claims files and computed annual use and health expenditures from the date of enrollment by site of care (inpatient, outpatient, emergency room, and pharmacy). The cost for each service is the sum of Prudential Healthcare’s reimbursement and the beneficiary’s out-of-pocket payment. Because there is a lack of consensus about the appropriate index for inflating historical medical costs to present levels, we leave costs stated in terms of 1997 dollars.

Statistical analysis

Univariate comparisons between respondents with inadequate, marginal, and adequate health literacy were performed using 1-way analysis of variance tests for continuous variables and chi-squared tests for categorical variables. The modified 2-part regression model proposed by Mullahy was used to measure differences in use and costs. Two-part models are the standard statistical frame-
work in empirical health economics for measuring the impact of covariates on medical costs. Because of the highly skewed distribution and presence of a large number of “0” values, standard regression methods yield inaccurate predictions of costs. The 2-part model attempts to more accurately mimic the empirical distribution of medical spending by splitting the distribution into 2 parts and allowing the impact of independent variables, such as health literacy, on the probability of using medical care to be independent of their impact on the costs of medical care for those who use it.

The first stage of the 2-part model is a logistic regression estimated on the entire sample, where the dependent variable equals 1 if the enrollee used care and 0 otherwise. The second part is a nonlinear regression estimated on only those enrollees who used care. This second component posits that medical expenditures equal the exponentiated sum of a linear combination of coefficients and covariates. Unlike the standard 2-part model, the modified 2-part model does not require that log dollars be retransformed into constant dollars, which is problematic when the data exhibit heteroskedasticity.

Both parts of the 2-part model include as covariates a dummy variable for health literacy level, as well as controls for age, sex, race/ethnicity (white, black, Spanish speaking, other), income (<$10 000, $10 000–$25 000, $25 000–$35 000, >$35 000, no response), education (<8 years of schooling, 9–11 years, 12 years, some college, college degree), tobacco (never, former, current) and alcohol (none, light to moderate, heavy) consumption, and self-reported comorbid conditions (heart attack, angina, stroke, high blood pressure, chronic obstructive pulmonary disease, cancer, diabetes, arthritis, depression).

Parameter estimates from the 2-part model are difficult to interpret, so we present results as mean differences (with 95% confidence intervals [CI]) between the inadequate and adequate groups and, separately, the marginal and adequate groups. The procedure for computing mean differences is as follows. First, predicted values are computed by using the regression model to estimate what use and spending would be for every respondent if they had inadequate, marginal, and adequate health literacy (for a total of 3 predicted values per respondent per outcome). Next, we subtract the respondent’s predicted value for adequate health literacy (the reference level) from the predicted values for marginal and inadequate health literacy. Finally, we compute the average of differences over the entire sample. Computing differences in this manner nets out the impact of observable covariates, such as age, on use and spending. In statistical terminology, this procedure yields the “sample average treatment effect” of marginal and inadequate health literacy on use and costs. Standard errors are computed via simulation. In keeping with previous analyses, total costs were analyzed by summing predicted values for site-specific expenditures.

Univariate analyses were performed in Stata (Stata Corporation; College Station, Texas). Multivariate analyses and simulations were performed in MATLAB (The MathWorks; Natick, Massachusetts).

Sensitivity analyses

Inclusion of education and comorbid conditions in the regression models may lead to estimates of the impact of health literacy on use and costs that are biased downward. In the case of education, variables measuring years of schooling and educational attainment may pick up some of the variation in use and costs that ought to be legitimately attributed to health literacy. In the case of comorbid conditions, low health literacy may be a cause of, rather than incidental to, the development and progression of chronic diseases. To examine the sensitivity of results to control variables, we re-estimated regression models without controls for education and, separately, without controls for comorbid conditions. We also estimated a model including a dichotomous variable indicating if the subject had mild cognitive impairment to test for confounding by cognitive status.

Results

Among the 3260 responders, there were 800 persons with inadequate health literacy, 366 with marginal health literacy, and 2094 with adequate health literacy. There were large and statistically significant differences by health literacy level in terms of age, race/ethnicity, income, education, alcohol and tobacco consumption, and comorbid conditions (Table 1). Taken together, these comparisons show that controlling for underlying differences in individual characteristics is important when examining the impact of health literacy on medical costs or other endpoints.

The medical cost data are highly skewed; the mean value for total costs ($7346) exceeds the 75th percentile ($6635) (Table 2). Costs for inpatient and emergency room care are particularly skewed; the vast majority of respondents did not use these services during the year following enrollment into Prudential Healthcare.

At least 95% of respondents in each health literacy group used some type of medical care during the year following enrollment (Table 3). The proportion did not differ significantly by health literacy level. Nearly one-third of respondents in each group used inpatient care, and respondents with inadequate health literacy were more likely to use inpatient care than respondents with adequate health literacy. Respondents with marginal or inadequate health literacy were no more likely than respondents with adequate health literacy to use outpatient care or fill prescription drugs, despite having more chronic conditions. Mean total costs were $7246 among
respondents with inadequate health literacy, $8484 among respondents with marginal health literacy, and $9614 among respondents with adequate health literacy. Mean inpatient costs exhibited a similar dose-response relationship by health literacy level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Inadequate (n = 800)</th>
<th>Marginal (n = 366)</th>
<th>Adequate (n = 2,094)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>75.6 ± 5.6</td>
<td>74.1 ± 6.3</td>
<td>71.6 ± 7.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>462 (58)</td>
<td>197 (54)</td>
<td>1213 (58)</td>
<td>0.33</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td>White</td>
<td>200 (25)</td>
<td>46 (13)</td>
<td>138 (7)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>465 (58)</td>
<td>249 (68)</td>
<td>1752 (84)</td>
<td></td>
</tr>
<tr>
<td>Spanish speaking</td>
<td>103 (13)</td>
<td>60 (16)</td>
<td>137 (7)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>26 (3)</td>
<td>11 (3)</td>
<td>58 (3)</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>&lt;$10,000</td>
<td>247 (31)</td>
<td>92 (25)</td>
<td>251 (12)</td>
<td></td>
</tr>
<tr>
<td>$10,000–$15,000</td>
<td>183 (23)</td>
<td>81 (22)</td>
<td>435 (21)</td>
<td></td>
</tr>
<tr>
<td>$15,000–$25,000</td>
<td>144 (18)</td>
<td>81 (22)</td>
<td>605 (29)</td>
<td></td>
</tr>
<tr>
<td>$25,000–$35,000</td>
<td>33 (4)</td>
<td>32 (9)</td>
<td>217 (10)</td>
<td></td>
</tr>
<tr>
<td>$&gt;35,000</td>
<td>25 (3)</td>
<td>13 (4)</td>
<td>293 (14)</td>
<td></td>
</tr>
<tr>
<td>No response</td>
<td>168 (21)</td>
<td>67 (18)</td>
<td>293 (14)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>&lt;8 years</td>
<td>326 (41)</td>
<td>88 (24)</td>
<td>149 (7)</td>
<td></td>
</tr>
<tr>
<td>9–11 years</td>
<td>194 (24)</td>
<td>93 (25)</td>
<td>311 (15)</td>
<td></td>
</tr>
<tr>
<td>12 years</td>
<td>182 (23)</td>
<td>110 (30)</td>
<td>801 (38)</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>68 (9)</td>
<td>48 (13)</td>
<td>497 (24)</td>
<td></td>
</tr>
<tr>
<td>College degree</td>
<td>28 (4)</td>
<td>25 (7)</td>
<td>334 (16)</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Never</td>
<td>361 (45)</td>
<td>156 (43)</td>
<td>801 (38)</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>343 (43)</td>
<td>164 (45)</td>
<td>1030 (49)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>96 (12)</td>
<td>46 (13)</td>
<td>263 (13)</td>
<td></td>
</tr>
<tr>
<td>Drinking</td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td>None</td>
<td>599 (75)</td>
<td>235 (64)</td>
<td>1221 (58)</td>
<td></td>
</tr>
<tr>
<td>Light to moderate</td>
<td>186 (23)</td>
<td>121 (33)</td>
<td>784 (37)</td>
<td></td>
</tr>
<tr>
<td>Heavy</td>
<td>13 (2)</td>
<td>7 (2)</td>
<td>84 (4)</td>
<td></td>
</tr>
<tr>
<td>Comorbid conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart attack</td>
<td>117 (15)</td>
<td>67 (18)</td>
<td>268 (13)</td>
<td>0.01</td>
</tr>
<tr>
<td>Angina</td>
<td>66 (8)</td>
<td>44 (12)</td>
<td>175 (8)</td>
<td>0.06</td>
</tr>
<tr>
<td>Stroke</td>
<td>102 (13)</td>
<td>34 (9)</td>
<td>146 (7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>407 (51)</td>
<td>175 (48)</td>
<td>939 (45)</td>
<td>0.01</td>
</tr>
<tr>
<td>COPD*</td>
<td>113 (14)</td>
<td>58 (16)</td>
<td>371 (18)</td>
<td>0.06</td>
</tr>
<tr>
<td>Asthma</td>
<td>54 (7)</td>
<td>30 (8)</td>
<td>157 (7)</td>
<td>0.65</td>
</tr>
<tr>
<td>Cancer</td>
<td>36 (5)</td>
<td>25 (7)</td>
<td>130 (6)</td>
<td>0.15</td>
</tr>
<tr>
<td>Diabetes</td>
<td>156 (19)</td>
<td>59 (16)</td>
<td>277 (13)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Arthritis</td>
<td>463 (58)</td>
<td>214 (58)</td>
<td>1056 (50)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Depression</td>
<td>153 (19)</td>
<td>52 (14)</td>
<td>243 (12)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*Chronic obstructive pulmonary disease.

Table 2  Mean and median costs by site of service

<table>
<thead>
<tr>
<th>Site of care</th>
<th>Mean</th>
<th>Median (25th to 75th percentiles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>$7346</td>
<td>$1968 ($662 to $6635)</td>
</tr>
<tr>
<td>Inpatient</td>
<td>$3740</td>
<td>$0 ($0 to $0)</td>
</tr>
<tr>
<td>Outpatient</td>
<td>$1687</td>
<td>$1687 ($171 to $1845)</td>
</tr>
<tr>
<td>Emergency room</td>
<td>$136</td>
<td>$0 ($0 to $0)</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>$687</td>
<td>$381 ($72 to $948)</td>
</tr>
</tbody>
</table>

After adjusting for age, sex, race/ethnicity, income, education, alcohol and tobacco consumption, and comorbid conditions, the probability of using inpatient care was significantly higher among persons with inadequate health literacy compared to persons with adequate health literacy, as was the probability of using emergency room care (Table 4). When compared to those with adequate health literacy, emergency room costs were significantly greater in those with inadequate health literacy, whereas differences in total and inpatient costs were marginally significant. The difference in outpatient costs was negative but insignificant. In general, adjustment for covariates reduced estimated cost differences. Persons with marginal health literacy incurred significantly lower costs than those with adequate healthy literacy for outpatient care and significantly higher costs for
emergency room services, but the difference in total costs was insignificant.

Estimated differences in total costs between the inadequate and adequate groups from models that did not adjust for education ($1713; 95% CI: $103 to $3322; \( P < 0.03 \)) and comorbid conditions ($1328; 95% CI: $397 to $2997; \( P < 0.13 \)) were similar to the estimate from the baseline model ($1551; 95% CI: $166 to $3267; \( P < 0.08 \)), which did adjust for these factors. The estimated difference in total costs from a model including a control for cognitive impairment was $1797 (95% CI: $36 to $3558; \( P < 0.05 \)).

**Discussion**

Recently released reports by the Institute of Medicine and Agency for Healthcare Research and Quality conclude that the 90 million adults in our country with limited health literacy cannot fully benefit from medical care and the health care system.\(^4,27\) The reports note various interventions that hold promise for improving health literacy in various clinical settings. These include educational tools designed specifically for patients with low health literacy, health education programs for students, and efforts to facilitate communication between patients and providers. Although none of these has been implemented on a wide scale, the American Medical Association and the American College of Physicians Foundation have national efforts underway to raise awareness among clinicians about the problems faced by patients with limited health literacy.

Though most interventions to address low health literacy are relatively inexpensive and “low-tech,” it is important in today’s health care environment that researchers evaluate health literacy-related programs from an economic perspective. Estimates of cost differences by health literacy level

---

### Table 3  Unadjusted proportion using care and mean costs by health literacy level and site of service

<table>
<thead>
<tr>
<th>Site of care</th>
<th>Health literacy level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inadequate (n = 800)</td>
</tr>
<tr>
<td>Number (%) using service</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1992 (95)</td>
</tr>
<tr>
<td>Inpatient</td>
<td>728 (35)</td>
</tr>
<tr>
<td>Outpatient</td>
<td>1874 (90)</td>
</tr>
<tr>
<td>Emergency room</td>
<td>620 (30)</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>1772 (85)</td>
</tr>
<tr>
<td>Mean (± SD) costs</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>$9614 ± $22 536</td>
</tr>
<tr>
<td>Inpatient</td>
<td>$6817 ± $21 049</td>
</tr>
<tr>
<td>Outpatient</td>
<td>$1970 ± $3477</td>
</tr>
<tr>
<td>Emergency room</td>
<td>$189 ± $551</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>$638 ± $1267</td>
</tr>
</tbody>
</table>

Table 4  Adjusted differences in mean use and costs by health literacy level and site of service

<table>
<thead>
<tr>
<th>Site of care</th>
<th>Comparison</th>
<th>Marginal versus adequate health literacy</th>
<th>Inadequate versus adequate health literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference in probability of use (95% confidence interval)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.00 (−0.02 to 0.03)</td>
<td>0.00 (−0.02 to 0.02)</td>
<td></td>
</tr>
<tr>
<td>Inpatient</td>
<td>0.04 (−0.01 to 0.09)</td>
<td>0.05 (0.00 to 0.09)†</td>
<td></td>
</tr>
<tr>
<td>Outpatient</td>
<td>−0.01 (−0.04 to 0.02)</td>
<td>−0.02 (−0.05 to 0.01)</td>
<td></td>
</tr>
<tr>
<td>Emergency room</td>
<td>0.04 (−0.01 to 0.09)</td>
<td>0.05 (0.01 to 0.10)†</td>
<td></td>
</tr>
<tr>
<td>Pharmacy</td>
<td>−0.04 (−0.08 to 0.00)‡</td>
<td>−0.03 (−0.06 to 0.00)‡</td>
<td></td>
</tr>
<tr>
<td>Difference in costs (95% confidence interval)*</td>
<td>$596 (−$1437 to $2630)</td>
<td>$1551 (−$166 to $3267)‡</td>
<td>$1543 (−$89 to $3175)‡</td>
</tr>
<tr>
<td>Inpatient</td>
<td>$748 (−$1252 to $2748)</td>
<td>$1543 (−$89 to $3175)‡</td>
<td>$213 (−$481 to $55)</td>
</tr>
<tr>
<td>Outpatient</td>
<td>−$350 (−$679 to −$20)†</td>
<td>−$213 (−$481 to $55)</td>
<td>$108 ($62 to $154)†</td>
</tr>
<tr>
<td>Emergency room</td>
<td>$80 ($28 to $132)†</td>
<td>$108 ($62 to $154)†</td>
<td>$27 ($55 to $110)</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>$35 (−$62 to $132)</td>
<td>$27 ($55 to $110)</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for age, sex, race/ethnicity, income, education, alcohol and tobacco consumption, and comorbid conditions.
†\( P < 0.05. \)
‡\( P < 0.10. \)
can be of value as researchers study the financial impact of specific interventions.

Using data collected from beneficiaries in a Medicare managed care plan, we find that elderly enrollees with inadequate health literacy incur higher medical costs compared to enrollees with adequate health literacy. Total costs do not differ significantly between elderly enrollees with marginal versus adequate health literacy. The pattern of use and costs differences by site of service indicates that enrollees with low health literacy consume an inefficient mix of health care services.

Taken together, these results suggest that programs to address low health literacy among patients have the potential to reduce costs and improve quality. However, trials of actual interventions are necessary to demonstrate cost savings conclusively, and advocates should be cautious when touting the benefits of untested programs. For example, interventions that improve participants’ ability to schedule office visits and take medications as directed may increase spending on prescription drugs and outpatient services without generating offsetting reductions in spending on inpatient and emergency room care.

When interpreting these findings, it is important to recognize that medical spending data are highly skewed, implying that the variance is large relative to the mean. We have used statistical methods designed specifically for the analysis of medical spending data, but caution readers that, even with state-of-the-art methods, very large sample sizes are needed to precisely identify differences in spending between groups. The estimates presented here are based on a sample that is large in comparison to those from previous studies. Inclusion of health literacy measures in large population surveys, such as the National Health Interview Survey or the Medicare Current Beneficiary Survey, would allow researchers to estimate attributable costs more precisely in the future.

There are several other important limitations to consider. First, the sample and setting are not representative of the U.S. population and health system. All respondents were elderly and self-selected into a Medicare managed care plan. Comparisons of this sample with the Medicare Current Beneficiary Survey sample, which is designed to be nationally representative of Medicare enrollees, indicate that the respondents are younger, have lower incomes, incur higher medical costs, and are more likely to be hospitalized than the typical elderly Medicare beneficiary. Additionally, as a managed care plan, Prudential Healthcare employed features designed to encourage beneficiaries to use preventive services, and, in doing so, probably diminished the impact of health literacy on use and costs compared to a fee-for-service environment.

Second, we cannot assess the degree to which associations between health literacy and medical costs are causal. Health literacy decreases with age in the sample, beyond what one would expect based on differences in education levels. A previous analysis of these data rejected the hypothesis that differences in health literacy across age groups are attributable to the onset of chronic conditions; yet, the differences in our population are so large that we must consider the possibility that poor health leads to deteriorations in cognitive functioning and, consequently, health literacy. If so, estimates overstate the impact of low health literacy on medical costs and utilization.

Encouragingly, the pattern of spending and use differences across the four types of care is consistent with a causal interpretation of results. Health literacy is defined partly by the ability to understand physician visit scheduling and dosing timetables. As would be expected, use of and spending on outpatient care and prescription drugs do not vary by health literacy level or are slightly lower for persons with low health literacy. If cost differences between groups were driven solely by unobserved differences in health status, then we would expect to find that costs for persons with low health literacy are uniformly higher across sites of care. Instead, we find that use and costs are higher mainly in those areas—emergency room and inpatient care—often associated with treatment of complications and advanced cases of disease.

We have elected to present the more conservative set of results based on models that include controls for education and comorbid conditions, but, as stated above, these may lead to estimates that are biased downwards. As noted by Berkman et al., “One limitation of the knowledge base to date is lack of appropriate specification for analytic models when variables being considered as potential confounders actually mediate the effect of reading ability on important health outcomes.” With further refinements to the concept of “health literacy,” it may be possible to narrow down the list of potential confounders for future cost studies.

References


CLINICAL RESEARCH STUDY

Community-acquired pneumonia as the initial manifestation of serious underlying diseases

Miquel Falguera, MD,a Mariela Martín, MD,a Agustín Ruiz-González, MD,a Ricard Pifarré, MD,a Mercè García, MDb

aDepartment of Internal Medicine and
bDepartment of Microbiology, Hospital Universitari Arnau de Vilanova.

Several studies have identified a variety of comorbid illnesses as important risk factors for community-acquired pneumonia.1,2 Diagnoses such as active malignancy, chronic obstructive pulmonary disease, chronic heart failure, diabetes mellitus or human immunodeficiency virus (HIV) infection are present in many patients with pneumonia.3 Moreover, significant associations between the etiology of community-acquired pneumonia and prior comorbid conditions have also been suggested, particularly for patients with alcoholism, hematologic malignancies, and chronic pulmonary or hepatic diseases.3-5

However, community-acquired pneumonia could also constitute the initial manifestation of these comorbid conditions. Thus, pneumonia or other infectious diseases have been recognized as potential indicators of having underlying hematologic malignancies.6 Similarly, a close relationship between pneumococcal disease and HIV infection has been established, particularly in young and previously healthy individuals.7 However, the magnitude of this association...
and factors that may be predictive of having an unknown underlying disease remain undetermined.

The objectives of the present study were to look for the association of community-acquired pneumonia with serious underlying diseases and to identify potential clinical or microbiologic parameters associated with a higher risk of finding a new coexisting condition.

**Subjects and methods**

**Study design and setting**

For a 5-year period (January 1998 to December 2002), all patients with community-acquired pneumonia diagnosed at the emergency room of the Arnau de Vilanova Hospital in Lleida (Catalonia, Spain) were included in a prospective study on epidemiologic, clinical, microbiologic, and outcome parameters. This study was examined and approved by the scientific and ethics committees of the institution.

**Study and follow-up of patients**

On enrollment (first visit), all patients underwent a complete clinical history and physical examination, a chest radiography, an intensive microbiologic evaluation as described in a previous paper, and chemistry and hematology tests, including measurements of blood glucose, urea nitrogen, creatinine, protein, sodium, potassium, chloride, alanine aminotransferase and aspartate aminotransferase levels, hematocrit, hemoglobin, mean corpuscular volume, white cell and differential count, platelet count, prothrombin time and oxygen saturation by pulse oximetry, or arterial blood gas testing. The Pneumonia Patient Outcome Research Team (PORT) prediction rule was employed to stratify patients in prognostic groups.

A particular effort was made to detect previously known relevant coexisting conditions. We considered the following illnesses to be relevant: neoplastic disease (any noncutaneous malignancy, active at the time of presentation or diagnosed within the last 12-month period), chronic liver disease (cirrhosis or chronic active hepatitis), chronic heart failure, cerebrovascular disease with severe neurologic sequelae at the time of presentation, diabetes mellitus, chronic renal disease, HIV infection, active tuberculosis, chronic obstructive pulmonary disease and severe autoimmune connective tissue disease or systemic vasculitides requiring immunosuppressive therapy. These comorbid conditions were classified as previously known under any of the following circumstances: 1) the existence of this underlying disease was notified at inclusion by the patient or family members; 2) current treatment (including pharmacologic or dietary therapies) was followed by the patient; 3) a review of the patient’s past hospital records revealed or suggested the presence of this prior coexisting condition.

Patients who required hospitalization were clinically monitored during admission; for outpatients, a second clinical visit was planned, within 48 to 72 hours after enrollment. Additional tests were performed only when an unsuccessful outcome was observed or when different or additional illnesses besides pneumonia were suspected, as described below.

For all patients, a final visit was made between 4 and 6 weeks later that included a clinical history and physical examination, chest radiography, a second sample for serologic studies, and a repeated battery of identical chemistry and hematology tests.

The initial visit and enrollment of patients in the study was handled by the physicians in the Emergency Department while the investigators were responsible for follow-up during hospitalization and care of outpatients. Any patients who failed to attend their appointments were contacted by phone.

**Evaluation and diagnostic criteria for new comorbid conditions**

When the clinical evaluation, radiologic findings, or biologic results suggested the presence of previously unrecognized comorbid conditions, appropriate diagnostic tests were ordered. Thus, radiologic, endoscopic, and/or histologic procedures were performed to evaluate a possible neoplasm; laboratory tests were used to study or to confirm diabetes mellitus, chronic hepatic or renal diseases, connective tissue diseases, or systemic vasculitides and HIV infection; microbiologic studies were used to detect active tuberculosis; computed tomography or magnetic resonance imaging procedures were performed to detect a cerebrovascular disease; echocardiograms were ordered for potential chronic heart failure; and a spirometric test was performed for patients with possible chronic obstructive pulmonary disease. In addition, HIV infection detection tests were ordered for all patients with bacteremic pneumonia or active tuberculosis. According to results provided by these diagnostic procedures, we defined:

- Neoplastic disease as the presence of a noncutaneous localized or disseminated neoplasm, histologically demonstrated
- Chronic liver disease as the presence of clinical features and/or laboratory findings of severe chronic liver disease
- Chronic heart failure as echocardiographic evidence of depressed left ventricular function (< 45%)
- Cerebrovascular disease as the presence of specific neurologic clinical features and abnormalities in neuroimaging procedures
- Diabetes mellitus as a fasting plasma glucose concentration > 125 mg/dL, or after ingestion > 200 mg/dL on at least 2 separate occasions
- Chronic renal disease as a reduced glomerular filtration rate (<35%) that was maintained in the second laboratory control
• HIV infection as positive ELISA and Western blot tests
• Active tuberculosis as a positive Ziehl-Neelsen stain and/or Löwenstein-Jensen culture of sputum or other biologic sample
• Chronic obstructive pulmonary disease as an irreversible reduced forced expiratory volume (FEV1) (<80%)
• Severe autoimmune connective tissue disease or systemic vasculitides as the presence of conventional clinical and laboratory criteria (for vasculitides, histologic confirmation was also required)

Classification of microorganisms

To perform comparative analyses, etiologies detected in microbiologic tests were distributed into 2 groups as follows: 1) bacterial etiology, including pneumonia caused by conventional bacteria such as Streptococcus pneumoniae, Haemophilus influenzae, Staphylococcus aureus, Streplococcus viridans and other less common gram-positive cocci, Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa and other less common gram-negative bacilli; and 2) atypical etiology, including pneumonia caused by atypical microorganisms such as Mycoplasma pneumoniae, Chlamydia pneumoniae, Chlamydia psittaci, Coxiella burnetii, Legionella pneumophila, and viruses.

Statistical analysis

Univariate statistical analysis was performed to compare groups, using, for qualitative variables, chi-square or Fisher exact tests, and for quantitative variables, t or Mann-Whitney tests. The level of significance was set at P < 0.05.

Results

We evaluated a total of 726 episodes of community-acquired pneumonia, based on the clinical and radiologic findings at entry. We excluded 66 patients for the following reasons: an erroneous diagnosis of pneumonia (59), recent hospitalization (4), and failure to obtain informed consent. In consequence, 660 patients constituted the final study group.

Five-hundred and twenty-nine (80%) were hospitalized and 131 were managed as outpatients. During follow-up, 58 (9%) patients died, all in the hospitalized group. For hospitalized patients, the final clinical visit, 4-6 weeks after inclusion (deceased patients excluded), was performed in 443 (94%) cases. For outpatients, the second clinical visit, 48-72 hours after inclusion, was performed in 125 (95%) patients; and the final visit, 4-6 weeks later, was performed in 104 (79%) patients. Microbiologically, the most common agent was S. pneumoniae; a polymicrobial infection was detected in 59 (9%) episodes; and for 204 (31%) patients, the etiology could not be determined.

Clinical records of patients on enrollment recognized 1 or more underlying diseases in 298 (45%) patients, with the following distribution: chronic obstructive pulmonary disease in 106 (16%) patients, diabetes mellitus in 92 (14%), chronic heart failure in 55 (8%), HIV infection in 41 (6%), malignancy in 29 (4%), chronic liver disease in 29 (4%), chronic renal disease in 12 (2%), and cerebrovascular disease in 6 (1%). Other underlying illnesses were found in 8 cases. Many patients had more than one comorbid condition.

In addition, 45 unrecognized comorbid conditions were found during the period of study in 41 (6%) patients. We established a diagnosis of diabetes mellitus in 14 patients with a mean age of 69 ± 15 years. In these patients, 9 (64%) of whom were male, bacteremia was detected in 3 (21%) and a bacterial etiology was obtained in 9 (64%). Twelve patients with a mean age of 60 ± 14 years suffered from a malignancy (lung cancer, 6; colon cancer, 1; gastric cancer, 1; primary liver cancer, 1; multiple myeloma, 1; lymphoma, 1; and disseminated cancer of unknown origin, 1). In this group of patients, 9 (75%) of whom were male; 2 of 10 had bacteremia and 7 (58%) had a bacterial etiology. A diagnosis of chronic obstructive pulmonary disease was established in 8 patients with a mean age of 62 ± 9 years. In this group of patients, 7 (87%) of whom were male; bacteremia was observed in 1 (14%), and 4 (50%) had a bacterial etiology. An HIV infection was diagnosed in 5 patients with a mean age of 34 ± 9 years. In this group of patients, 4 (80%) of whom were male, bacteremia was present in 2 (40%), and 4 (80%) had a bacterial infection. Finally, chronic heart failure (2 patients), active tuberculosis (2), liver cirrhosis (1), and acute stroke (1) diagnoses were established for the remaining patients.

Figure 1 shows a flowchart documenting the type of comorbid condition and the moment when these conditions were detected. As this figure demonstrates, the diagnosis of diabetes mellitus was usually established in the initial evaluation, based on the first laboratory screening. By contrast, the diagnosis of the remaining comorbid conditions was usually established during follow-up. In the final visit, neoplasm constituted the most frequently diagnosed underlying disease.

Table 1 compares the characteristics of both groups (patients with and without new comorbid conditions). There were no differences between samples in relation to baseline epidemiological findings and clinical variables; however, the diagnosis of new comorbid conditions was significantly associated with a bacterial etiology, bacteremia, and hospitalization.

Discussion

This study shows that community-acquired pneumonia can be the first manifestation of previously unknown underlying diseases. Thus, we found 1 or more new serious comorbid

380 The American Journal of Medicine, Vol 118, No 4, April 2005
conditions in more than 6% of our patients; diabetes mel-
itus, neoplasm (particularly lung and hematologic malign-
ancies), chronic obstructive pulmonary disease, and HIV 
infection were the most frequently recognized coexisting 
illnesses. The detection of unknown comorbid conditions 
was particularly frequent in patients who had a bacterial 
etiology, had positive blood cultures, or were hospitalized.

It is hardly surprising that we found more underlying 
diseases among patients with bacterial etiology. Several 
investigators had already found a correlation between 
certain underlying diseases and specific bacterial micro-
organisms. Ruiz et al. showed that pneumococcal pneu-
monia was more likely in patients with pulmonary or hepatic comorbid illnesses, and gram-negative enteric bacilli were associated with the presence of chronic pul-
monary diseases.3 Similarly, Porath et al observed a 
significant association between comorbid conditions, par-
ticularly chronic obstructive pulmonary disease, and sev-
eral bacterial etiologies.10 Furthermore, studies on the 
etiology of pneumonia in patients with HIV infection or 
hematologic malignancies have detected a low incidence 
of atypical microorganisms.4,11

In addition, it is reasonable to associate comorbid con-
ditions and bacteremia. We know that the rate of pneumo-
coccal bacteremia is increased in patients with HIV infec-
tion and may be more than 100 times higher than that found 
in age-matched populations;12 a higher incidence of sepsis 
accounts in patients with hematologic malignancies;4 and an 
increased rate of bacteremia has been suggested among 
diabetic patients with community-acquired pneumonia.13

Hospitalization was also statistically related to a 
higher probability of the detection of underlying diseases. 
Several explanations for this relationship can be pro-
posed. Thus, we may speculate that, in some cases, the 
decision to hospitalize was reached because, at the time 
of admission, some clinical or biologic parameters were 
identified that led physicians to suspect the existence of a 
comorbid condition. Alternatively, hospitalization may 
have allowed physicians to detect the clinical or biologic 
signs associated with a comorbid condition. Finally, pa-
tients with more severe infections who are more likely to 
be hospitalized may have a greater risk of having un-
known coexisting illnesses. Undoubtedly, we cannot con-
clude that hospitalization is mandatory for all patients 
with community-acquired pneumonia; however, the de-
tection of a comorbid condition could be a potential 
benefit provided by more intensive medical care.

Figure 1 Follow-up of patients and detection of new comorbid conditions in 660 episodes of community-acquired pneumonia (some patients had more than 1 comorbid condition).
Many of these underlying illnesses may be clinically silent during the initial period, precisely when therapies could have greater efficacy in delaying the progression of underlying disease and avoiding the development of complications. Therefore, the recognition of a comorbid condition in patients with community-acquired pneumonia may provide important benefits for the patient and even for the community.

Our study has some important limitations. First, it may be hypothesized that a careful screening would demonstrate the presence of unknown comorbid conditions in a proportion of hospitalized patients, regardless of the causative reason for the admission, or even among a sample of asymptomatic individuals. However, such high incidences as we found, particularly for neoplasm and HIV infection, appear to be improbable in other circumstances. Second, the causative agent of community-acquired pneumonia remains unknown in a high percentage of patients, even when extensive diagnostic testing by conventional methods is performed. In consequence, an etiologic diagnosis of bacterial pneumonia frequently cannot be established, although newly developed diagnostic tools, such as antigen detection tests or especially the polymerase chain reaction–based method, show considerable promise for the future. Finally, we should recognize that patients with new underlying diseases constituted, as we describe in the results, a heterogeneous sample. Thus, HIV-infected patients were certainly younger and had a higher rate of bacteremia, when compared with patients with diabetes mellitus or chronic obstructive pulmonary disease.

In summary, community-acquired pneumonia constitutes a good opportunity to identify unknown but clinically serious comorbid conditions. Greater risk of detecting unknown underlying diseases was found in patients with bacterial etiology and those with bacteremia. Patients with community-acquired pneumonia should have an extensive diagnostic workup targeted to detect the presence of an unknown comorbid condition.

Acknowledgment

We thank Roman Pallares, MD for his critical review of the manuscript.

References


CLINICAL RESEARCH STUDY

Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia

Drahomir Aujesky, MD, MSc,a,d Thomas E. Auble, PhD,b Donald M. Yealy, MD,b Roslyn A. Stone, PhD,c,d D. Scott Obrosky, MSc,a,d Thomas P. Meehan, MD, MPH,e,f Louis G. Graff, MD,e,g,h Jonathan M. Fine, MD,i Michael J. Fine, MD, MSca,d

aDivision of General Internal Medicine, Department of Medicine,
bDepartment of Emergency Medicine,
cDepartment of Biostatistics, Graduate School of Public Health, University of Pittsburgh;
dVA Center for Health Equity Research and Promotion and VA Pittsburgh Healthcare System;
eQualidigm;
fYale University School of Medicine;
gDepartment of Traumatology and Emergency Medicine, University of Connecticut School of Medicine;
hDepartment of Emergency Medicine, New Britain General Hospital; and
iSection of Pulmonary and Critical Care Medicine, Norwalk Hospital.

ABSTRACT

PURPOSE: We assessed the performance of 3 validated prognostic rules in predicting 30-day mortality in community-acquired pneumonia: the 20 variable Pneumonia Severity Index and the easier to calculate CURB (confusion, urea nitrogen, respiratory rate, blood pressure) and CURB-65 severity scores.

SUBJECTS AND METHODS: We prospectively followed 3181 patients with community-acquired pneumonia from 32 hospital emergency departments (January–December 2001) and assessed mortality 30 days after initial presentation. Patients were stratified into Pneumonia Severity Index risk classes (I–V) and CURB (0–4) and CURB-65 (0–5) risk strata. We compared the discriminatory power (area under the receiver operating characteristic curve) of these rules to predict mortality and their accuracy based on sensitivity, specificity, predictive values, and likelihood ratios.

RESULTS: The Pneumonia Severity Index (risk classes I–III) classified a greater proportion of patients as low risk (68% [2152/3181]) than either a CURB score <1 (51% [1635/3181]) or a CURB-65 score <2 (61% [1952/3181]). Low-risk patients identified based on the Pneumonia Severity Index had a slightly lower mortality (1.4% [31/2152]) than patients classified as low-risk based on the CURB (1.7% [28/1635]) or the CURB-65 (1.7% [33/1952]). The area under the receiver operating characteristic curve was higher for the Pneumonia Severity Index (0.81) than for either the CURB (0.73) or CURB-65 (0.76) scores (P < 0.001, for each pairwise comparison). At comparable cut-points, the Pneumonia Severity Index had a higher sensitivity and a somewhat higher negative predictive value for mortality than either CURB score.

CONCLUSIONS: The more complex Pneumonia Severity Index has a higher discriminatory power for short-term mortality, defines a greater proportion of patients at low risk, and is slightly more accurate in identifying patients at low risk than either CURB score.

© 2005 Elsevier Inc. All rights reserved.

KEYWORDS: Clinical prediction rule; Community-acquired pneumonia

This study was supported by grant RO1 HS10049-03 from the Agency for Healthcare Research and Quality, Rockville, Maryland. Dr. Aujesky was supported in part by the Novartis Research Foundation and Dr. MJ Fine was supported in part by a K24 career development award from the National Institute of Allergy and Infectious Diseases.

Requests for reprints should be addressed to Drahomir Aujesky, MD, MSc, Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System, Building #28 Suite 1A129, University Drive C, Pittsburgh, PA 15240.

E-mail address: aujesky@swissonline.ch

0002-9343/$ -see front matter © 2005 Elsevier Inc. All rights reserved.
Physicians often use models of prognosis to quantify severity of illness and guide the initial site of treatment decision for patients with community-acquired pneumonia (designated pneumonia). The Pneumonia Severity Index was derived to identify patients with pneumonia who are at low risk for short-term mortality and potential candidates for outpatient care. Based on 20 clinical variables (Figure 1A), patients are assigned into 5 risk classes with an increasing risk of 30-day mortality. Patients in risk classes I–III are defined as low risk for mortality and are considered potential candidates for outpatient treatment. Three prior interventional studies have demonstrated that the clinical application of the Pneumonia Severity Index safely decreases the proportion of hospital admissions for low-risk patients. Utilization of the Pneumonia Severity Index safely decreases the proportion of hospital admissions for low-risk patients. Utilization of the Pneumonia Severity Index safely decreases the proportion of hospital admissions for low-risk patients. Utilization of the Pneumonia Severity Index safely decreases the proportion of hospital admissions for low-risk patients. Utilization of the Pneumonia Severity Index safely decreases the proportion of hospital admissions for low-risk patients.

Figure 1 Prediction rules for 30-day mortality for patients with community-acquired pneumonia. (A) Risk Class Assignment based on the Pneumonia Severity Index. (B) Risk Stratum Assignment based on the CURB and CURB-65 severity scores. In all analyses, missing values were assumed to be normal. CURB = confusion, urea nitrogen, respiratory rate, and blood pressure.
nia Severity Index for the initial risk assessment has been widely endorsed by organizations such as the Infectious Diseases Society of America and others.1,2,5,8

Based upon prediction rules originally developed to identify patients with severe pneumonia,13,14 Lim et al described a prognostic model that stratifies patients into 5 strata of increasing risk of mortality.15 In Lim’s CURB score (confusion, urea nitrogen, respiratory rate, blood pressure), a total point score ranging from 0–4 is calculated based upon these 4 prognostic variables (Figure 1B). In a recent single site validation study, patients with a CURB score <1 had a low 30-day mortality that was comparable to that found in Pneumonia Severity Index risk classes I–III, leading to the conclusion that the relatively simple CURB score may replace the more complex Pneumonia Severity Index for identifying low-risk patients with pneumonia.16 However, this study excluded 30% of the study patients in comparisons of the CURB and the Pneumonia Severity Index,16 resulting in a potential selection bias.

A modified version of the CURB score, which added age ≥65 years as a fifth prognostic variable and stratifies patients into 6 strata of increasing risk of mortality, was recently derived and internally validated (Figure 1B).17 The 2004 update of the British Thoracic Society pneumonia guidelines states that patients with a CURB-65 score <2 may be suitable for outpatient treatment.1,18

Despite such guideline recommendations, the CURB-65 score has never been externally validated. Nor has either CURB score been directly compared with the more complex Pneumonia Severity Index in an independent patient sample of pneumonia patients. Therefore, we sought to compare the performance of the Pneumonia Severity Index and the 2 versions of the CURB severity scores in predicting 30-day mortality in a large cohort of patients with pneumonia. Our a priori hypotheses were that the Pneumonia Severity Index would more accurately identify patients who are at low risk of 30-day mortality, whereas the two CURB scores would be superior at identifying pneumonia patients at high risk of mortality.

Methods

Study sites and patients

This study was conducted between January and December 2001 as part of a clinical trial to assess the effectiveness of 3 practice guideline implementation strategies (low, moderate, and high intensity) for pneumonia in 32 hospital emergency departments in Pennsylvania (n = 16) and Connecticut (n = 16).19 The institutional review boards of all participating study sites approved all study procedures. The study design and characteristics of participating emergency departments were described previously.19

Potential study subjects were identified in all 32 participating emergency departments. Eligible patients were 18 years of age or older with a clinical diagnosis of pneumonia and a new radiographic pulmonary infiltrate. Patients were excluded if they were considered to have hospital-acquired pneumonia, immunosuppression or co-morbid conditions that distinguished them diagnostically or therapeutically from pneumonia, or psychosocial problems incompatible with outpatient treatment, enrollment, or follow-up. Patients who were pregnant, previously enrolled, or enrolled in a competing research protocol were also excluded.19 Patients with all inclusion criteria and no exclusion criteria documented were approached for informed consent to participate in the study.

Patient baseline assessment

For all enrolled patients, baseline demographic information (age, sex, nursing home residence) and clinical data were collected by medical record review by trained research nurses and recorded on a standard data collection instrument. The physical examination findings (pulse, respiratory rate, systolic and diastolic blood pressure, temperature, and mental status), comorbid conditions (neoplastic disease, liver disease, congestive heart failure, cerebrovascular disease, renal disease), and laboratory and radiographic results (arterial pH, blood urea nitrogen, sodium, glucose, hematocrit, level of arterial oxygenation, and pleural effusion) that comprise the Pneumonia Severity Index and the CURB severity scores were abstracted as part of the medical record review. For all physical examination, laboratory, and radiographic findings, the first available measurement after the time of presentation in the emergency department was recorded.

Patient outcomes assessment

All-cause mortality was assessed for all patients 30 days after initial presentation by patient interview and medical record review. Interviews were performed via telephone and were administered by trained interviewers. For this project, we excluded those patients for whom mortality could not be ascertained.

Clinical prognostic models

Based on patient demographics and baseline clinical data obtained by chart review, we determined the presence of the 20 Pneumonia Severity Index risk factors (Figure 1A) and all individual risk factors comprising the CURB and the CURB-65 severity scores (Figure 1B). For the 2 CURB scores, “presence of confusion” is defined using an Abbreviated Mental Test Score ≤8 or new disorientation to person, place, or time.17 Due to the absence of these variables in the present study, “altered mental status” was used as a proxy measure for confusion. For any of the variables constituting the Pneumonia Severity Index or CURB scores, missing values were assumed to be normal. This strategy is widely used in the clinical application of prognostic prediction rules and reflects the methods used in the original derivation and validation of the Pneumonia Severity Index.9,20

Based upon a 2-step algorithm (Figure 1A), all enrolled patients were classified into Pneumonia Severity Index risk classes I–V. Likewise, all patients were assigned to
5 risk strata based on the 4 prognostic factors in the original CURB score and 6 risk strata based on the 5 prognostic factors in the more recent CURB-65 score (Figure 1B).

**Statistical analysis**

We described 30-day mortality and mortality likelihood ratios by Pneumonia Severity Index risk class, as well as by CURB and CURB-65 risk strata. Likelihood ratios represent the degree to which the classification into a given risk class or stratum modifies the pretest probability of 30-day mortality.\(^{21}\)

Based on commonly accepted definitions of low-risk patients (Pneumonia Severity Index risk classes I–III; CURB scores <1; and CURB-65 scores <2), pairwise comparisons of the proportions of patients classified as low risk by the 3 rules were made using McNemar’s test.\(^{22}\)

To determine the accuracy of each rule to predict 30-day mortality, we estimated sensitivity, specificity, and positive and negative predictive values for each possible cut-point to define high risk. We assessed the discriminatory power of each rule by calculating the area under each receiver operating characteristic curve, performing pairwise comparisons of the areas under the 3 curves.\(^{23}\)

In a secondary comparison, we also tested whether a 2-step approach as used in the more complex Pneumonia Severity Index would improve the predictive performance of the CURB severity scores relative to the Pneumonia Severity Index. Patients in Pneumonia Severity Index risk class I were assigned a CURB and CURB-65 score of 0 (Step 1). The CURB and the CURB-65 scores were calculated in the remaining patients by adding 1 to the original score values (Step 2), converting the CURB score to a 5-point, 6-strata scale and the CURB-65 score to a 6-point, 7-strata scale. Overall 30-day mortality rates, mortality likelihood ratios, and the area under the receiver operating characteristic curves were calculated for each new CURB score.

We calculated an exact binomial 95% confidence interval (CI) for each test performance measure. For all analyses, a 2-sided \(P\) value <0.05 was considered to be statistically significant.

**Results**

**Study sample**

Of the 4506 identified patients with pneumonia who met eligibility criteria, 3615 (80%) were initially enrolled. Nonenrolled individuals tended to be older (mean age 74 vs. 63 years) and more likely to be resident in nursing homes (36% vs. 4.9%) than enrolled patients.

**Comparison of mortality**

The Pneumonia Severity Index classified a significantly greater proportion of patients as low risk (68%...
than the CURB (51% [1635/3181]) and the CURB-65 (61% [1952/3181]) scores (P<0.001 for each pairwise comparison) (Table 2). Low-risk patients identified based on the Pneumonia Severity Index had a slightly lower aggregate 30-day mortality of 1.4% (31/2152) compared with 1.7% (28/1635) for low-risk patients identified using the CURB score and 1.7% (33/1952) for low-risk patients using the CURB-65 score. High-risk patients based on the Pneumonia Severity Index had a somewhat higher mortality of 11.1% (114/1029) compared with 7.6% (117/1546) for high-risk patients based on the CURB score and 9.1% (112/1229) for high-risk patients based on the CURB-65 score. Although patients in the highest CURB and CURB-65 risk strata showed a higher mortality than those in the highest Pneumonia Severity Index risk class V (43% vs. 24%), these 2 CURB risk strata each included only 7 patients (0.2%) compared with 200 (6%) patients in Pneumonia Severity Index risk class V.

Comparison of predictive accuracy and discriminatory power

At every given threshold, the Pneumonia Severity Index had a higher sensitivity and a lower specificity than the two CURB scores (Table 3). The negative predictive values were high (>95%) across all thresholds for all prediction rules; the positive predictive values were low. The Pneumonia Severity Index had a greater discriminatory power to predict 30-day mortality than either CURB score (Figure 2).

Secondary comparison

If CURB scores <2 and CURB-65 scores <3 were used to define low-risk in the revised, 2-step severity CURB scores, 30-day mortality of low-risk patients was 1.6% (28/1711) for CURB and 1.7% (33/1955) for CURB-65 (Table 4). The corresponding areas under the receiver operating characteristic curves were 0.75 (95% CI: 0.71–0.78) for the revised CURB and 0.77 (95% CI: 0.74–0.80) for the revised CURB-65 scores, both significantly lower than for the Pneumonia Severity Index (0.81) (P<0.001, for each pairwise comparison).

Discussion

Our comparison shows that the Pneumonia Severity Index has a higher discriminatory power for predicting 30-day mortality than either CURB severity score. Low-risk patients identified using the Pneumonia Severity Index have a slightly lower mortality and a slightly higher negative predictive value for death than low-risk patients identified using either CURB severity score. Even when the CURB severity scores were converted into more complex, 2-step prediction rules using the Pneumonia Severity Index to identify the lowest risk strata, these results did not change markedly. Although the absolute difference in 30-day mortality based on the 3 prediction rules is small and of uncertain clinical relevance, the Pneumonia Severity Index classifies a significantly larger proportion of patients with
pneumonia as low risk than either CURB score. Because over 4 million cases of pneumonia occur in the United States per year, utilization of the Pneumonia Severity Index would identify an additional 650,000 low-risk patients compared with the CURB and an additional 250,000 low-risk patients compared with the CURB-65, many of whom would be potential candidates for outpatient treatment. Given an average cost of care of $7500 for inpatients and $350 for outpatients with pneumonia, the greater number of low-risk patients identified by the Pneumonia Severity Index is likely to result in substantial cost savings.

Although the Pneumonia Severity Index accurately identifies patients with pneumonia who are at low risk of short-term mortality and potential candidates for outpatient treatment, decisions pertaining to the initial site of treatment should include other clinical factors such as the presence of arterial hypoxemia, other coexisting illnesses that warrant hospital admission, and psychosocial problems that preclude outpatient care.

Table 3 Measures of performance in predicting 30-day mortality by prediction rule

<table>
<thead>
<tr>
<th>Cut-points by prediction rule</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia Severity Index risk classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥II</td>
<td>99 (95–100)</td>
<td>23 (21–24)</td>
<td>6 (5–7)</td>
<td>99.7 (99–100)</td>
</tr>
<tr>
<td>≥III</td>
<td>97 (92–99)</td>
<td>48 (46–50)</td>
<td>8 (7–10)</td>
<td>99.7 (99–100)</td>
</tr>
<tr>
<td>≥IV*</td>
<td>79 (71–85)</td>
<td>70 (68–72)</td>
<td>11 (9–13)</td>
<td>99 (98–99)</td>
</tr>
<tr>
<td>V</td>
<td>32 (25–41)</td>
<td>95 (94–96)</td>
<td>24 (18–30)</td>
<td>97 (96–97)</td>
</tr>
<tr>
<td>CURB scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1*</td>
<td>81 (73–87)</td>
<td>53 (51–55)</td>
<td>8 (6–9)</td>
<td>98 (98–99)</td>
</tr>
<tr>
<td>≥2</td>
<td>47 (39–55)</td>
<td>85 (84–87)</td>
<td>13 (11–17)</td>
<td>97 (96–98)</td>
</tr>
<tr>
<td>≥3</td>
<td>10 (6–17)</td>
<td>98 (97–98)</td>
<td>19 (11–29)</td>
<td>96 (95–97)</td>
</tr>
<tr>
<td>4</td>
<td>2 (0.4–6)</td>
<td>99.9 (99.7–100)</td>
<td>43 (10–82)</td>
<td>96 (95–96)</td>
</tr>
<tr>
<td>CURB-65 scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥1</td>
<td>96 (91–99)</td>
<td>34 (33–36)</td>
<td>7 (6–8)</td>
<td>99.4 (99–100)</td>
</tr>
<tr>
<td>≥2*</td>
<td>77 (70–84)</td>
<td>63 (62–65)</td>
<td>9 (7–11)</td>
<td>98 (98–99)</td>
</tr>
<tr>
<td>≥3</td>
<td>45 (37–53)</td>
<td>87 (86–88)</td>
<td>14 (11–18)</td>
<td>97 (96–98)</td>
</tr>
<tr>
<td>≥4</td>
<td>10 (5–16)</td>
<td>98 (98–99)</td>
<td>20 (11–31)</td>
<td>96 (95–97)</td>
</tr>
<tr>
<td>5</td>
<td>2 (0.4–6)</td>
<td>99.9 (99.7–100)</td>
<td>43 (10–82)</td>
<td>96 (95–96)</td>
</tr>
</tbody>
</table>

CURB is an acronym for confusion, urea nitrogen, respiratory rate, and blood pressure.
*Cut-points that designate commonly accepted thresholds to define low vs. high-risk patients.

Figure 2 Receiver operating characteristic curves for 30-day mortality for the Pneumonia Severity Index and the 2 CURB severity scores. The areas under the receiver operating characteristic curves were 0.81 (95% confidence interval [CI]: 0.78–0.84) for the Pneumonia Severity Index, 0.73 (95% CI: 0.68–0.76) for the CURB, and 0.76 (95% CI: 0.73–0.80) for the CURB-65 severity score (P <0.001, for each pairwise comparison). CURB = confusion, urea nitrogen, respiratory rate, and blood pressure.
Our study also demonstrates that the recently developed CURB-65 severity score is a better tool for predicting mortality than the CURB score: the CURB-65 score showed a higher overall discriminatory power for mortality than the original CURB score. This finding underscores the importance of older age as a predictor of short-term mortality in pneumonia. However, in contrast to the study by Lim et al, in which a CURB-65 score had a sensitivity of 93% in the derivation and 100% in the validation cohort, the performance of the CURB-65 score was less impressive in our study: a CURB-65 severity score had a sensitivity of 77% in our cohort. Indeed, the CURB-65 severity score, derived and internally validated in inpatients (with a 30-day mortality of 9%) appears to perform less well in an independent patient sample that includes both inpatients and outpatients.

A potential advantage of the CURB severity scores over the Pneumonia Severity Index is their simplicity and potential for greater ease of use in the clinical setting. The CURB and the CURB-65 scores consist of 4 and 5 clinical variables, including only 1 laboratory variable (blood urea nitrogen). In contrast, the Pneumonia Severity Index is comprised of 20 predictor variables, including 7 laboratory and radiographic variables. On the other hand, patients in Pneumonia Severity Index risk class I, who represent a substantial proportion of our study cohort (22%), can be identified solely on the basis of history and physical examination findings, without the need for any laboratory tests such as blood urea nitrogen. This makes the Pneumonia Severity Index useful in clinical settings where laboratory tests may not be available. The use of pocket cards, electronic handheld devices, or Internet support systems could further facilitate the application of the Pneumonia Severity Index in clinical practice.

There are several potential limitations to our study that should be acknowledged. First, our study sample may not reflect the full prognostic spectrum of patients with pneumonia because enrolled patients were younger than nonenrolled patients and less likely to be admitted from nursing homes. Thus, we cannot exclude the possibility that the 3 prediction rules would have performed differently in more severely ill patients. However, the 80% enrollment rate that we achieved is laudable for a multicenter clinical trial, and the enrolled patients reflect a broad spectrum of patient demographic and clinical characteristics. Second, in our analysis we assumed missing values for any of the predictor variables constituting the Pneumonia Severity Index and the CURB scores to be normal, a strategy previously validated for the Pneumonia Severity Index but not for the CURB scores. Because information about blood urea nitrogen was not available in 21% of patients, we cannot exclude the possibility that disease severity may have been underestimated by the CURB scores that we calculated for this subgroup with missing values. Third, this study was performed within a clinical trial to compare 3 guideline implementation strategies of incremental intensity, and was not originally designed to compare different prediction rules for mortality. Thus, the CURB variable “presence of confusion,” defined as an Abbreviated Mental Test Score ≤8 or new disorientation to person, place, or time, was not available and “altered mental status” had to be used as a proxy measure. Because 30-day mortality did not vary across study intervention arms, it is unlikely that this study design affected our study results.

<table>
<thead>
<tr>
<th>Prediction rule risk strata</th>
<th>Patients (n = 3181)</th>
<th>Deaths</th>
<th>Likelihood ratio for mortality</th>
<th>Parameter (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CURB score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>686 (22)</td>
<td>2 (0.3)</td>
<td>0.06 (0.03–0.2)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1025 (32)</td>
<td>26 (2.5)</td>
<td>0.54 (0.4–0.8)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>962 (30)</td>
<td>49 (5.1)</td>
<td>1.1 (0.9–1.4)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>428 (13)</td>
<td>53 (12)</td>
<td>3.0 (2.3–3.7)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>73 (2)</td>
<td>12 (16)</td>
<td>4.1 (2.3–7.5)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7 (0.2)</td>
<td>3 (43)</td>
<td>16 (3.5–70)</td>
<td></td>
</tr>
<tr>
<td>CURB-65 score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>686 (22)</td>
<td>2 (0.3)</td>
<td>0.06 (0.03–0.2)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>441 (14)</td>
<td>4 (0.9)</td>
<td>0.19 (0.07–0.5)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>828 (26)</td>
<td>27 (3.3)</td>
<td>0.71 (0.5–1.0)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>772 (24)</td>
<td>47 (6.1)</td>
<td>1.4 (1.1–1.7)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>383 (12)</td>
<td>51 (13)</td>
<td>3.2 (2.5–4.1)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>64 (2)</td>
<td>11 (17)</td>
<td>4.3 (2.3–8.1)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7 (0.2)</td>
<td>3 (43)</td>
<td>16 (3.5–70)</td>
<td></td>
</tr>
</tbody>
</table>

CURB is an acronym for confusion, urea nitrogen, respiratory rate, and blood pressure.

* Patients in Pneumonia Severity Index risk class I were assigned a CURB and CURB-65 score of 0 (Step 1). In patients with a Pneumonia Severity Index of II–V (Step 2), the CURB and the CURB-65 scores were calculated by adding 1 to the original score values, converting the CURB score to a 5-point, 6-strata scale and the CURB-65 score to a 6-point, 7-strata scale.

† The denominator used was the number of patients per risk stratum. Mortality rates could not be statistically compared across prediction rules because the denominators differed.
In conclusion, the Pneumonia Severity Index has a significantly higher discriminatory power for predicting 30-day mortality than the 2 CURB severity scores in this large cohort of inpatients and outpatients with pneumonia. The more recently developed CURB-65 severity score recommended by the British Thoracic Society is superior to the CURB score in predicting mortality. The Pneumonia Severity Index is both more efficient and slightly more accurate in identifying low-risk patients with pneumonia who are potential candidates for outpatient care and at least as accurate as the CURB severity scores in identifying high-risk patients with this illness.

Acknowledgments

The authors would like to acknowledge the important contributions of the individuals listed below for serving the following key roles on the project:

Study Site Investigators. These individuals were responsible for facilitating the initiation and conduct of the trial at the local hospital level:

Connecticut: Michael L. Carius, MD, Thomas Ceddia, MD, Ian Cummings, MD, Robert Femia, MD, Bernard J. Ferguson IV, MD, William Gemmell, MD, Robert J. Grant, MD, Steven T. Holland, MD, Thomas J. Koobatian, MD, Lawrence P. Levine, MD, Robert D. Powers, MD, Mark R. Prete, MD, Eric Salk, MD, John Scarfo, MD, John Schriver, MD, Jay Walshon, MD, Michael J. Werdmann, MD, and C. Steven Wolf, MD

Pennsylvania: Thomas P. Campbell, MD, Theodore R. Delbridge, MD, Christopher Dooley, MD, Roderick B. Gromes, MD, F. Richard Heath, MD, William G. Kristan, MD, Rani K. Kumar, MD, Bruce A. MacLeod, MD, Robert J. Maha, MD, Jeffrey Moldovan, DO, James E. Nicholas, MD, Edwin H. Page Jr., MD, Joel D. Rosenbloom, DO, Jerald A. Solot, DO, and Robert R. Whipkey, MD.

Research Staff. These individuals were responsible for data entry, database development and maintenance, and statistical analyses: Teri Foreman, Deljo Gannon, Sarah Geis, Shih-Yieh Ho, PhD, Nan Klein, Jennifer Paterline, BS, and Linda Quinn.

Study Research Nurses. These individuals were responsible for serving as liaisons between the research office and the 32 study sites, guiding the emergency departments in screening and enrolling patients, retrieving forms from the sites, and assisting in patient interviews and follow-up: Barbara Albro, RN, MBA, Mary Jansen, RN, BSN, Charlotte Jerney, RN, MSN, Susan Lamson, RN, BA, and Donna Simpson, CRNP.

References


CLINICAL RESEARCH STUDY

Prolonging the return visit interval in primary care

Gordon Schectman, MDa Gary Barnas, MD,a Prakash Laud, PhD,b Laura Cantwell, MD,a Monica Horton, MS, MPP,c Edwin J. Zarling, MDd

aMilwaukee Veterans Affairs Medical Center and bDepartment of Biostatistics, Medical College of Wisconsin; cSocioeconomic Policy Development Department, American Medical Association; and dVeterans Integrated Service Network #12 Loyola University Medical Center.

ABSTRACT

PURPOSE: Extending the scheduled return visit interval has been suggested as one means to improve clinic access to the provider. However, prolonging the return visit interval may affect quality of care if prevention measures and chronic disease management receive less attention as clinic visits become less frequent. The purpose of this study was to determine whether a comprehensive education program could encourage providers to lengthen their return visit interval without compromising performance on key quality indicators.

SUBJECTS AND METHODS: This was a prospective cohort study monitoring scheduling and performance data of primary care providers at the Milwaukee Veterans Affairs Medical Center. Following collection of baseline data (January through June 1999), providers were encouraged to lengthen the return visit interval while increasing reliance on nurses and other clinic staff for interim management of chronic disease. Provider-specific feedback of return visit interval and performance data was utilized to motivate behavioral change. Scheduling and clinical data were abstracted from random medical record audits performed at baseline and from July through December in the years 2000 and 2001.

RESULTS: Compared with the baseline period, the percent of patients scheduled ≥6 months was significantly increased among staff providers and medicine residents at 2 years (Staff providers: 31% vs. 62%, P < 0.001; Medicine residents: 22 vs. 44%, P < 0.001). Colorectal screening, pneumonia immunizations, and achievement of therapeutic goals for diabetes, hypertension, and lipid disorders significantly improved at 2 years compared with baseline measurements.

CONCLUSIONS: Educational interventions can successfully retrain providers to extend the return visit interval and reduce the scheduling of routine and perhaps unnecessary appointments. This can be accomplished without compromising important performance monitors for diabetes, lipid disorders, hypertension, and prevention.

© 2005 Elsevier Inc. All rights reserved.
Training providers to lengthen the scheduled return visit interval is an important change concept recommended to increase provider access. The return visit interval is the time interval determined by the provider to be appropriate between the current visit and the next visit and is usually determined by the clinician near the end of each clinic visit. Lengthening the return visit interval has the potential to reduce the number of routine and possibly unnecessary clinic visits, thereby enhancing provider availability to address urgent/emergent patient issues.

Studies investigating the return visit interval in clinical practice suggest several important determinants, including patient-specific factors (such as how ill the patient appeared to the clinician) and clinician management factors (such as if a decision to order tests or change patient management occurred at the visit). However, provider practice style also explained much of the variance of the scheduled return visit interval, a finding that persisted even after adjustments for important patient characteristics. Provider practice style may reflect scheduling habits acquired from previous training independent of the patient’s medical needs. For example, providers have often been trained to schedule their patients every 3 or 4 months routinely, regardless of disease severity. There is also a common belief among clinicians that frequent provider-patient visits are needed to achieve therapeutic goals in many common chronic illnesses. Thus, patient-provider visits were often encouraged to facilitate closer chronic disease monitoring. Although prolonging the return visit interval may help promote patient access to health care, data suggesting that clinicians can be retrained to accomplish this without compromising important chronic disease outcomes are not available.

In primary care clinics at the Milwaukee Veterans Affairs Medical Center, we sought to influence clinician choice of the return visit interval through provider education; clinicians were encouraged to adjust the return visit interval according to medical necessity rather than upon a fixed time interval. We wished to test the hypothesis that the routine 3- or 4-month return visit interval commonly found among our providers (and encouraged by their previous training) could be modified through relatively simple interventions without compromising clinical performance. Internal medicine residents were included in this program, as appropriate return visit scheduling was considered to be an important element of their clinical training.

Methods

Staff providers at the Milwaukee Veterans Affairs Medical Center and associated community-based outreach clinics who had at least 2 clinics weekly were included. Resident physicians were also included, although each had only one primary care clinic weekly for the 3-year duration of their training. Approximately 15% of primary care patients seen at the Medical Center were managed by resident physicians. Beginning in January 1999, 120 chart audits from staff and 30 from residents, selected at random, were reviewed per 6-month observation period. Clinic policy dictated that all patients were to have a return appointment scheduled within 2 weeks of the clinic visit. Chart audits were conducted 2 weeks after a clinic visit. Patients without a return visit interval scheduled by this time were excluded. Data were abstracted from medical data available on the electronic record through the VA’s Veteran Health Information System Technology Architecture (VISTA) database. The return visit interval was defined as the number of months between the date of the clinic visit and the next primary care appointment date entered into the scheduling package for the subsequent return visit.

Patients were identified as having diabetes if either a past glycated hemoglobin was greater than 7% or the patient was currently taking 1 or more glucose-lowering medications. Glycemic goals were considered achieved if the patient’s most recent glycohemoglobin level performed within the past year was less than 8%.

Patients were considered to have hyperlipidemia if either a lipid profile was above goal levels or prescriptions for lipid-lowering medications were present. Lipid goals were derived from National Cholesterol Education Program recommendations and included minor modifications consistent with HEDIS (Health Plan Employer and Data Information Set). Lipid goals were considered achieved if the patient met the following criteria:

- 0 or 1 coronary risk factors: LDL cholesterol ≤4.91 mm/L (190 mg/dL) and triglycerides ≤4.52 mm/L (400 mg/dL)
- Two or more coronary risk factors: LDL cholesterol ≤4.14 mm/L (160 mg/dL) and triglycerides ≤4.5 mm/L (400 mg/dL)
- Presence of coronary heart or peripheral vascular disease: LDL ≤3.36 mm/L (130 mg/dL), TG ≤3.39 mm/L (300 mg/dL), and HDL ≥0.78 mm/L (30 mg/dL)

Patients were identified as having hypertension by a review of diagnostic codes from patient problem lists or discharge diagnoses from either clinic or hospital within the past 5 years. Although less accurate than pharmacy records, diagnostic codes were used to identify hypertension as antihypertensive medications often have multiple indications. A review of diagnostic coding accuracy routinely demonstrated a sensitivity and specificity for the actual presence of hypertension of >90%. Hypertension goals were achieved if the mean blood pressure of the last 2 blood pressure determinations was less than 140/90 mm Hg. Although stricter clinical goals have been recommended for dyslipidemia, diabetes and hypertension management, we selected more liberal performance goals to allow for the potential for poor compliance, known variability of drug efficacy, and gaps in research supporting intensification of drug therapy...
when LDL cholesterol, glycated hemoglobin or blood pressure are only minimally abnormal.\textsuperscript{12}

Colorectal cancer screening was required for all individuals \(\geq 50\) years old and considered accomplished if fecal occult blood screening was performed within the past year, colonoscopy was performed within the past 10 years, or flexible sigmoidoscopy was performed in the past 5 years. Pneumonia immunization was required every 10 years for patients \(\geq 55\) years old or at particularly high pneumonia risk.\textsuperscript{13}

**Intervention to lengthen the return visit interval**

A performance improvement initiative was implemented in June 1999 to encourage primary care providers to prolong the return visit interval whenever medically feasible. This concept was promoted as a means to reduce congestion in busy primary care clinics and decrease workload of providers and staff. To facilitate lengthening of the return visit interval without compromising quality of care, increased reliance on the allied health professionals working within the primary care team was recommended. Each primary care team included registered nurses, licensed practical nurses, and medical assistants; clinical pharmacists and dietitians were also available on a more limited basis. Additional support personnel were not added to the primary care teams during this initiative; rather, ancillary support staff adopted efficiencies designed to optimize utilization consistent with advanced access principles.\textsuperscript{4}

To encourage return visit interval lengthening, a series of meetings targeting important stakeholders in the primary care clinic were initiated. These included:

- **Primary care providers for two 1-hour meetings:** Clinicians were encouraged to arrange return visits when a face-to-face patient visit would be particularly useful, such as to check wound healing or to evaluate resolution of an acute illness; the scheduling of “routine” visits every 3 or 4 months independent of the patient’s medical status was discouraged. Providers were instructed to plan a specific agenda for the next visit; if only laboratory or blood pressure monitoring were required, the provider was encouraged to consider arranging these outside of a provider visit. During this initiative, quarterly reports were disseminated to all clinicians disclosing their own return visit interval for the past quarter, ranked anonymously with all other primary care providers.

- **Registered nurses and clinical pharmacists for two 1-hour meetings:** National VA clinical practice guidelines for diabetes, hypertension and lipid disorder were reviewed and effective collaboration with their providers was encouraged to facilitate guideline implementation.

- **Licensed practical nurses and medical assistants for four 1-hour meetings:** techniques to provide interim surveillance between extended return visits were taught, including routine blood pressure, laboratory monitoring, and medication compliance checks. Implementing preventive measures according to algorithms for cancer screening (such as fecal occult blood testing) and immunizations (such as for pneumonia and influenza) was also encouraged.

- **Patients:** A 1-page brochure was distributed to patients informing them that their provider may wish to lengthen their return visit interval if appropriate, possibly substituting greater participation in their care by ancillary team members. Patients were encouraged to call the clinic to be seen sooner by their provider for any medical need or concern.

Data were collected prospectively beginning in January 1999. For purposes of analysis, this was divided into a preintervention baseline period (January to June 1999) and three postintervention periods (July to December 1999, July to December 2000, and July to December 2001). To assess the impact of the prolongation of return visit interval on the utilization of primary care clinic staff, emergency department, and specialty clinics, we retrospectively reviewed a simple random sample of patient visits occurring during the baseline (January to June 1999) and final periods of data collection (July to December 2001).

The return visit interval was grouped into 2 dichotomous categories, \(< 4\) months and \(\geq 6\) months. To assess the impact of multiple clinical variables on the return visit interval, a logistic regression model using simultaneous entry of all independent variables was used. Analyses from patients sampled within providers were adjusted for clustering by provider (Tables 1 and 2). The outcome “return visit interval” was analyzed using a normal linear model with a provider-specific random effect. All other outcomes being binary, clustering for these was accounted for by using generalized estimating equations.\textsuperscript{14} Analyses were carried out using the Statistical Analysis System (SAS; SAS Institute Inc.; Cary, NC).

**Results**

During the baseline period, 3525 patients having scheduled appointments with 24 staff providers and 22 medicine residents between January and June 1999 were initially reviewed; 3418 were scheduled for a return visit (97%). The mean baseline return visit interval for all patients was 4.4 ± 2.4 months. Patients had a mean age of 64 ± 13 years and a mean of 2.2 ± 1.0 cardiac risk factors.

During the baseline period (January–June 1999) for staff physicians, 57% of patients were scheduled for return appointments at \(>4\) months and 31% were scheduled for return appointments at \(>6\) months (Table 1). Following the baseline period, strategies to prolong the return visit interval, including team-building, provider and team education, and provider-specific performance reports concerning the return visit interval were implemented (see Methods). The percent of patient return visits scheduled at 4 and 6 months increased steadily through December 2001 (Table 1). This
lengthening of the return visit interval observed subsequent to the baseline period was highly significant. Increases in the return visit interval for resident physicians paralleled those observed for staff through December 2000, although the return visit interval for residents was shorter for each period of data collection. Data collection through December 2001 showed continued lengthening of the return visit interval for staff but not for resident providers.

Individual staff return visit interval for baseline and 2 follow-up periods are shown in Figure 1. During the base-

### Table 1  Effect of team training and provider feedback on the return visit interval*

<table>
<thead>
<tr>
<th>Time period</th>
<th>Baseline January to June 1999</th>
<th>Post-intervention July to December 1999</th>
<th>Post-intervention July to December 2000</th>
<th>Post-intervention July to December 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Patients</td>
<td>2899</td>
<td>2950</td>
<td>2968</td>
<td>3758</td>
</tr>
<tr>
<td>Patient age in years (mean ± SD)</td>
<td>64 ± 13</td>
<td>64 ± 13</td>
<td>65 ± 13</td>
<td>65 ± 13</td>
</tr>
<tr>
<td>Return visit interval ≥4 months (%)</td>
<td>1648 (57)</td>
<td>1917 (65)†</td>
<td>2207 (74)†</td>
<td>3077 (82)†</td>
</tr>
<tr>
<td>Return visit interval ≥6 months (%)</td>
<td>889 (31)</td>
<td>1143 (39)‡</td>
<td>1556 (52)‡</td>
<td>2325 (62)‡</td>
</tr>
<tr>
<td>Mean return visit interval (months)</td>
<td>4.3 ± 2.4</td>
<td>4.7 ± 2.4†</td>
<td>5.4 ± 2.6†</td>
<td>5.7 ± 2.4†</td>
</tr>
<tr>
<td><strong>Residents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># Patients</td>
<td>519</td>
<td>643</td>
<td>762</td>
<td>726</td>
</tr>
<tr>
<td>Patient age (years)</td>
<td>63 ± 13</td>
<td>65 ± 14</td>
<td>64 ± 13</td>
<td>67 ± 13</td>
</tr>
<tr>
<td>Return visit interval ≥4 months (%)</td>
<td>259 (50)</td>
<td>374 (58)‡</td>
<td>503 (66)§</td>
<td>461 (64)§</td>
</tr>
<tr>
<td>Return visit interval ≥6 months (%)</td>
<td>113 (22)</td>
<td>222 (35)‡</td>
<td>330 (43)†</td>
<td>318 (44)†</td>
</tr>
<tr>
<td>Mean return visit interval (months)</td>
<td>3.9 ± 2.1</td>
<td>4.4 ± 2.1†</td>
<td>4.7 ± 2.2†</td>
<td>4.7 ± 2.2†</td>
</tr>
</tbody>
</table>

*24 staff and 22 resident providers monitored during baseline period, expanding to 25 staff and 26 resident providers during first follow-up period and second follow-up period, and 31 staff and 26 resident providers during final follow-up period.

†P < 0.001.
‡P < 0.01.
§P < 0.05 compared to baseline period (January to June 1999).

### Table 2  Performance measures during implementation of intervention to lengthen the return visit interval

<table>
<thead>
<tr>
<th>Time period</th>
<th>Baseline January to June 1999</th>
<th>Post-intervention July to December 1999</th>
<th>Post-intervention July to December 2000</th>
<th>Post-intervention July to December 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percent (n/N)</strong>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Staff</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycated hemoglobin &lt;8%</td>
<td>54 (359/667)</td>
<td>53 (386/729)</td>
<td>59 (420/717)</td>
<td>70 (597/849)‡</td>
</tr>
<tr>
<td>Achieved lipid goals</td>
<td>57 (701/1220)</td>
<td>60 (767/1289)</td>
<td>64 (941/1481)</td>
<td>71 (1326/1872)‡</td>
</tr>
<tr>
<td>Achieved blood pressure goals†</td>
<td>43 (533/1247)</td>
<td>45 (910/2006)</td>
<td>49 (1105/2235)</td>
<td>57 (1313/2311)†</td>
</tr>
<tr>
<td>Patients screened for colorectal cancer</td>
<td>31 (756/2457)</td>
<td>39 (1013/2587)§</td>
<td>56 (1510/2687)†</td>
<td>70 (2323/3331)†</td>
</tr>
<tr>
<td>Patients receiving Pneumovax</td>
<td>33 (636/1907)</td>
<td>43 (848/160)§</td>
<td>85 (1746/2065)§</td>
<td>93 (2306/2486)‡</td>
</tr>
<tr>
<td><strong>Residents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycated hemoglobin &lt;8%</td>
<td>51 (61/120)</td>
<td>63 (93/147)</td>
<td>66 (122/185)</td>
<td>69 (122/177)§</td>
</tr>
<tr>
<td>Achieved lipid goals</td>
<td>50 (80/161)</td>
<td>57 (144/251)</td>
<td>64 (216/338)</td>
<td>66 (236/356)§</td>
</tr>
<tr>
<td>Achieved blood pressure goals†</td>
<td>43 (132/306)</td>
<td>48 (217/483)</td>
<td>53 (328/627)</td>
<td>53 (253/476)†</td>
</tr>
<tr>
<td>Patients screened for colorectal cancer</td>
<td>39 (174/444)</td>
<td>44 (239/547)</td>
<td>55 (367/673)§</td>
<td>67 (438/646)‡</td>
</tr>
<tr>
<td>Patients receiving Pneumovax</td>
<td>31 (103/337)</td>
<td>47 (203/435)§</td>
<td>69 (340/491)§</td>
<td>86 (444/519)‡</td>
</tr>
</tbody>
</table>

* n/N—numerator n represents those patients achieving the defined goal (for example the number of patients achieving glycated hemoglobin levels <8% or the number of patients screened for colorectal cancer) and denominator N represents those eligible to be included in the sample (for example, the number of patients with diabetes or the number of patients eligible for colorectal cancer screening).
† During the baseline period, use of electronic data entry for blood pressure measurements had just been initiated and only 54% of blood pressures were entered electronically into the medical record and available for review. This number increased to >98% by the end of the study.
‡P < 0.001.
§P < 0.01.
¶P < 0.05 compared to baseline period (January to June 1999).
line period, marked heterogeneity in provider practice was observed in the return visit interval, as the proportion of patients per provider with a return visit interval \(\geq 6\) months varied from 3\% to 70\%. Following feedback and training, the proportion of patients scheduled \(\geq 6\) months increased in 22 of 24 providers. The largest increases were noted in providers who initially had the shortest return visit interval (correlation coefficient between initial return visit interval and magnitude of the change in the return visit interval from baseline through December 2001: \(-0.70\); 95\% confidence interval \(0.39–0.87\), \(P<0.001\)). However, even with this improvement, marked provider heterogeneity persisted through the last 6 months of data collection (July–December 2001), as the proportion of patients with a return visit interval scheduled \(\geq 6\) months varied from 17\% to 84\%.

Despite prolongation of the return visit interval in primary care, performance measures monitored subsequent to the baseline period showed improvement (Table 2). Compared with baseline, a greater proportion of patients achieved therapeutic goals for diabetes, hypertension, and lipid disorder management. In addition, more patients received appropriate colorectal screening and pneumovax immunization. Patients managed by both internal medicine trainees and primary care staff providers showed similar improvements.

The impact of the intervention to prolong the return visit interval on clinical resources is shown in Table 3. Compared with the baseline period, provider visits to the primary care team were reduced by 27\%, visits to specialty care were reduced by 14\%, and visits to the emergency department for urgent care were unchanged. In contrast, patient visits to the primary care registered nurse increased by 100\% and telephone calls to the primary care nurse increased by 150\%. Patient visits to the licensed practical nurse and to the clinical pharmacists in the follow-up period were not significantly different compared with the baseline period. Telephone contacts for licensed practical nurses and clinical pharmacists were not recorded.

### Discussion

In this study, we assessed the impact of an intervention designed to affect clinician behavior regarding return visit interval scheduling and also addressed the important question of whether lengthening the return visit interval would be associated with deleterious effects on selected hypertension, diabetes, lipid, and prevention outcomes. We found that an intervention to prolong the return visit interval utilizing provider and ancillary staff education

![Figure 1](image)

**Figure 1** Effect of feedback and counseling on the proportion of patients scheduled for a return visit at 6 or more months. Black bars: baseline period 1/99–6/99; White bars: postintervention period 7/00–12/00; Gray bars: postintervention period 7/01–12/01. Includes only staff providers present during both baseline and at least the first postintervention period.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Baseline January to June 1999</th>
<th>Post-intervention July to December 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients sampled</td>
<td>918</td>
<td>980</td>
</tr>
<tr>
<td>Visits to primary care registered nurse</td>
<td>0.26 ± 0.64</td>
<td>0.56 ± 1.28</td>
</tr>
<tr>
<td>Telephone contacts with primary care registered nurse</td>
<td>0.24 ± 0.83</td>
<td>0.86 ± 2.24</td>
</tr>
<tr>
<td>Visits to primary care licensed practical nurse</td>
<td>0.41 ± 1.52</td>
<td>0.29 ± 1.31</td>
</tr>
<tr>
<td>Visits to primary care clinical pharmacist</td>
<td>0.54 ± 2.46</td>
<td>0.64 ± 2.74</td>
</tr>
<tr>
<td>Visits to primary care provider</td>
<td>2.91 ± 2.30</td>
<td>2.11 ± 1.89</td>
</tr>
<tr>
<td>Visits to emergency department provider</td>
<td>0.60 ± 1.55</td>
<td>0.69 ± 1.71</td>
</tr>
<tr>
<td>Visits to specialty clinic provider</td>
<td>5.52 ± 7.27</td>
<td>4.72 ± 6.76</td>
</tr>
</tbody>
</table>

Data analysis performed by independent samples t-test.
and performance feedback was associated with substantial lengthening of the return visit interval. No deterioration in diabetes, lipid disorders, hypertension outcomes were found, and prevention measures showed substantial improvement. These changes were observed in both staff and resident physician practices, and persisted during a 2.5-year period of observation. These data suggest that primary care providers can be educated to modify their practice patterns to extend the return visit interval without compromising quality of care.

Prolonging many routine primary care appointments beyond 6 months raises the concern that patients, perceiving incorrectly that access to their provider is reduced, may shift their care to either an urgent or specialty care setting. In our study, this concern did not materialize. There was no significant increase in emergency department utilization, and specialty care visits actually decreased. Although non-VA health care utilization was not monitored, most veterans would have faced much higher charges obtaining their care from this sector, and thus major shifting of care in this direction is unlikely. However, within the primary clinic, there appeared a marked shift in resource utilization from the primary care provider to the primary care registered nurse. The ability to maintain or improve important performance measures despite fewer contacts with the provider was likely due to the expanded role of the registered nurse in the clinic. Sharing of the clinical workload within a primary care team of health professionals is an important concept of advanced access clinical systems.

Our quality improvement program to prolong the return visit interval was multifaceted and was not designed to evaluate the relative importance of each specific component. However, it did appear that individual provider education and provider-specific feedback were particularly important for providers who initially had the shortest return visit intervals. These providers showed the greatest return visit interval prolongation, perhaps because their initial return visit interval scheduling was more heavily dependent upon their training and least adapted to individual patient characteristics such as disease burden or severity. On the other hand, providers who had a longer return visit interval at baseline appeared to benefit more from interdisciplinary training of the primary care team. These providers were more likely to require greater involvement of the primary care team ancillary staff to further extend the return visit interval.

Certain limitations of this study should be mentioned. Our study was conducted at a large Veterans Affairs Medical Center and therefore results may not be generalizable to other institutions. This may be particularly true for institutions with a fee-for-service, rather than a capitated reimbursement model. VA providers therefore have an incentive to efficiently manage the health care of their patients independent of the number of clinic visits. This study was not a randomized controlled trial and therefore, observed associations may not be causal. For example, patient variables not collected may affect patient case-mix (such as total number of comorbidities, functional status, and number of medications) and thereby influence the provider return visit interval. Patient case-mix was likely to be similar between providers, as new patients were not selectively distributed to providers based upon specific patient or provider characteristics. An additional concern is that the before-after design utilized in this study cannot adjust for time-dependent factors. Thus, if all providers had recently been assigned new patients at the start of the observation period, increasing familiarity with their medical problems may have resulted in longer return visit intervals and improved performance measures independent of the training intervention. However, this is unlikely because staff physicians at program initiation had relatively mature patient panels with little turnover during the observation period. However, a randomized controlled trial would be necessary to address these potential confounders, to more completely assess the impact of specific interventions to increase the return visit interval, and to more fully evaluate the relationship between the return visit interval and clinical outcomes. The serious methodological difficulties in the design of such studies have been previously discussed.

In conclusion, our data suggest that provider decision-making regarding the return visit interval can be significantly modified by feedback and education without compromising performance. Because many ambulatory care settings in the United States are characterized by full provider clinic schedules and limited patient access to health care, lengthening the return visit interval may be one important approach to create provider access and improve health care. Increasing reliance on ancillary health professionals, and particularly registered nurses, working closely together within a primary care setting, allowed clinicians to reduce the frequency of patient visits while maintaining high standards of care.

Acknowledgment

The authors acknowledge the many hours of hard work and dedication of Ms. Leslie Voigt, who coordinated and supervised data entry for this study.

References

Adjustment for do-not-resuscitate orders reverses the apparent in-hospital mortality advantage for minorities

Naomi Bardach, MD, a Shoujun Zhao, MD, PhD, a Steven Pantilat, MD, b S. Claiborne Johnston, MD, PhD c

a Department of Neurology
b Division of General Internal Medicine, Department of Medicine, and
c Department of Neurology and Epidemiology, University of California.

ABSTRACT
PURPOSE: The use of do-not-resuscitate (DNR) orders may differ by sex or ethnicity, and DNR status may be associated with outcomes for hospitalized patients. Thus, we sought to determine whether differences in rates of DNR by sex and ethnicity influenced differences in mortality.

SUBJECTS AND METHODS: We included all patients admitted to nonfederal California hospitals in 1999 with stroke, congestive heart failure, pneumonia, chronic obstructive pulmonary disease, chronic renal failure, angina, or diabetes mellitus. Rates of physician orders for DNR written within 24 hours of hospital admission and in-hospital mortality were compared between sexes and ethnicities after adjustment for age, admission source and diagnosis, payment type, and comorbidity scores in multi-variable logistic regression models.

RESULTS: Of 327,890 patients included, 25,196 (7.7%) had DNR orders. In adjusted models, women were more likely to have DNR orders than men (odds ratio [OR] 1.19; 95% confidence interval 1.16–1.23; P <0.001) and non-Hispanic whites were more likely to have DNR orders than other ethnicities (OR 1.75; 1.69–1.82; P <0.001). Overall, 13,549 (4.1%) patients died in the hospital. Risk of death was greater in those with a DNR order (OR 7.0; 6.7–7.3; P <0.001). Non-Hispanic whites appeared to have a greater risk of in-hospital death in adjusted models (OR 1.09; 1.04–1.12; P <0.001) when DNR status was ignored; however, the risk of death appeared to be lower in non-Hispanic whites in the complete model with DNR included (OR 0.94; 0.90–0.99; P = 0.01). A survival advantage for women was also more apparent after including DNR status in the adjusted model.

CONCLUSIONS: Women and non-Hispanic whites are more likely to have DNR orders. DNR status affected the measurement of sex-ethnic differences in mortality risk.

© 2005 Elsevier Inc. All rights reserved.
Few studies have evaluated the predictors and impact of DNR orders. Several studies have found that utilization of DNR orders varies by sex and ethnicity, but others have not demonstrated an association. Two studies demonstrated that patients with DNR orders are more likely to die in the hospital, even after adjustment for characteristics such as age and severity of illness. DNR appears to influence outcomes because care is less aggressive and because it is a measure of disease severity not captured with other variables.

Ethnicity is an important predictor of death after diagnosis or admission for several diseases. Differences in access to care and socioeconomic status have been hypothesized as possible explanations, but the cause of ethnic differences in mortality is largely unknown. Interestingly, 2 recent studies found lower in-hospital mortality in African-Americans than in non-Hispanic whites, but the influence of DNR was not evaluated. Similarly, without consideration of DNR, rates of in-hospital mortality are often found to be lower for women. We sought to test this hypothesis; thus, we asked, is DNR a confounder in the association of sex and ethnicity with in-hospital mortality?

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. DNR (%)</th>
<th>Associated odds of DNR</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>Univariate</td>
<td>Adjusted*</td>
</tr>
<tr>
<td>&lt;65</td>
<td>127 095</td>
<td>2.24</td>
<td>1.00</td>
</tr>
<tr>
<td>≥65</td>
<td>200 795</td>
<td>11.13</td>
<td>5.47</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>153 881</td>
<td>6.51</td>
<td>1.00</td>
</tr>
<tr>
<td>Female</td>
<td>174 009</td>
<td>8.72</td>
<td>1.37</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>210 075</td>
<td>9.64</td>
<td>1.00</td>
</tr>
<tr>
<td>Asian</td>
<td>22 029</td>
<td>4.84</td>
<td>0.48</td>
</tr>
<tr>
<td>Hispanic</td>
<td>38 513</td>
<td>4.15</td>
<td>0.41</td>
</tr>
<tr>
<td>African-American</td>
<td>33 373</td>
<td>3.95</td>
<td>0.39</td>
</tr>
<tr>
<td>Admission Source:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-nursing home</td>
<td>322 739</td>
<td>7.51</td>
<td>1.00</td>
</tr>
<tr>
<td>Nursing home</td>
<td>5151</td>
<td>18.35</td>
<td>2.77</td>
</tr>
<tr>
<td>Payment source</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>186 761</td>
<td>10.84</td>
<td>1.00</td>
</tr>
<tr>
<td>Medi-Cal/indigent</td>
<td>60 110</td>
<td>2.91</td>
<td>0.25</td>
</tr>
<tr>
<td>HMO</td>
<td>48 444</td>
<td>3.99</td>
<td>0.34</td>
</tr>
<tr>
<td>Non-HMO private</td>
<td>29 007</td>
<td>4.01</td>
<td>0.34</td>
</tr>
<tr>
<td>Comorbidity score‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>132 850</td>
<td>4.79</td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td>54 390</td>
<td>8.72</td>
<td>1.90</td>
</tr>
<tr>
<td>2</td>
<td>61 960</td>
<td>7.88</td>
<td>1.70</td>
</tr>
<tr>
<td>3 or 4</td>
<td>52 772</td>
<td>10.54</td>
<td>2.35</td>
</tr>
<tr>
<td>≥5</td>
<td>25 918</td>
<td>14.06</td>
<td>3.25</td>
</tr>
<tr>
<td>Admission diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angina</td>
<td>54 953</td>
<td>1.91</td>
<td>1.00</td>
</tr>
<tr>
<td>Diabetes</td>
<td>17 077</td>
<td>3.12</td>
<td>1.65</td>
</tr>
<tr>
<td>COPD§</td>
<td>34 934</td>
<td>6.97</td>
<td>3.84</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>53 113</td>
<td>8.83</td>
<td>4.97</td>
</tr>
<tr>
<td>Chronic renal failure</td>
<td>12 174</td>
<td>9.20</td>
<td>5.20</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>79 207</td>
<td>9.34</td>
<td>5.28</td>
</tr>
<tr>
<td>Stroke</td>
<td>76 432</td>
<td>10.43</td>
<td>5.97</td>
</tr>
</tbody>
</table>

*Adjusted by age, sex, ethnicity, admission source, payment source, comorbidity scores, and admission diagnosis, as categorized above. DNR = do-not-resuscitate orders, OR = odds ratio, CI = confidence interval. †P values calculated using logistic regression. ‡Comorbidity scores were based on the Charlson comorbidity index adapted for administrative databases. §Chronic obstructive pulmonary disease.
Methods

Subjects

The Human Subjects Committee of the University of California, San Francisco, approved this study. California’s Office of State Health Planning and Development (OSHPD) maintains a statewide database of inpatient discharge abstracts for all nonfederal hospitals. Based on a list of common diagnoses included in prior studies of ethnicity and mortality, we selected a cohort of patients discharged between January 1 and December 31, 1999, with 1 of 7 primary diagnoses: stroke (International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM] code 430–431, 432.0, 432.1, 433–434.9, 436–437.9), angina (diagnosis-related group [DRG] 140, 143), congestive heart failure (DRG 127), pneumonia (DRG 89–91), chronic obstructive pulmonary disease (DRG 88), chronic renal failure (ICD-9 code 585), and diabetes mellitus (DRG 294, 295). To mimic a population-based study, patients residing outside the state of California were excluded. For patients with multiple admissions, only the first admission was included so that one patient could not contribute multiple data points.

Predictor variables and outcomes

Ethnicity was based on patient self-report of race, which is selected from the following categories on hospital admission—white, black, Asian/Pacific Islander, Native American/Alaskan/Aleutian, and others—and on self report of Hispanic origin. For this analysis, we defined white as whites not of Hispanic origin, African-American as blacks not of Hispanic origin, Hispanic as all those patients who identified themselves as being of Hispanic ethnicity, Asian as Asian/Pacific Islanders not of Hispanic origin, and Other/unknown as all others. The nonwhite group consisted of African-Americans, Asians, and Hispanics.

Beginning in 1999, OSHPD required hospitals to indicate the presence of a physician DNR order within 24 hours of hospital admission. The DNR variable is abstracted by staff in hospitals and indicates whether the inpatient medical chart within 24 hours of admission contains a physician directive to limit resuscitative efforts (such as chest compressions, intubation, assisted ventilation, or defibrillation) in the event of a patient’s cardiac or pulmonary arrest.

For payment type categorization, we combined self-payers with those listed as indigent or receiving Medicaid and combined worker’s compensation with private, non-HMO insurance in order to simplify listing of results. Comorbidity scores were developed using a database version of the Charlson comorbidity index and are a summary of major secondary diagnoses weighted by severity. The index predicts 1-year mortality based on the presence of coded comorbidities. We categorized scores as 0, 1, 2, 3–4, or >4 to simplify presentation of results. OSHPD provides ongoing training and quality-improvement projects to hospital medical record staff to enhance reliability of coding.

Statistical analysis

We defined predictors of DNR in univariate and multivariable analyses with logistic regression, assessing the following characteristics: age, sex, ethnicity, admission source, payment type, and comorbidity scores. All of these factors were included in each multivariable logistic regression model. Analyses were performed for the entire cohort and for each diagnosis separately; in the overall cohort analysis, a variable for diagnostic group was included in the multivariable model.

To determine whether DNR influenced the association of in-hospital mortality with sex and ethnicity, we evaluated DNR as a confounder in stratified analysis and by compar-
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Full cohort</th>
<th>Stratified by DNR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ignoring DNR</td>
<td>Adjusted for DNR</td>
</tr>
<tr>
<td></td>
<td>(n = 327 890)</td>
<td>(n = 25 169)</td>
</tr>
<tr>
<td>Age, per decade</td>
<td>1.30</td>
<td>1.17</td>
</tr>
<tr>
<td>Female</td>
<td>0.92</td>
<td>0.85</td>
</tr>
<tr>
<td>Ethnicity (reference group: white)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>0.87</td>
<td>0.98</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.87</td>
<td>0.97</td>
</tr>
<tr>
<td>Asian</td>
<td>1.04</td>
<td>1.24</td>
</tr>
<tr>
<td>Payment (reference group: Medicare)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medi-Cal/indigent</td>
<td>1.35</td>
<td>1.32</td>
</tr>
<tr>
<td>HMO</td>
<td>1.01</td>
<td>0.96</td>
</tr>
<tr>
<td>Non-HMO private</td>
<td>1.20</td>
<td>1.15</td>
</tr>
<tr>
<td>Admission source (reference group: non-nursing home)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nursing Home</td>
<td>1.75</td>
<td>1.52</td>
</tr>
<tr>
<td>Comorbidity score (reference group: 0)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.14</td>
<td>1.05</td>
</tr>
<tr>
<td>2</td>
<td>0.91</td>
<td>0.85</td>
</tr>
<tr>
<td>3 or 4</td>
<td>1.02</td>
<td>0.72</td>
</tr>
<tr>
<td>5 or greater</td>
<td>1.51</td>
<td>1.42</td>
</tr>
<tr>
<td>Admission diagnosis (reference group: angina)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>10.1</td>
<td>8.49</td>
</tr>
<tr>
<td>COPD‡</td>
<td>10.5</td>
<td>8.31</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>20.5</td>
<td>16.0</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>20.5</td>
<td>18.2</td>
</tr>
<tr>
<td>Stroke</td>
<td>91.2</td>
<td>83.8</td>
</tr>
<tr>
<td>Chronic renal failure</td>
<td>73.8</td>
<td>62.7</td>
</tr>
<tr>
<td>DNR</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*All models adjusted by age, sex, race, admission source, payment source, comorbidity scores, and admission diagnosis, as categorized above. DNR = do-not-resuscitate orders, OR = odds ratio, CI = confidence interval.
†Comorbidity scores were based on the Charlson comorbidity index adapted for administrative databases.
‡Chronic obstructive pulmonary disease.
ing associations with and without DNR included in multivariable models that included all variables. We compared sex-mortality and ethnicity-mortality associations in those with and without a DNR order; \( P \) values of these associations were determined using the chi-squared test. In multivariable analysis, we used logistic regression to define predictors of in-hospital mortality with and without including DNR in the models. Multivariable analysis was performed for the entire cohort and in groups stratified by DNR status. All demographic variables were included in the models, and separate analyses were performed for each diagnostic group, as described above. We did not test for statistical differences in the mortality odds ratios before or after DNR was included in the model, as statistically insignificant differences could have clinical significance.

We were concerned that adjustment for prognostic factors was incomplete in multivariable models of in-hospital mortality, so we performed a sensitivity analysis using propensity scores for DNR in the mortality model, using both continuous and categorical scores.\(^9\) Propensity scores summarize the association of multiple potential confounders in a single variable; in this study, the propensity score summarized the association of demographics and other patient characteristics with DNR status. The Statistical Analysis System (SAS) software (8.1 release, SAS Institute Inc.; Cary, NC) and the Stata statistical package (version 7.0, Stata Corporation; College Station, TX) were used for statistical analysis.

**Results**

**Cohort**

In 1999, there were 466 539 admissions to nonfederal hospitals in California for 1 of the 7 diagnoses studied, 130 329 of which were not first admissions during the period of study. A cohort of 327 890 patients remained after excluding admissions for non-California residents (n = 8320). DNR orders were written within 24 hours for 7.7% of admissions (n = 25 195). The average age was 64.5 years (median 70 years, intraquartile range, 54–80 years); women constituted 53.0% of the cohort, whites 64.0%, African-Americans 10.2%, Hispanics 11.7%, and Asians 6.7%. Comorbidity scores ranged from 0 to 28, with 57.1% of patients having a score of 0–1, consistent with limited life-threatening comorbidity.

**Predictors of DNR**

In unadjusted analyses, patients who were white, female, and older were more likely to have DNR orders (Table 1). Patients with higher comorbidity scores and admission from a nursing home were also more likely to have DNR orders. These associations persisted in the analyses adjusted by age, sex, ethnicity, admission source, payment source, comorbidity scores, and admission diagnosis (Table 1). In univariate analysis, rates of DNR orders were higher in those with Medicare coverage, but this association was not apparent after adjustment. Rates of DNR orders varied by admission diagnosis (Table 1). Women and non-Hispanic whites were more likely to be DNR for all admission diagnoses (Table 2).

**Predictors of mortality**

Overall mortality for the entire cohort was 4.1% (n = 13 549). Those with a DNR order were more likely to die (odds ratio [OR] 9.55, 95% confidence intervals [CI], 9.20–9.92, \( P < 0.001 \)), and the association persisted in the multivariable analysis (Table 3). When the score for propensity of DNR orders based on demographic and other patient characteristics was used in the mortality model instead of the individual demographic variables, DNR remained strongly associated with mortality (overall OR 7.17, \( P < 0.001 \); OR 4.48–15.7 among the different diagnostic groups, all with \( P < 0.001 \)).

**Influence of DNR on predictors of mortality**

Among those without a DNR order, in-hospital mortality was similar for women compared with men (2.6% vs. 2.8%, \( P < 0.001 \)). Among those with a DNR order, in-hospital mortality was lower in women than in men (19.4% vs. 23.7%, \( P < 0.001 \)). Similarly, in comparing in-hospital mortality between whites and other ethnicities, mortality rates were similar (non-whites vs. whites, 2.6% vs. 2.8%, \( P = 0.005 \)) among those without DNR orders. However, among those with DNR orders, in-hospital deaths were more frequent in non-whites (22.8% vs. 20.8%, \( P = 0.003 \)). The directions of these associations were similar after adjustment (Table 3).

After adjustment, women were less likely to die in the hospital overall and in most diagnostic groups (overall women vs. men: OR 0.92, 95% CI 0.88–0.95, \( P < 0.001 \); Table 3 and Figure 1). The difference between women and men was generally greater when DNR was included in the analysis, and women appeared to survive more frequently than before DNR was included in the model (overall women vs. men: OR 0.85, 95% CI 0.82–0.88, \( P < 0.001 \); Table 3 and Figure 1).

When DNR was omitted from multivariable models, in-hospital mortality appeared lower in non-whites compared with whites (overall non-whites vs. whites: OR 0.92, 95% CI 0.89–0.96, \( P < 0.001 \)). However, when DNR status was included, the apparent survival advantage for non-whites was reduced in every diagnostic group (Figure 2) and was eliminated overall (overall non-whites vs. whites: OR 1.06, 95% CI 1.01–1.11, \( P = 0.01 \); Table 3).

Multivariable analyses stratified by those with DNR orders and those without (Table 3) showed that women were more likely than men to survive in both groups (women vs. men with DNR orders: OR 0.76, 95% CI 0.72–0.81, \( P < 0.001 \); without DNR orders: OR 0.89, 95% CI 0.85–
and that non-whites were as likely to die as whites in both groups (non-whites vs. whites with DNR orders: OR 1.08, 95% CI 0.99–1.17, P = 0.08; without DNR orders: OR 1.01, 95% CI 0.96–1.06, P = 0.66).

**Discussion**

In this study of Californian in-patients, we found that a physician DNR order within 24 hours of admission was associated with in-hospital death even after adjustment for age, demographic characteristics, and comorbidity score, and in an analysis with propensity scores. Because cardio-pulmonary resuscitation has a limited impact on survival, we assume that DNR reflects less aggressive care overall, as demonstrated in other studies. Although the presence of a DNR order within 24 hours of admission would be expected to be associated with disease severity as well, it also indicates a patient’s preferences for care. Thus, the presence of a DNR order likely captures information not otherwise represented in administrative databases.

After adjustment, we found that odds of having a DNR order were 19% greater in women compared with men, and 43% greater in non-Hispanic whites compared with other ethnicities. These findings confirm results of several previous studies. The one large study that did not show a difference in DNR usage between whites and non-whites required informed consent to participate, and patients who are willing to participate in a study may not be representative of all patients. Ethnic differences in use of advanced directives likely reflect cultural preferences, such as whether death should occur in the hospital. Although details of patient management are limited in our study, it is the largest to evaluate DNR use in a well-defined, socio-economically diverse multiethnic population, and provides new data on frequency of DNR orders in Hispanics and Asian-Americans.

We found that DNR status was a confounder in the association of ethnicity and mortality. Non-whites were less likely to die in the hospital in an analysis ignoring DNR status, but this apparent mortality benefit disappeared after adjustment for DNR status. Thus, higher rates of DNR orders in whites may create the appearance

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>Mortality Odds Ratios (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Before inclusion of DNR: 0.92 (0.88-0.95)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.85 (0.82-0.88)</td>
</tr>
<tr>
<td>Angina</td>
<td>Before inclusion of DNR: 0.67 (0.40-1.12)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.64 (0.38-1.08)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Before inclusion of DNR: 0.80 (0.58-1.11)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.78 (0.56-1.08)</td>
</tr>
<tr>
<td>COPD</td>
<td>Before inclusion of DNR: 0.78 (0.66-0.93)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.74 (0.62-0.89)</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>Before inclusion of DNR: 0.90 (0.82-0.99)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.85 (0.77-0.94)</td>
</tr>
<tr>
<td>Chronic Renal Failure</td>
<td>Before inclusion of DNR: 1.00 (0.88-1.12)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.94 (0.83-1.07)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>Before inclusion of DNR: 0.86 (0.79-0.93)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.80 (0.74-0.87)</td>
</tr>
<tr>
<td>Stroke</td>
<td>Before inclusion of DNR: 1.03 (0.97-1.08)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.95 (0.90-1.00)</td>
</tr>
</tbody>
</table>

**Figure 1**  Adjusted odds ratios for in-hospital mortality, women vs. men, in the 7 diagnostic groups and the total cohort before (gray) and after (black) inclusion of DNR status in the model. Odds ratios were consistently lower when DNR status was included in the model. DNR = do-not-resuscitate orders, COPD = chronic obstructive pulmonary disease.
of a survival advantage for other ethnicities, and unmeasured differences in use of DNR orders may mask ethnic and sex disparities. In addition, odds of in-hospital mortality tended to be lower for women, and the survival advantage for women was even greater after DNR status was included in the models. Failure to capture DNR in the model could misrepresent the component of outcome that is due to disparities in the delivery of medical care. Furthermore, DNR may be a confounder in other studies of risk factors for mortality.

Previous studies have ignored DNR status when evaluating outcome differences among ethnicities. In these studies of a variety of diseases and populations, some have documented lower mortality rates for whites,19-21,34-39 some found no difference,39-43 and some reported higher mortality rates for whites.18,43-45 Differences in follow-up duration, study settings, and included diseases may explain these variable results. In this study, we found that ethnic disparities varied broadly among the diagnostic categories, but that whites tended to have higher overall rates of in-hospital mortality when DNR status was ignored. However, when DNR was included, the apparent benefit of being nonwhite was reduced for every disease category studied and, overall, whites had lower mortality rates than other ethnicities.

This study is limited by its use of administrative data. Although coding of most variables included in the analysis is probably reliable, the DNR item, new in 1999, may be more vulnerable to errors as abstractors gain experience with it. However, random errors in coding tend to reduce the impact of the variable. Another major limitation of this study is incomplete knowledge about severity of disease, which is inadequately captured by demographic characteristics and the comorbidity score we employed. In fact, we found that both DNR and mortality were associated with measures of disease severity, including age and a comorbidity score, so DNR may be acting as a surrogate for unmeasured disease severity. However, this would not fully explain the shift in ethnic disparity we demonstrated because rates of DNR were greater in whites, a group that generally has lower disease severity at admission.46-48 Furthermore, if DNR was simply a marker for unmeasured disease severity, adjusting for it should always reduce the apparent sex-mortality and ethnicity-mortality associations,49 and this was not the case. Finally, we considered

### Table: Mortality Odds Ratios

<table>
<thead>
<tr>
<th>Diagnostic group</th>
<th>Mortality Odds Ratios (95% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Before inclusion of DNR: 0.92 (0.89-0.96)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 1.06 (1.01-1.11)</td>
</tr>
<tr>
<td>Angina</td>
<td>Before inclusion of DNR: 1.06 (0.61-1.86)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 1.16 (0.66-2.05)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Before inclusion of DNR: 1.06 (0.75-1.50)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 1.21 (0.85-1.72)</td>
</tr>
<tr>
<td>COPD</td>
<td>Before inclusion of DNR: 0.55 (0.42-0.71)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.72 (0.55-0.94)</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>Before inclusion of DNR: 0.61 (0.54-0.69)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.71 (0.63-0.81)</td>
</tr>
<tr>
<td>Chronic Renal Failure</td>
<td>Before inclusion of DNR: 0.83 (0.73-0.95)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.92 (0.80-1.06)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>Before inclusion of DNR: 0.75 (0.67-0.83)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 0.92 (0.82-1.02)</td>
</tr>
<tr>
<td>Stroke</td>
<td>Before inclusion of DNR: 0.95 (0.90-1.02)</td>
</tr>
<tr>
<td></td>
<td>After inclusion of DNR: 1.08 (1.02-1.16)</td>
</tr>
</tbody>
</table>

Figure 2 Adjusted odds ratios for in-hospital mortality, non-whites vs. whites, in the 7 diagnostic groups and the total cohort before (gray) and after (black) inclusion of DNR status in the model. Odds ratios were consistently greater when DNR status was included in the model. DNR = do-not-resuscitate orders, COPD = chronic obstructive pulmonary disease.
DNR orders only within the first 24 hours of admission, when a previously established advanced directive, rather than disease course, is likely to be more important. Thus, unmeasured disease severity is not likely to explain fully the association of DNR with sex and ethnicity or the shifting of sex-mortality and ethnicity-mortality associations when DNR is included in the models.

In this study, consideration of DNR status was important in estimations of sex and ethnic differences. If we had not adjusted for DNR status, we would have concluded that non-whites had lower rates of in-hospital mortality than whites. However, with adjustment for DNR, we found just the opposite. With adjustment, we found higher rates of death in Asians but similar rates in African-Americans and Hispanics compared with whites. Administrative data similar to that in the OSHPD database has been an important source of information about outcome differences between ethnicities and sexes. Failure to include DNR in previous studies of mortality could have led to a misrepresentation of ethnic and sex differences in outcomes. In addition, hospital mortality rates generated from administrative data are often used to assess hospital quality. The adjustment for DNR may be important in evaluating hospital quality because frequency of use is likely to vary between institutions and to contribute to differences in adjusted mortality rates. Future studies should routinely consider the association of DNR status with patient outcomes.

Acknowledgments

Many thanks to Eugene Bardach, Heather Fullerton, and Mendocino Steele for their editorial contributions.

References


Iatrogenic events resulting in intensive care admission: Frequency, cause, and disclosure to patients and institutions

Lisa Soleymani Lehmann, MD, PhD, MSc, Ann Louise Puopolo, BSN, RN, Shimon Shaykevich, MS, Troyen A. Brennan, MD, JD, MPH

Division of General Medicine and Primary Care, Brigham and Women’s Hospital and Harvard Medical School.

ABSTRACT

PURPOSE: To identify the frequency and type of iatrogenic medical events requiring admission to an intensive care unit. To assess the consequences of iatrogenic medical events for patients and institutions. To assess the prevalence of disclosure of iatrogenic medical events to patients, surrogates, and institutions.

METHODS: The project on Care Improvement for the Critically Ill enrolled 5727 patients to 8 intensive care units at 4 Boston teaching hospitals. To determine the nature, consequences, and disclosure of iatrogenic medical events, we did a retrospective chart review on all patients whose admission to an intensive care unit was precipitated by an iatrogenic event.

RESULTS: Sixty-six patients (1.2 %) were identified by an intensive care unit’s clinical team as having an iatrogenic medical event as the primary reason for admission to the unit. The majority (29, or 45%) of iatrogenic medical events were secondary to technical error, but a high percentage (21, or 33%) was due to iatrogenic drug events. Twenty-two (34%) cases were assessed by the investigators to have been preventable. In 60 (94%) cases there was no documentation in the patient’s chart of communication to the patient regarding the reason for admission to the intensive care unit. In 11 (17%) cases there was documentation of a discussion with the surrogate about the reason for admission to the unit. In only 3 (5%) cases was there documentation that the patient or surrogate was informed that an iatrogenic medical event was the reason for admission to the intensive care unit. Incident reports or malpractice claims were filed in only 4 (6 %) cases.

CONCLUSION: The frequency of iatrogenic medical events resulting in admission to intensive care units is lower than previous studies have reported. Iatrogenic drug events continue to be an important source of error. A considerable percentage of iatrogenic events may be preventable. Health care professionals rarely document disclosure of iatrogenic events to patients and surrogates.

© 2005 Elsevier Inc. All rights reserved.

KEYWORDS: Iatrogenic medical events; Intensive care; Disclosure

In its report To Err is Human, the Institute of Medicine estimated that medical errors kill between 44 000 and 98 000 people a year in U.S. hospitals.1 This report, in conjunction with the growing empirical data detailing medical errors in this country,2,3,4 has generated a national policy debate and proposals for system changes to reduce the rate of medical errors.5 Although some have challenged the accuracy of the Institute of Medicine estimate,6,7 con-
cern about errors in medicine remains at the forefront of public consciousness.

Although all medical errors should be a concern for our society, errors that either result in serious consequences for patients or that are preventable are of particular concern. We focused on errors that necessitated admission to an intensive care unit. Our goal was to better understand the errors that were most likely to have a negative impact on patient outcomes and to determine what type of communication with patients or surrogates occurs after an error that results in admission to an intensive care unit.

Methods

Background

This study was conducted as part of the project on Care Improvement for the Critically Ill, organized in 1997 by a consortium from the intensive care units at Harvard teaching hospitals in response to emerging controversies in critical care medicine. The complete methods for this project have been published elsewhere.8

Settings

The project on Care Improvement for the Critically Ill was a prospective observational study that examined decision making and satisfaction with care provided in 8 intensive care units at 4 teaching hospitals. This included 4 surgical intensive care units, 3 medical intensive care units, and 1 pediatric intensive care unit. The institutional review boards at participating hospitals approved the study.

Eligibility

All patients admitted to the intensive care unit during the study period were eligible for enrollment. Only a patient’s first admission was included to avoid counting 2 outcomes for the same individual. Any patient already admitted to the intensive care unit before the start of the study was excluded. Clinicians in the intensive care unit identified patients with an iatrogenic event. Each morning a trained research assistant attended rounds and queried the clinical team regarding whether a patient was admitted to the intensive care unit because of an iatrogenic event. If a member of the team questioned what was meant by the term “iatrogenic event,” the following definition was provided: “an injury or event induced inadvertently by a physician, other clinical staff or a medical treatment or diagnostic procedure resulting in a patient’s admission to the intensive care unit.”

Data collection

Data were collected between November 1998 and November 1999. Research nurses, who attended a training session and adhered to a detailed coding instruction manual, abstracted data from the medical record. Variables abstracted from the medical record included demographic information and measures of patient acuity, including the Simplified Acute Physiology Score II,9 a prognostic scoring system that estimates probability of outcome from physiologic measures.

Physicians reviewed the medical records using a structured abstraction instrument developed for this study to collect information on the details of iatrogenic medical events. The abstraction instrument included information on the etiology requiring intensive care unit admission, a classification of iatrogenic drug events, an assessment of whether the event had been avoidable, and documentation of communication about the event. Physicians were instructed to select all the possible reasons for an iatrogenic event. The entire medical record was available to reviewers. Reviewers were instructed to carefully review the record for at least 2 days prior to the intensive care unit admission through discharge from the unit. To assess inter-rater reliability, 1 physician reviewed all patient charts and a second physician reviewed a random subset of 28 (44%) medical records.

Definitions

An iatrogenic event was defined as an unintended injury or harm to a patient resulting from health care management rather than a disease process. An iatrogenic event was considered preventable if it was avoidable using any means currently available, unless those means were not considered standard of care.2 Iatrogenic events were categorized as either technical, diagnostic, or medication errors. Technical error was defined as medical procedure events such as injury occurring during an operation or bleeding.10 Diagnostic error was defined as delayed or incorrect diagnosis or therapy.11 In order to place our results within the context of previous data in this field, the relationship between a medication error and an iatrogenic event was categorized as definite, probable, possible, conditional, or doubtful.12

Statistical analysis

The characteristics of patients are expressed as means ± standard deviation or as the frequency of patients or events. Kappa statistics were performed to determine inter-rater reliability for judgments concerning preventability of iatrogenic medical events. All statistical calculations were performed using SAS version 8 statistical software (SAS Institute; Cary, NC).

Results

Characteristics of patients

A total of 5727 patients were admitted to the intensive care units during the period of November 1998 to March
Among these, 873 (15%) patients were enrolled in the Care Improvement for the Critically Ill study. In 1999 Darchy et al. did a retrospective review of admissions during a 1-year period to a French general hospital and found a similar percentage of admissions (68 or 10.9%) to the intensive care unit resulted from iatrogenic medical events.14 Because many errors may not be documented in the medical record15,4 and implicit review was performed for 64 (97%) of the 66 patients. Two charts (1.2%) were identified as having an iatrogenic medical event. In only 1 (2%) case did we find the use of the term “iatrogenic” in the chart to describe the event that resulted in admission to the intensive care unit. In 60 (94%) cases, there was no documentation in the chart of communication to the patient regarding the reason for admission to the unit. Because 768 (88%) patients enrolled in the Care Improvement for Critically Ill study did not have the capacity to make decisions for themselves, we assessed disclosure of iatrogenic events to patient surrogates. In 11 (17%) cases, there was documentation of a discussion with the surrogate about the reason for admission to the intensive care unit (Table 2).

**Etiology of iatrogenic medical events resulting in intensive care unit admission**

For 20 patients (31%), the primary reason for intensive care unit admission was respiratory decompensation. The majority of iatrogenic medical events (29, or 45%) were secondary to technical error, but a high percentage (21, or 33%) were also due to adverse drug events (Table 1). Adverse drug events were further categorized into dosage error for 9 (43%) events, idiosyncratic reaction for 7 (33%), frequency error for 2 (10%), errors that did not fit any general classification for 2 (10%), and wrong drug to patient for 1 (5%). Narcotic analgesics were the most common type of drug (9 events, or 43%) resulting in an iatrogenic event, with sedative hypnotics being responsible for 5 (24%) of adverse events.

Invasive procedures were related to the iatrogenic event in 51 (80%) cases. The types of procedures involved were surgical 28 (44%), gastrointestinal 9 (14%), line placement 4 (6%), renal 3 (5%), cardiac catheterization 3 (5%), obstetrical and gynecological 2 (3%), and pulmonary 2 (3%). When procedures were involved, perforation (12, or 19%) and hypotension (12, or 19%) were the most common iatrogenic outcomes resulting from the procedure. Reviewers were asked to evaluate whether the iatrogenic medical event was preventable. The inter-rater reliability for determining preventability was only fair. In 22 (34%) cases, reviewers thought the event was preventable; in 9 (14%) cases, the iatrogenic medical event was thought not to be preventable; and in 33 (52%) cases, the preventability of the iatrogenic medical event could not be evaluated with certainty (kappa=0.5).

**Documentation of disclosure of iatrogenic medical event**

In only 1 (2%) case did we find the use of the term “iatrogenic” in the chart to describe the event that resulted in admission to the intensive care unit. In 60 (94%) cases, there was no documentation in the chart of communication to the patient regarding the reason for admission to the unit. Because 768 (88%) patients enrolled in the Care Improvement for Critically Ill study did not have the capacity to make decisions for themselves, we assessed disclosure of iatrogenic events to patient surrogates. In 11 (17%) cases, there was documentation of a discussion with the surrogate about the reason for admission to the intensive care unit (Table 2).
Similarly, explicit nonphysician review is also limited. It is therefore possible that the teaching hospitals are less likely to suffer preventable iatrogenic drug events. It is possible that some patients with an iatrogenic event were admitted to the intensive care unit after morning rounds. It is possible that some patients with an iatrogenic event were admitted to the intensive care unit after morning rounds and died or were discharged prior to the next morning rounds. Thomas et al. also showed that patients in major teaching hospitals are less likely to suffer preventable iatrogenic drug events. It is therefore possible that the inclusion of major academic teaching institutions within our study is part of the reason for our finding a lower frequency of iatrogenic events.

Most surprising among our data was the low rate of documented discussions concerning the reason for admission to the intensive care unit with patients or surrogates. It is possible that health care professionals discussed iatrogenic events with patients without documenting their discussions. The extremely low rate of documentation of disclosure, however, raises questions about health care professionals’ understanding that disclosure is vital to changing systems and the resultant improvement in health care.

Disclosure of error should be encouraged from both an ethical and legal perspective. Failing to disclose errors undermines patients’ trust in physicians and the public’s trust in medicine as a profession. Furthermore, failure to disclose known errors is a manifestation of disrespect for patients because it entails deception. Nondisclosure of error also undermines efforts to improve the safety of medical practice. Institutions that are unaware of errors cannot make efforts to improve the systems that resulted in the development of the error. Previous studies have indicated that patients have a clear preference for disclosure of errors even when the error was minor. The Joint Commission on Hospital Accreditation recently introduced patient safety and medical/health care error reduction standards for hospital practitioners. These stipulate that “patients and, when appropriate, their families are informed about the outcomes of care, including unanticipated outcomes.”

There are many reasons why physicians may not disclose errors to patients. Lack of knowledge about reporting rules, uncertainty about how to disclose errors, concerns about upsetting patients, and a fear of the consequences of disclosure may influence physicians’ response to an iatrogenic event. If we are to succeed in improving the quality of care, physicians must be educated about the importance of disclosure as a catalyst for system change, as a protection against malpractice, and as information that patients prefer to receive. Most importantly, we need to be certain that physicians have the communication skills to sensitively disclose iatrogenic events.

Our results are limited by our inclusion of only major academic teaching institutions. Our reliance on implicit physician review to make judgments concerning the preventability of iatrogenic events is another limitation.

Our data about disclosure of error is limited by medical record review. It is possible that health care providers had more frequent discussions with patients or surrogates regarding iatrogenic events and the reason for admission to the intensive care units and did not document these discussions in the medical record. The low rate of incidence report filing, however, suggests that, even when health care providers do discuss errors with patients, they do not disclose those conversations to the institution. Patients were also enrolled in our study prior to the Institute of Medicine report To Err Is Human. It is possible that this report may have led to an increase in the frequency of disclosure of iatrogenic events to patients and families.

Iatrogenic medical events have negative consequences for patients, population health, and the cost of medical care. Our study suggests that iatrogenic events continue to be an important cause of admission to intensive care units and in many cases may be preventable. The challenge that remains is for health care institutions to develop systems to better identify errors, reduce the incidence of preventable iatrogenic events, and ultimately provide better quality care.

Table 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Iatrogenic” present in chart</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Reason for intensive care unit admission discussed with patient</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Reason for intensive care unit admission discussed with surrogate</td>
<td>11 (17)</td>
</tr>
<tr>
<td>Iatrogenic nature of intensive care unit admission discussed with patient or surrogate</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Incident report filed with hospital</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Medical record suggests litigation</td>
<td>3 (5)</td>
</tr>
</tbody>
</table>

The American Journal of Medicine, Vol 118, No 4, April 2005
to our patients. Developing a system of anonymous physician self-reporting of errors might come closest to capturing the true rate of iatrogenic events within an institution. Improving the systems by which drugs are ordered and administered could prevent some iatrogenic drug events.3,4

Further data is needed on the disclosure of iatrogenic medical events. Are physicians having conversations with patients or surrogates about iatrogenic events and, if so, what is the content and process of that communication? Future research should also be directed at assessing patients’ perspectives on the ideal way to respond to iatrogenic medical events.

Acknowledgements

We would like to thank Eric Poon, MD, for his assistance with chart review and Jeffrey P. Burns, MD, MPH, for his comments on a draft of this manuscript.

References

Helicobacter pylori eradication in the management of patients with idiopathic thrombocytopenic purpura

Roberto Stasi, MD,a Zaccaria Rossi, MD,b Elisa Stipa, MD,c Sergio Amadori, MD,c Adrian C. Newland, MA,d Drew Provan, MDd

aDepartments of Medical Sciences and bGastroenterology, “Regina Apostolorum” Hospital; cDepartment of Hematology, University of Rome “Tor Vergata”; and dDepartment of Hematology, St Bartholomew’s and The Royal London School of Medicine and Dentistry.

ABSTRACT

PURPOSE: To investigate the relation between Helicobacter pylori infection and the clinical features of idiopathic thrombocytopenic purpura (ITP), and to examine the effects of H. pylori eradication on platelet counts.

METHODS: A 13C urea breath test for H. pylori infection was performed in a cohort of 137 consecutive patients with ITP. Patients who tested positive received standard eradication therapy if their platelet count was \( \geq 50 \times 10^9/L \) or if they had symptoms of dyspepsia.

RESULTS: H. pylori infection was detected in 64 patients (47%), and was not associated with dyspepsia or other clinical or laboratory features. Eradication therapy was successfully administered to 52 patients. Platelet responses were observed in 17 (33%) of these patients, which lasted for more than 1 year in 11 patients. Duration of ITP was shorter among responders than nonresponders. Only one response was observed among patients with severe thrombocytopenia (platelet count \( \geq 30 \times 10^9/L \)).

CONCLUSION: The prevalence of H. pylori infection in patients with ITP is similar to that found in the general population. Infection is not associated with distinctive features of the disease. H. pylori eradication may improve the platelet counts in adults in whom the ITP is of recent onset and in those with less severe degrees of thrombocytopenia, but was not effective in patients with chronic severe ITP.

© 2005 Elsevier Inc. All rights reserved.

KEYWORDS: Idiopathic thrombocytopenic purpura; Helicobacter pylori; Dyspepsia

Helicobacter pylori is a gram-negative bacillus that colonizes the mucous layer of the human stomach. The bacterium has been causally linked with a diverse spectrum of gastrointestinal disorders, including gastritis, peptic ulcer disease, nonulcer dyspepsia, gastric adenocarcinoma, and mucosa-associated lymphoid tissue lymphoma. Several investigators have studied whether H. pylori causes non-digestive diseases, but these associations, if any, are uncertain.

Several studies have described an improvement in thrombocytopenia among patients with chronic idiopathic thrombocytopenic purpura (ITP) following H. pylori eradication, but other studies have not. However, most of these previous studies involved a relatively small number of patients, the median observation following eradication was often less than 1 year, and the effects of prior therapies were unclear. In addition, studies usually included patients with mild thrombocytopenia who would not ordinarily have been treated. Therefore, the effects of H. pylori eradication...
in the management of patients with ITP remains undetermined.

In this study, we investigated the prevalence of *H. pylori* infection in a large cohort of consecutive patients with ITP and the relation between infection and the clinical features of the disease, and examined the effects of *H. pylori* eradication on platelet counts.

### Methods

#### Patients

Between December 1999 and January 2003, we investigated the presence of gastric *H. pylori* infection and dyspepsia in 137 consecutive adult patients (70 in the United Kingdom and 67 in Italy) attending our clinics with a diagnosis of ITP for at least 3 months. All patients were white. The diagnosis of ITP was made using standard criteria.19 All patients signed an institutional review board-approved informed consent form.

#### Definition and assessment of dyspepsia

Dyspepsia was defined as intermittent or persistent pain or discomfort centered in the upper abdomen. This definition excluded patients with heartburn or acid regurgitation as the predominant symptom, as these symptoms were thought to be due to gastroesophageal reflux disease.20 The severity of dyspepsia during the 6 months preceding the visit was assessed using the Glasgow dyspepsia severity score,21 which evaluates the frequency of symptoms (0 to 5 scale); the effect of dyspepsia on normal activities (0 to 2); the number of days of work missed because of dyspepsia (0 to 2); and the frequency of medical consultations (0 to 2), home visits by a physician (0 to 2), tests for dyspepsia (0 to 2), use of over-the-counter medications (0 to 2), and use of prescribed medications (0 to 3). The scoring system ranges from 0 to 20, with higher scores indicating more severe dyspepsia.

#### *H. pylori* infection assessment

The BreathMAT 13C Urea Breath Test (Finnigan MAT, Bremen, Germany) was performed in all patients. This assay has a sensitivity and specificity of 90% to 95% for *H. pylori* infection.1 The 13C enrichment in the expired breath was measured by automated breath 13C analysis by means of continuous-flow isotope ratio mass spectrometry. Test results were considered positive when the delta-over-base-line value was greater than 4%. To minimize false-negative results, patients had not received antacids or antibiotics for at least 2 weeks.

### *H. pylori* eradication therapy

Patients were given antimicrobial therapy only if they tested positive for *H. pylori* infection and had a platelet count <50 × 10^9/L or dyspepsia. The regimen for *H. pylori* eradication consisted of amoxicillin (1000 mg twice daily), clarithromycin (500 mg twice daily), and omeprazole (20 mg twice daily) for 7 days. To assess the efficacy of eradication therapy, the urea breath test was repeated 6 to 8 weeks following antimicrobial therapy. If *H. pylori* was not eradicated after initial treatment, a second regimen that included metronidazole was used.

### Laboratory studies

The baseline platelet count was the average of three platelet counts during the 2 weeks before eradication therapy. However, the three counts had to be in the same range of values, that is <50 × 10^9/L or ≥50 × 10^9/L. In fact, platelet counts may show ample variations from week to week, which can make patients easily shift from severe forms of thrombocytopenia to milder ones. In addition to a complete blood count, laboratory evaluation included serum chemistry profiles; direct and indirect Coombs’ (antiglobulin) tests; prothrombin time; partial thromboplastin time; fibrinogen levels; and serologic tests for hepatitis B and C, human immunodeficiency virus, cytomegalovirus, and toxoplasmosis. Immunology studies included tests for C3 and C4 fractions of complement, antinuclear antibodies, antidiagonal-stranded DNA antibodies and anticardiolipin antibodies, lymphocyte subsets, and serum immunoglobulin (Ig) (IgG, IgA, and IgM) concentrations. Anticardiolipin antibodies and platelet-associated IgG concentrations were determined as described previously.22

### Platelet response criteria

Because disease severity varied among our patients, we developed response criteria that differed from those used commonly among patients with severe thrombocytopenia (<30 × 10^9/L).23 We defined a platelet response as either a major response (a rise of greater than 50 × 10^9/L relative to baseline) or a partial response (a rise of greater than 30 but less than 50 × 10^9/L relative to baseline). This definition was based on clinical considerations, since patients with counts in the range of 30 to 50 × 10^9/L may have occasional bruises, and may require platelet transfusions if they are bleeding or before surgery or obstetric delivery.19

### Statistical analysis

Analyses were carried out using the STATISTICA for Windows (StatSoft, Inc., Tulsa, Oklahoma) software package. Comparisons of continuous variables between the various groups were performed using nonparametric tests (Mann-Whitney U test, Kruskal-Wallis analysis of vari-
ance), as appropriate, and correlation analysis with the Spearman rank correlation test. The Fisher exact test was used to evaluate the relation between two dichotomous variables. Changes in platelet counts and platelet-associated IgG levels over time were examined with the Wilcoxon matched-pairs test. A value of p ≤.05 was designated as statistically significant.

Results

About half of the patients had undergone previous treatment for ITP (Table 1). *H. pylori* infection was diagnosed in 64 (47%) of the 137 patients. The prevalence of infection was similar in the United Kingdom (47% [33/70]) and Italy (46% [31/67]). Infection was more common in older patients (Figure 1). The median age of infected patients was 59 years (range, 22 to 78 years), compared with 41 years (range, 15 to 80 years) in uninfected patients (P <.001). There was no association between *H. pylori* infection and other clinical variables (Table 2). Dyspepsia was reported by 70 patients (51%) and was associated with a greater number of previous therapies for ITP (P = 0.03), but not with *H. pylori* infection (P = 0.30), age (P = 0.72), sex (P = 0.12), or disease duration (P = 0.97).

### Platelet response to eradication therapy

Fifty-two patients with *H. pylori* infection received standard eradication therapy. The baseline platelet count was <30 × 10^9/L in 22 of these patients, 30 to 49 × 10^9/L in 19 patients, and ≥50 × 10^9/L in 11 patients. Sixteen patients were also treated concurrently with prednisone to maintain a safe platelet count. Only 1 of these 16 patients achieved a platelet response to eradication therapy. Prednisone was administered for 2 weeks in this patient, while it was given for as long as 3 months in the other 15 patients. Other forms of immunosuppressive therapy had been discontinued for at least 2 weeks before eradication therapy, and was not resumed until 3 months from the start of eradication therapy.

Eradication therapy was successful at treating *H. pylori* infection in all 52 patients. Antibiotics were discontinued by 3 patients before the full course had been administered because of adverse effects (abdominal pain or diarrhea). A second regimen that included metronidazole was required by 5 patients to achieve eradication. A major platelet response was obtained in 14 (27%) of the 52 patients, and a partial response in 3 patients (6%), for an overall response rate of 33%. The response rate was 54% (6/11) in patients with a baseline platelet count >50 × 10^9/L, and 53% (10/19) in those with a platelet count between 30 and 49 × 10^9/L. Only 1 patient (of 22) with a platelet count <30 × 10^9/L achieved a response. This patient, a 62-year-old woman with a 6-month history of severe thrombocytopenia, had been on therapy with prednisone and intravenous immunoglobulin. She was still being treated with prednisone at a daily dose of 25 mg when she started *H. pylori* eradication therapy.

Full details for responders are shown in Table 3. Responses occurred rapidly, and a marked improvement in platelet count was usually seen 2 weeks after eradication therapy had been completed.

---

**Table 1** Characteristics of patients at study entry, by country

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>United Kingdom (n = 70)</th>
<th>Italy (n = 67)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex</td>
<td>29 (41)</td>
<td>28 (42)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>52 ± 14</td>
<td>50 ± 18</td>
</tr>
<tr>
<td>ITP duration (months)</td>
<td>25 ± 19</td>
<td>24 ± 19</td>
</tr>
<tr>
<td>Previous therapies for ITP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prednisone</td>
<td>34 (49)</td>
<td>36 (54)</td>
</tr>
<tr>
<td>No treatment</td>
<td>36 (51)</td>
<td>31 (46)</td>
</tr>
<tr>
<td>Intravenous</td>
<td>19 (29)</td>
<td>15 (23)</td>
</tr>
<tr>
<td>Splenectomy</td>
<td>8 (11)</td>
<td>9 (13)</td>
</tr>
<tr>
<td>Pulse cyclophosphamide/lymphoma-like therapy</td>
<td>8 (11)</td>
<td>12 (18)</td>
</tr>
<tr>
<td>Azathioprine</td>
<td>15 (21)</td>
<td>9 (13)</td>
</tr>
<tr>
<td>High-dose dexamethasone</td>
<td>0</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Cyclosporine</td>
<td>1 (1)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>0</td>
<td>4 (6)</td>
</tr>
<tr>
<td>Vincristine</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Rituximab</td>
<td>2 (3)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Baseline platelet count (× 10^9/L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50–99</td>
<td>45 ± 25</td>
<td>43 ± 22</td>
</tr>
<tr>
<td>30–49</td>
<td>29 (41)</td>
<td>24 (36)</td>
</tr>
<tr>
<td>&lt;30</td>
<td>21 (30)</td>
<td>20 (30)</td>
</tr>
</tbody>
</table>

*ITP* = idiopathic thrombocytopenic purpura.

---

**Figure 1** Prevalence of *Helicobacter pylori* infection and dyspepsia, by age in patients with idiopathic thrombocytopenic purpura.
Predictors of response

Responders were younger than nonresponders (55 ± 8 years vs. 60 ± 14 years, \( P = 0.03 \)), had a higher baseline platelet count (48 ± 21 x 10^9/L vs. 31 ± 20 x 10^9/L, \( P = 0.001 \)), had a shorter disease duration (14 ± 7 months vs. 34 ± 23 months, \( P < 0.001 \)), and had not received prior therapies for ITP or had received therapies including only prednisone or intravenous immunoglobulin (Table 3; \( P < 0.001 \)). In particular, a response after eradication therapy was observed in 9 (64%) of 14 patients who had not received prior treatment, and in 8 (47%) of 17 patients who had received prednisone or intravenous immunoglobulin. All responders were below 65 years of age; 15 (88%) of 17 had a disease duration of less than 2 years.

Platelet-associated IgG levels

High pretreatment levels were demonstrated in 37 (84%) of the 44 patients in whom these were measured. After eradication, platelet-associated IgG levels showed a marked decrease compared with the levels before eradication among responders (140 ± 95 ng/10^7 platelets vs. 369 ± 277 ng/10^7 platelets, \( P = 0.04 \)). In nonresponders, platelet-associated

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>H. pylori-positive (n = 64)</th>
<th>H. pylori-negative (n = 73)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (%) or mean ± SD</td>
<td></td>
<td></td>
<td>---------</td>
</tr>
<tr>
<td>Male sex</td>
<td>24 (37)</td>
<td>33 (45)</td>
<td>0.39</td>
</tr>
<tr>
<td>Age (years)</td>
<td>58 ± 12</td>
<td>42 ± 16</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ITP (months)</td>
<td>26 ± 20</td>
<td>23 ± 18</td>
<td>0.24</td>
</tr>
<tr>
<td>Number of previous therapies</td>
<td>1.7 ± 1.8</td>
<td>1.4 ± 1.7</td>
<td>0.31</td>
</tr>
<tr>
<td>Platelet count (× 10^9/L)</td>
<td>42 ± 25</td>
<td>46 ± 23</td>
<td>0.38</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>36 (56)</td>
<td>36 (47)</td>
<td>0.30</td>
</tr>
<tr>
<td>Glasgow Dyspepsia Severity Score</td>
<td>7.4 ± 2.4</td>
<td>6.7 ± 2.1</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Table 2 Patients’ characteristics in relation to Helicobacter pylori infection

Table 3 Clinical characteristics of patients with idiopathic thrombocytopenic purpura who had a platelet response after Helicobacter pylori eradication

<table>
<thead>
<tr>
<th>Country</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Disease duration (months)</th>
<th>Previous therapies for ITP</th>
<th>Glasgow dyspepsia severity score</th>
<th>Platelet counts (× 10^9/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>61</td>
<td>F</td>
<td>23</td>
<td>Prednisone</td>
<td>0</td>
<td>96 140 MD MD 177 196 254 221</td>
</tr>
<tr>
<td>Italy</td>
<td>32</td>
<td>F</td>
<td>33</td>
<td>—</td>
<td>4</td>
<td>136 169 188 193 245 267 219 291</td>
</tr>
<tr>
<td>UK</td>
<td>51</td>
<td>M</td>
<td>6</td>
<td>Prednisone, intravenous immunoglobulin</td>
<td>6</td>
<td>97 220 287 251 234 215 227 195</td>
</tr>
<tr>
<td>Italy</td>
<td>50</td>
<td>F</td>
<td>17</td>
<td>—</td>
<td>4</td>
<td>88 133 156 178 171 185 167 193</td>
</tr>
<tr>
<td>UK</td>
<td>55</td>
<td>M</td>
<td>13</td>
<td>—</td>
<td>6</td>
<td>74 93 105 124 156 171 143 189</td>
</tr>
<tr>
<td>UK</td>
<td>57</td>
<td>F</td>
<td>12</td>
<td>—</td>
<td>5</td>
<td>110 131 151 151 151 164 173 156</td>
</tr>
<tr>
<td>Italy</td>
<td>52</td>
<td>F</td>
<td>5</td>
<td>—</td>
<td>7</td>
<td>73 81 94 130 141 165 153 156</td>
</tr>
<tr>
<td>UK</td>
<td>55</td>
<td>M</td>
<td>27</td>
<td>—</td>
<td>14</td>
<td>63 87 MD 121 MD 135 129 144</td>
</tr>
<tr>
<td>Italy</td>
<td>60</td>
<td>M</td>
<td>11</td>
<td>—</td>
<td>8</td>
<td>81 103 121 146 138 169 146 131</td>
</tr>
<tr>
<td>UK</td>
<td>62</td>
<td>F</td>
<td>6</td>
<td>Prednisone, intravenous immunoglobulin</td>
<td>0</td>
<td>34 98 137 105 151 123 116 120</td>
</tr>
<tr>
<td>Italy</td>
<td>62</td>
<td>M</td>
<td>16</td>
<td>—</td>
<td>4</td>
<td>83 107 MD MD 125 131 113 102</td>
</tr>
<tr>
<td>UK</td>
<td>59</td>
<td>F</td>
<td>7</td>
<td>Prednisone</td>
<td>8</td>
<td>58 79 67 84 93 88 106 97</td>
</tr>
<tr>
<td>UK</td>
<td>52</td>
<td>M</td>
<td>8</td>
<td>—</td>
<td>0</td>
<td>79 125 MD 194 MD 202 236 67</td>
</tr>
<tr>
<td>UK</td>
<td>46</td>
<td>M</td>
<td>20</td>
<td>Prednisone</td>
<td>10</td>
<td>43 51 65 74 MD 71 67 59</td>
</tr>
<tr>
<td>Italy</td>
<td>63</td>
<td>F</td>
<td>11</td>
<td>Prednisone</td>
<td>0</td>
<td>78 59 81 86 73 91 64 53</td>
</tr>
<tr>
<td>Italy</td>
<td>48</td>
<td>M</td>
<td>12</td>
<td>Prednisone, intravenous immunoglobulin</td>
<td>6</td>
<td>46 55 61 69 63 71 58 51</td>
</tr>
<tr>
<td>UK</td>
<td>64</td>
<td>F</td>
<td>10</td>
<td>Prednisone, intravenous immunoglobulin</td>
<td>6</td>
<td>56 104 135 104 148 225 77 47</td>
</tr>
</tbody>
</table>

F = female; ITP = idiopathic thrombocytopenic purpura; M = male; MD = Missing data; UK = United Kingdom.

*This patient was diagnosed with H. pylori reinfection, and was successfully re-treated with metronidazole-based eradication therapy.
IgG levels did not change significantly (399 ± 235 ng/10^7 platelets vs. 408 ± 251 ng/10^7 platelets, P = 0.46; Figure 2).

**Follow-up of responders**

Median follow-up among the 17 responders was 25 months (range, 7 to 42 months). Responses were sustained in 11 patients. Six patients experienced a relapse; all were retested for *H. pylori* infection and reinfection was diagnosed in 1. Successful re-treatment with eradication therapy resulted in a new sustained platelet response. The 3 patients with a partial response had a slow decline in the platelet count after peak values 3 to 6 months after eradication therapy.

**Follow-up of patients who did not receive eradication therapy**

Twelve patients had mild ITP (platelet count ≥50 × 10^9/L) and *H. pylori* infection, but no symptoms of dyspepsia. These patients did not receive eradication therapy, but were followed monthly to assess symptoms of dyspepsia or fluctuations of the platelet counts as indications for a trial of eradication therapy. During a median follow-up of 27 months (range, 8 to 41 months), no patient reported symptoms of dyspepsia or had a sustained improvement in platelet counts (>30 × 10^9/L). In 6 patients, the platelet counts dropped below 50 × 10^9/L during follow-up, but none of these changes lasted longer than 1 month.

**Discussion**

*H. pylori* colonization of the gastric mucosa in humans is common. Seropositivity for the bacterium is associated with increasing age and lower socioeconomic sta-

**Figure 2** Changes in platelet-associated immunoglobulin G (IgG) levels in patients with idiopathic thrombocytopenic purpura who had undergone eradication of *Helicobacter pylori* infection.
Acknowledgment

We are grateful to David Baker, Department of Medical Microbiology, Barts and London NHS Trust, for providing us with the U.K. *H. pylori* data.

References

EDITORIAL

Infectious agents and human immune diseases: Lessons from Helicobacter pylori

The search for new therapeutic approaches for the control and possibly the cure of chronic, self-perpetuating immune diseases is a challenging goal for clinicians and for clinical investigators. Included in this search is chronic immune thrombocytopenic purpura (ITP), an acquired bleeding disease, prevalent in adults, in whom autoantibodies bind to the platelet’s surface and cause their destruction in the reticulo-endothelial system. The mechanisms that trigger the production of platelet autoantibodies remain poorly understood. Unlike children, who often have a self-limiting disease, ITP in adults is usually a chronic disorder that is often refractory to standard therapy and is punctuated by frequent relapses and a high rate of complication. Current treatment protocols include immunosuppressive drugs, often chronically, and splenectomy.

In 1998, Akiyama and Onozawa reported a 2- to 3-fold increase in platelet counts in an elderly patient with chronic ITP who received omeprazole treatment for gastric ulcer. Subsequently, several studies have reported that the eradication of Helicobacter pylori by a proton pump inhibitor and antibiotics might be followed by variable improvement of the platelet counts in a fraction of infected patients with ITP. A careful review of the literature showed that among 750 patients with ITP from different countries, 65.5% tested positive for H. pylori infection; eradication was successful in 71.5% of these patients, and 53.5% of patients with successful eradication had either a complete or partial improvement of platelet counts. Of note, the majority of these studies were in relatively homogenous ethnic groups from Italy and Japan with rather homogeneous findings. Another study of pediatric chronic ITP patients in Taiwan showed a positive response in 55.5% of infected children.

In this issue of The American Journal of Medicine, Stasi and colleagues extend the previous findings by confirming the positive effects of eradicating H. pylori on platelet counts in a large cohort of patients with ITP, half of whom were from the United Kingdom. Consistent with a pathogenetic role of H. pylori in ITP, an improvement of platelet counts has not been observed in H. pylori–negative patients who have been treated with the same eradication regimen or in H. pylori–positive patients who have not had successful eradication. Some authors, including Stasi et al suggest that thrombocytopenia may be more responsive in patients with ITP of recent onset, with less severe disease (platelets >30 x 10⁹/L) and younger (<65 years). By comparison, others have reported that the duration of disease did not differ between responders and nonresponders; that a good response, varying from 40% to 73%, has been obtained in patients with platelet count below 30 x 10⁹/L, and that age did not predict response rates. One group has also reported a 100% response in patients aged >60 years, compared with 22% in those aged <50 years. Resolution of thrombocytopenia has also been reported in patients whose ITP was previously refractory to immunosuppression or splenectomy. Two groups of investigators, one from Spain and another from the United States, found no benefit of eradicating H. pylori in 130 patients.

The variability in the clinical benefits of eradicating H. pylori in patients from diverse countries and the differences in clinical features of responsive patients raise the issue of whether H. pylori infection causes the development or persistence of ITP. This issue is tremendously complicated because ITP is a more heterogeneous disease than initially suspected, with platelet destruction due to phagocytosis by macrophages, antibody-dependent cell-mediated cytotoxicity, and T-cell mediated destruction. Second, H. pylori shows extensive genetic diversity and variability due to frequent intraspecific recombination during mixed infection, potentially changing its interaction with the host.

How might H. pylori take part in the pathogenesis of ITP? Based on the observation that eradication of H. pylori can induce regression of B-cell gastric MALT lymphoma, platelet autoantibodies might be produced by autoreactive clonal B-cells that are induced by the chronic immunologic stimulus of H. pylori. Genetic factors of the host could be evoked in the pathogenesis of disease and in the susceptibility to infection. Differences have been reported in human leukocyte antigen (HLA) class II allele patterns between patients who have ITP with and without H. pylori infection, but rigorous studies on
specific HLA alleles in very large series of ITP patients have not yet been performed. Molecular studies of \textit{H. pylori} have identified strains of varying virulence. For example, the blood-group antigen-binding adhesion gene (babA), encoding adhesion (BabA) targeting human Lewis (Le) surface epitopes, is associated with gastric adenocarcinoma. Le antigens absorb to platelets, and their expression might contribute to the dysfunction of specific compartments of the cellular immune system of the host. Again, some strains of \textit{H. pylori} induce platelet activation mediated by \textit{H. pylori} bound to von Willebrand factor (vWF), interact with glycoprotein Ib (GPIb), and might contribute to platelet consumption.\textsuperscript{11} A study from Japan on the \textit{H. pylori}–induced alteration in T helper1 (Th1)–type immune response and cytokine profiles showed that the ratio of cells positive for intracellular interferon gamma to those positive for interleukin 4 (Th1/Th2 ratio) was significantly increased after eradication therapy, except in nonresponsive patients positive for interleukin 4 (Th1/Th2 ratio) was significantly increased after eradication therapy, except in nonresponsive patients with ITP.\textsuperscript{12} Another study from the same group found that platelet eluates from patients with ITP recognized \textit{H. pylori} cytotoxin–associated gene A (CagA) protein and that levels of anti-CagA antibody in eluates declined after eradication treatment in responsive patients. These results suggest that autoimmunity might be mediated by molecular mimicry between \textit{H. pylori} and platelet antigens,\textsuperscript{5} although a small French study could not confirm such a mechanism.\textsuperscript{13}

One investigation of the \textit{H. pylori} virulence profile—phenotyping for cagA, vacuolating cytotoxin gene A (vacA), and induced by contact with epithelium gene A (iceA)—showed no differences in strain expression between patients and controls (without ITP) with peptic ulcer or non-ulcer disease.\textsuperscript{8} In contrast, we found cagA, vacA, s1/m2, and iceA (“triple-positive” strains) in gastric biopsy specimens in 5 of 8 ITP patients with ITP who had asymptomatic \textit{H. pylori} gastritis (chronic or atrophic) but not in 8 control patients who had symptomatic \textit{H. pylori} gastritis but did not have ITP.

The above-mentioned studies seem to be the first step toward a better understanding of the pathogenetic mechanisms underlying the relationship between \textit{H. pylori} infection and ITP. Genotyping studies of \textit{H. pylori} strains and their association with a particular phenotype, as in the case of ulcer and gastric tumors, are worth pursuing. Such investigations should be performed using stringent methods approaches in large, prospective, randomized, placebo-controlled trials spread over different geographic areas.

The causative role of infectious agents in autoimmune diseases has been a subject of research and controversy for decades. The debate over the causative role of \textit{H. pylori} in ITP is reminiscent of the controversial association between viruses and autoimmune diseases, such as Epstein-Barr virus and Sjogren’s syndrome,\textsuperscript{14} human herpervirus-6 and multiple sclerosis,\textsuperscript{15} and parvovirus B 19 and rheumatoid arthritis.\textsuperscript{16} Nevertheless, the eradication of \textit{H. pylori} has proved to be of significant clinical value in some patients with ITP. Thus, the screening of all ITP patients for \textit{H. pylori} at diagnosis and the eradication of the bacterium in \textit{H. pylori}–positive patients should precede other specific treatment for ITP whenever possible. This strategy is a useful and inexpensive way to avoid the discomfort and side effects of immunosuppressive treatments and splenectomy in at least a subset of patients with ITP.

Giovanni Emilia, MD
Mario Luppi, MD
Giuseppe Torelli, MD

Department of Oncology and Hematology, Division of Hematology, University of Modena and Reggio Emilia, Modena, Italy

References
The effect of self monitoring of blood glucose concentrations on glycated hemoglobin levels in diabetic patients not taking insulin: A blinded, randomized trial

Mayer B. Davidson, MD, Maria Castellanos, RN, Don Kain, MS, MA, Petra Duran, BS

Clinical Trials Unit, Charles R. Drew University.

Self monitoring of blood glucose concentrations has the potential to lower glycemia and to decrease diabetic retinopathy, nephropathy and neuropathy. Self monitoring has been helpful in insulin-requiring patients because glycated hemoglobin (A1C) levels are inversely related to the frequency of glucose testing. In one study, increased frequency of self monitoring of blood glucose concentrations resulted in lower A1C levels only in those who self adjusted their insulin doses, not in the insulin-requiring patients who did not (strongly suggesting that acting on the values is necessary). The same cannot be said for patients on oral anti-diabetes drugs or diet alone. Most studies do not demonstrate any beneficial effect on glycemic control. In the few that do other factors (discussed below) influenced the results. However, glucose values in these patients could serve to educate and motivate them to adopt healthier life styles regarding diet and exercise.

There are three possible explanations for the lack of an effect of glucose monitoring in patients. First, patients receive little or no feedback on their results. Second, they are not taught the self-management skills to use to lower the measured glucose values. Third, the vast majority of patients measure their glucose level either fasting or preprandially, rather than postprandially. Fasting values serve neither to educate (there is no information on the effect of the meal composition or size) nor to motivate well (postprandial values are much higher).

There have been calls for blinded, randomized studies to answer the important question of whether self monitoring of blood glucose concentrations improves A1C responses. Therefore, a study in which the patients not taking insulin were randomized into a monitoring group or a control group with the provider (a specially trained nurse following detailed glycemic management algorithms) making the clinical management decisions blinded to the group assignment was carried out in a Diabetes Managed Care Program (DMCP) at a community clinic associated with King-Drew Medical Center.

Methods

All type 2 diabetic patients not taking insulin who were either currently enrolled or on entrance into the DMCP were approached to enter the study. All 89 patients consented and were randomized. One patient did not return after being randomized to see the nurse or the dietitian (who provided nutritional counseling to all patients enrolled in the DMCP) and was not included in the study. All patients were followed for 6 months. Patients were instructed to measure glucose levels before and between 1 and 2 hours after eating meals 6 days a week; 2 breakfasts, 2 lunches, and 2 suppers, and to record what they ate at those meals. A second value within 30 minutes of the first was considered a “duplicate” (probably either to confirm the preceding one or to assess the response to treatment of possible hypoglycemia). Only the first one was counted. Patients in both groups were scheduled to meet with the dietitian 5 times; at randomization and 2, 4, 8, and 12 weeks later. The dietitian utilized the glucose values and the meal descriptions in his nutritional

This study was supported by grants from Eli Lilly and Company and the NIH (DK 54047).

Requests for reprints should be addressed to Mayer B. Davidson, MD, Clinical Trials Unit, Charles R. Drew University, 1731 East 120th Street, Los Angeles, California 90059.

E-mail address: madavids@cdrewu.edu.
counseling to educate the patient on the effects of the meal components and portion sizes on the rise of postprandial glucose concentrations.

The nurse, who was unaware of whether the patient was randomized to the monitoring group or not, followed detailed algorithms to make her therapeutic decisions. The first goal was to lower the fasting plasma glucose concentrations to <130 mg/dL by stepwise increases of metformin or a sulfonylurea agent every 2 weeks until either that goal was attained or a maximum dose of the oral drug was reached. If the goal was not met with a maximal dose of one of these drugs, the second one was added with subsequent stepwise increases until the goal was reached. Once the fasting goal was attained, A1C levels were measured every 2 months with a goal value of <7.5%. (This study was started at a time when the American Diabetes Association’s guideline for A1C levels was to “suggest action” if the value were >8.0%.) If maximal (tolerated) doses of both metformin and a sulfonylurea agent were prescribed without reaching the 7.5% A1C goal, a maximal dose of a thiazolidinedione was added (triple oral therapy).

A1C levels were measured at entry into the study and every 2 months. If the A1C level were >7.5% 4 months after adding a maximal dose of the thiazolidinedione or at subsequent 2-month intervals, the thiazolidinedione was stopped and bedtime NPH insulin was started. In the 10 patients who failed triple oral therapy, the A1C value just before initiating insulin was used in the statistical analysis. A final A1C level was measured at the end of the study in all of the 78 patients remaining on oral antidiabetes drugs without insulin.

An intention to treat analysis was used. Paired and unpaired t tests were carried out as appropriate on continuous data and a chi-squared analysis on dichotomous data with statistical significance accepted at the 0.05 level (2-tailed). All data are presented as means ± SD. The study was approved by the Drew Institutional Review Board.

### Results

There were no differences in the baseline characteristics of the patients randomized to the monitoring group and those who were randomized to the control group (Table 1). Patients in the monitoring group averaged 4.0 ± 1.0 dietary visits compared with 3.2 ± 0.9 visits in the control group. The average number of tests in patients who monitored was 129. There were no changes in weight or body mass index during the study in either group (Table 2). Medications at the end of the study were similar in both groups of patients, indicating that the two were treated similarly by the nurse. Although final A1C levels were similar after 6 months in patients randomized to the monitoring group (7.7% ± 1.6) and those in the control group (7.8% ± 1.5), they fell significantly in both (monitoring group P <0.001; control group P = 0.05). The decrease, however, was not significantly different between the monitoring group (−0.8%) and the control group (−0.6%). The 95% confidence interval of the change in A1C levels was −1.1 to +0.6%.

### Discussion

The results of this randomized, single blind study in type 2 diabetic patients not taking insulin failed to demonstrate lower A1C levels in those that perform pre- and postprandial glucose testing. This lack of an effect is consistent with 11 other studies as well as a meta-analysis of 6 of the randomized ones. Four studies did show an effect. In one, a “therapy decision scheme” was applied only to the group that tested but not to the control group, casting doubt on whether testing per se was responsible for the difference. In two studies, A1C levels were lower in patients who opted to perform glucose testing. However, self selection could conceivably explain these results. This interpretation is supported by the results of a self-administered questionnaire or a computer-assisted telephone interview given to the plan members in one of the studies, 83% of whom responded. Self-care practices and healthy lifestyle behaviors were more common in individuals who self monitored more frequently.

In the final study, A1C levels fell 1.0% in 113 patients randomized to monitor before and after main meals twice a week compared with 0.5% in 110 control patients (P <0.01) over 24 weeks. Patients were seen every 4 weeks with nurses. However, an eating diary and a structured counseling program were provided only to the monitoring group. Furthermore, there was no information concerning who.

| Table 1 Baseline characteristics* of patients randomized to monitoring and controls |
|-----------------------------------|------------------|------------------|
|                                  | Monitoring group | Control group     |
|                                  | (n = 43)         | (n = 45)         |
| Women (number [%])               | 34 (79)          | 31 (67)          |
| Age (years)                      | 50.9 ± 11.0      | 49.8 ± 11.2      |
| Weight (kg)                      | 83.9 ± 23.3      | 80.7 ± 17.0      |
| Body mass index (kg/m²)          | 33.4 ± 7.0       | 31.7 ± 6.7       |
| Duration of diabetes (years)     | 5.8 ± 5.8        | 5.5 ± 4.7        |
| Ethnicity (number [%])           |                  |                  |
| African-American                 | 9 (21)           | 10 (22)          |
| Latino                           | 32 (74)          | 34 (76)          |
| Other                            | 2 (5)            | 1 (2)            |
| Medications (number [%]))        |                  |                  |
| None (diet alone)                | 0 (0)            | 2 (4)            |
| Metformin alone                  | 12 (28)          | 12 (27)          |
| Sulfonylurea† alone              | 0 (0)            | 4 (9)            |
| Metformin plus sulfonylurea       | 22 (51)          | 24 (53)          |
| Triple oral therapy‡             | 9 (21)           | 3 (7)            |
| A1C levels (%)                   | 8.5 ± 2.2        | 8.4 ± 2.1        |

*There are no significant differences between the two groups.
†glyburide or glipizide.
‡metformin, sulfonylurea, and thiazolidinedione.
made therapeutic decisions during the 6-month period and whether that provider knew the monitoring status of the patient. Hence, the present study may be the only randomized and truly blinded one to address the question of whether glucose monitoring in type 2 diabetic patients not taking insulin leads to improved glycemia.

Glucose monitoring is expensive. Although it is not possible to completely isolate monitoring costs for diabetic patients not taking insulin, the Medicare B fee-for-service program affords a fairly accurate estimate of this cost. The ICD-9 code 250.00 (type 2 diabetes, uncomplicated, not uncontrolled) is the one most often used for diabetic patients on either diet alone or taking oral antidiabetes medications. The total cost in 2002 for reagent strips, lancets, lancing devices, meters, batteries, and calibration solutions or chips was $465 500 000, which represented 58.8% of the total outlay of the Medicare B program for the ICD-9 code of 250.00 (personal communication, staff, Center for Medicare & Medicaid Services). If type 2 diabetic patients receiving insulin were given this ICD-9 code, this cost would be an underestimate. Conversely, if type 2 diabetic patients not taking insulin were given another ICD-9 code, this cost would be an overestimate. However, because this cost does not include the 10% of Medicare beneficiaries enrolled in HMO Managed Medicare, this figure is certainly an underestimate of the total cost for glucose monitoring in type 2 diabetic Medicare patients not taking insulin. Given that this nearly one-half billion dollars is only for Medicare patients, the total cost for glucose monitoring for all type 2 diabetic patients not taking insulin is obviously much higher.

Because of the wide 95% confidence interval in the change of A1C levels, a possible limitation of this study is that a significant effect may have been missed. A second limitation is the possibility that this mostly poor and poorly educated minority population might have been unable to limitation is the possibility that this mostly poor and poorly educated minority population might have been unable to take advantage of the information afforded by the nutritional counseling that utilized pre-and postprandial glucose values. These negative findings occurred in a practice in which intensive follow-up and treatment are the norm, indicating that monitoring does not improve A1C levels in this setting. However, the lack of a clear-cut beneficial effect in the published literature suggests that monitoring may not be helpful in other practice settings either. That this blinded, randomized study could not show any beneficial glycemic effects in type 2 diabetic patients not taking insulin is consistent with the conclusions of several reviews and a meta-analysis. Thus, there is no convincing evidence that self monitoring of blood glucose concentrations in type 2 diabetic patients not taking insulin leads to better glycemic outcomes.

Acknowledgments

The authors gratefully acknowledge Lifescan, Inc., for donating glucose meters and strips and the statistical help of Mohsen Bazargan, PhD.

References


Intravenous cyclophosphamide in refractory polyneuropathy associated with IgM monoclonal gammopathy: An uncontrolled open trial

Mohamed A. Hamidou, MD, Cristina Belizna, MD, Sandrine Wiertlewsky, MD, Marie Audrain, MD, Charlotte Biron, MD, Jean-Yves Grolleau, MD, Jean-Marie Mussini, MD

Monoclonal gammopathy-associated demyelinating neuropathy is a common and probably under-recognized peripheral neuropathy in elderly patients. The presence of serum anti-myelin-associated glycoprotein antibodies, immunoglobulin M (IgM) and complement deposits on the myelin sheaths, as well as animal models, favor an immune-mediated disease supporting the use of immune therapies. Controlled randomized clinical trials and open pilot studies did not demonstrate a great benefit of different therapeutic modalities and chlorambucil is still the standard first line of treatment for symptomatic patients. Cyclophosphamide has proven its efficiency in various clinical immunologic situations. We therefore performed an open prospective study to evaluate the efficacy and safety of intravenous cyclophosphamide pulses. The diagnosis of monoclonal gammopathy of undetermined significance was established from the presence of a serum low grade monoclonal IgM, no clinical evidence of organ enlargement, and normality of clinical examination, hemogram, CT thoracic and abdominal scan, and bone marrow biopsy. The electrophysiologic examination included the study of the nerve conduction of peroneal, tibial, ulnar and median nerves (motor and sensory nerve conduction velocity, distal latency, amplitude, terminal latency index) and the study of sensory conductions. Demyelinating peripheral neuropathy was assessed by electromyography, with symmetrical reduction of motor and sensory nerve conduction velocities, and prolonged distal motor latencies, without conduction blocks. Sural nerve biopsy showed a pattern of segmental demyelination with typical widening of the myelin lamellae, without inflammatory infiltrates; it excluded other causes of peripheral neuropathy. Anti-myelin-associated glycoprotein antibodies were positive in the 9 patients based on enzyme-linked immunosorbent assay testing (Bühlmann Laboratories, Allschwil, Switzerland). Tests for human immunodeficiency virus, hepatitis B and C virus, antinuclear, antineutrophilic cytoplasmic antibodies, and cryoglobulinemia were negative. Patients with clinical or laboratory evidence of other cause of neuropathy were excluded. The mean motor nerve conduction velocity (ulnar nerve) was 26

Requests for reprints should be addressed to Mohamed Hamidou, MD, Department of Internal Medicine, Hôtel-Dieu, Place Alexis Ricordeau, Nantes 44 035, France.
E-mail address: mohamed.hamidou@chu-nantes.fr

0002-9343/$ -see front matter © 2005 Elsevier Inc. All rights reserved.
meters per second. Strength decreased progressively in all patients during the 12 months before cyclophosphamide, with worsening disability and impairment of daily life. After oral patient information, cyclophosphamide was administered intravenously at 600 mg/m^2 of body surface every 4 weeks for 6 months, and subsequently every 8 weeks during 12 months. Antiemetic olasetron therapy and the medication mesna to prevent hemorrhagic cystitis were used. No patient received corticosteroids.

Strength was assessed by analytical testing using a summated Medical Research Council strength (0–5) score in 9 muscles as previously described.4,5 Functional status was assessed using the modified Rankin disability scale: 0 = no symptoms at all, 1 = nondisabling symptoms that do not interfere with daily activities, 2 = minor disability, unable to carry out all activities, but able to walk without assistance; 3 = moderate disability, requiring some help, but able to walk without assistance and unable to attend to own bodily needs without assistance; and 5 = severe disability, totally dependent, requiring constant nursing care, 6 = death. The evaluation of the clinical status was done on the basis of the clinical score (Rankin scale and Medical Research Council strength). The clinical assessment, quantification of serum IgM, detection of anti-myelin associated glycoprotein antibodies and electromyography, were performed before inclusion in the study, at 6 and 18 months for a majority of the patients. Improvement was defined as an increase of at least 1 point in the Rankin disability scale. Stabilization was defined as no change on Rankin scale after 6 courses of therapy. Progression of the disease was defined as a deterioration of Rankin scale.

Statistical analysis

Data are expressed as the mean ± SD. Baseline, 6, and 18 months values were compared with a paired Student’s t test. P values <0.05 were considered statistically significant.

Results

Baseline characteristics

The characteristics of patients are shown in Table 1. Patients included were 7 men and 2 women, with a mean age of 63 years and a mean evolution time of neuropathy of 3.5 years before cyclophosphamide treatment. All 9 patients received at least a 6-month course of 0.1 mg/kg/day oral chlorambucil and a 2-month dose of twice-a-week plasma exchange. At the time of inclusion, the 9 patients had sensory and motor neuropathy, and were functionally severely disabled. Therefore, 4 patients had a Rankin score of 4, 4 had a score of 3, and 1 had a score of 5 (Table 2). Before treatment, 7 patients needed devices, 4 a walking cane, and 3 a walking cane and ankle braces.

Clinical response to therapy

With the response criteria defined above, 7 patients had improvement in strength and functional status, with a decrease of at least 1 point on the Rankin scale (Table 2). Two patients had disability with stabilization of the symptoms at the 18-month end-point. At month 18, only 2 patients needed a walking cane and ankle braces.

Biological and electrophysiological responses to treatment

Despite clinical improvement, there were no significant changes in the electrophysiologic measures and antimyelin-associated glycoprotein antibodies. Total IgM values were reduced in all the patients, and for 3 patients, they fell below the normal range (0.5 to 1.8 g/L). No side effect was noticed in our group. The mean duration of the follow-up since the end of treatment was 24 months (range 18 to 30 months), and no patient actually relapsed.

Discussion

We showed, in a group of patients with refractory sensorimotor neuropathy associated with monoclonal IgM monoclonal gammopathy of undetermined significance and antimyelin-associated glycoprotein antibodies, the benefits of intravenous pulse cyclophosphamide. A clinical improvement occurred in 7 of 9 cases, and stabilization was noted in 2 cases with a mean follow-up of 28 months. No correlation was found with electromyography and antibodies,3,8,9,17 whereas the M protein decreased significantly in the 9 patients.

Current standard therapy of IgM neuropathy is mainly based on chlorambucil, despite only a 30% response rate.
Table 2  Assessment parameters at baseline, 6, and 18 months of cyclophosphamide treatment

<table>
<thead>
<tr>
<th>Patient</th>
<th>Rankin scale</th>
<th>Motor serve conduction velocity (meters by second)</th>
<th>Immunoglobulin M (mg/L)</th>
<th>Anti-myelin-associated glycoprotein antibodies (Bülhmann titer units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Month 0</td>
<td>Month 6</td>
<td>Month 18</td>
<td>Month 0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>76</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>78</td>
</tr>
<tr>
<td>Mean</td>
<td>3.66</td>
<td>2.33</td>
<td>2.22</td>
<td>77.1</td>
</tr>
<tr>
<td>± SD</td>
<td>(0.7)</td>
<td>(0.5)</td>
<td>(0.44)</td>
<td>(3.88)</td>
</tr>
<tr>
<td>P</td>
<td>0.004</td>
<td>0.003</td>
<td>0.007</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Rankin scale: 0: no significant disability to 6: death.
Muscle strength as assessed by a summated Medical Research Council strength score (0–5) of 9 muscles (maximum score: 90).
in an open prospective study. Plasma exchanges are a commonly used option, even though steroids, intravenous immunoglobulins, interferon alpha, and fludarabine showed no significant response rate. Consequently, these poor results argue for the development of new therapeutic strategies, and the monoclonal humzized anti-CD20 antibody rituximab presents promising perspectives.

Cyclophosphamide has not often been evaluated in the monoclonal gammopathy-associated immune neuropathy, and the literature data are controversial. Gorson et al showed that 1000 mg/m² intravenous cyclophosphamide monotherapy (6 patients) or combined with plasma exchange (8 patients) was not more effective than other treatment regimens. Blume et al showed an improvement in 4 patients, with a combined protocol including a cyclic administration of 1000 mg/m² cyclophosphamide and plasma exchanges during 5 to 7 months. In a prospective study, Notermans et al demonstrated, in 16 patients, the benefit of 6 months intermittent oral cyclophosphamide, with a follow-up of 36 months. Nobile-Orazio et al did not draw the same conclusion in 7 patients, as only 2 had a partial response.

Our study is a monocentric prospective open label trial, testing long duration (18 months) intravenous cyclophosphamide monotherapy in refractory patients to chlorambucil and plasmapheresis. It was homogeneous, in regard to the inclusion criteria and treatment schedule. In contrast to other trials in the literature, this monotherapy protocol allowed us to avoid the interferences of associated treatments. Therefore, it tested the specific effect and tolerance of intravenous cyclophosphamide. Intravenous pulse cyclophosphamide regimen is commonly used in systemic immune diseases, showing fewer side effects than oral continuous administration. Thus, we did not record any side effects, even in the older patients. The ideal duration of treatment is unknown, and we arbitrarily chose a long duration. However, our comparative evaluation at 6 and 18 months did not show significant modifications at these 2 end-points. We wonder if, as concluded by Notermans et al, it is necessary to extend immunosuppressive therapy for more than 6 months. On the other hand, it is interesting to note that our 7200-mg/m² total cumulative intravenous cyclophosphamide dose for 18 months was equivalent to the dose administered in Notermans’ study, but with a different design (300 mg/m² oral cyclophosphamide for 4 days during 6 months). Our trial is an uncontrolled open-label study and therefore a placebo effect, or regression to the mean, could not be excluded. However, all the patients had 2 lines of treatment before cyclophosphamide without any benefit. Classically, anti-myelin-associated glycoprotein antibodies-associated monoclonal IgM neuropathy had a slowly progressive course and spontaneous regressions are exceptional.

This study is limited by its small sample size and the lack of randomization of a control group. Nevertheless, this prospective study shows that intravenous cyclophosphamide without steroids is safe, low cost, and short-term effective in patients with refractory demyelinating sensorymotor polyneuropathies associated with anti-myelin-associated glycoprotein IgM monoclonal gammopathy of undetermined significance.

References


BRIEF OBSERVATION

Dispensing of proton pump inhibitor medication is independently associated with subsequent asthma emergency hospital utilization

Michael Schatz, MD, MS,a Peter T. Chen, MS,b Eric Macy, MD,a Robert S. Zeiger, MD, PhD

Departments of Allergy and Pharmacy Analytic Services, Kaiser-Permanente Medical Care Program.

Patients with asthma frequently have nonrespiratory comorbid conditions, such as gastroesophageal reflux disease,1 cardiovascular disease,2,3 and anxiety-depression.4,5 Some of these comorbid conditions can directly exacerbate asthma, such as gastroesophageal reflux disease1 and anxiety/depression,4,5 whereas the treatment of others (eg, beta-blockers and possibly angiotensin-converting enzyme inhibitors) could make asthma worse. The relationships of these comorbid conditions to subsequent emergency department visits or hospitalizations for asthma have been incompletely studied, especially as independent risk factors. The purpose of this study was to evaluate the relationship of common treatment-determined comorbid conditions to subsequent emergency department or hospital utilization for asthma, accounting for other relevant risk factors.

Methods

This study was approved by the Southern California Kaiser Permanente Institutional Review Board. Patients were identified from computerized administrative data as having asthma if they met one or more of the following criteria:

1) Any discharge diagnosis (principal or other diagnosis) of asthma in the hospitalization database (ICD-9 code: 493.xx);
2) Two or more asthma-related medication dispensings (excluding oral steroids) in a 1-year period in the prescription database, including beta-agonists (excluding oral terbutaline), inhaled steroids, other inhaled anti-inflammatory drugs, and oral leukotriene modifiers;
3) Any emergency department or regular clinic asthma-related visit in the outpatient diagnosis and procedures database.

Subjects in the present cohort were patients included in the asthma clinical identification database during 1999 who were aged 18–64 years, were assigned to the San Diego medical center, were continuously enrolled (no breaks in membership >2 months) in 1999 and 2000, and had prescription drug coverage as a benefit. To identify a purely asthmatic sample, potential subjects with a diagnosis of chronic obstructive pulmonary disease (ICD-9 codes 491.xx, 492.xx, or 496.xx) were excluded from the analysis. The outcome was emergency hospital care for asthma during the year 2000, defined as at least one hospitalization or emergency department visit for asthma (ICD-9 code 493.xx), expressed as a yes/no variable.

The predictors of interest were treatment-defined comorbid conditions identified from the prescription database during 1999: 1) acid gastrointestinal (H2 blockers, proton pump inhibitors), 2) cardiovascular (beta-blockers, ACE-inhibitors, other), 3) diabetes (oral agents, insulin), and 4) depression (SSRIs, tricyclics, other).

All analyses were conducted using SAS statistical software (SAS version 8.2 for Windows; SAS Institute, Inc, Cary, NC). Hypothesis testing for binary variables in unad-
justed analyses was by means of chi-square analysis. Mantel-Haenszel chi-square was used to test for a significant linear trend. Multivariable analyses were conducted using multiple logistic regression analyses, adjusting for sex, subsidized insurance, asthma emergency hospital care in the prior year, and number of beta agonist canisters, inhaled corticosteroid canisters, and oral corticosteroid dispensions in the prior year. Hypothesis testing in logistic regression models was by means of Wald chi-square. A 2-sided $P$ value of $<0.05$ was considered statistically significant in univariate and multivariable analyses.

### Results

The cohort included 6310 subjects, 21 (0.3%) of whom were hospitalized for asthma during 2000, 229 (3.6%) of whom required an asthma emergency department visit during 2000, and 240 (3.8%) of whom required emergency hospital care in the form of at least one emergency department visit or hospitalization during 2000.

The characteristics of the sample are shown in Table 1. Treatment-defined acid gastrointestinal disease, cardiovascular disease, and depression were fairly common (17–23%), whereas diabetes was less common (~5.0%). Acid gastrointestinal disease was associated with a significant increase in emergency hospital utilization, which appeared to be specifically due to proton pump inhibitor treatment (Table 2). The relationship to protein pump inhibitor treatment persisted after adjustment for potential confounders (Table 2). The population attributable fraction, calculated as the rate of disease in the total sample minus the rate in the unexposed group divided by the rate in the unexposed group, was 3.8%. There were no significant relationships between any of the other treatment-defined comorbid conditions or specific drug classes and subsequent asthma emergency hospital care in crude or adjusted analyses.

The median number of dispensions of proton pump inhibitors in adults who used them in 1999 was 3 (range 1–21). A significant linear relationship ($P <0.006$) was found between asthma emergency hospital care and proton pump inhibitor use, trichotomized into none (3.7% of 6003), 1–3 (5.6% of 180), and $>3$ dispensions (7.9% of 127).

### Discussion

Prior studies have identified a number of risk factors for asthma emergency hospital care, including demographic, socioeconomic, clinical (eg, asthma severity), laboratory (eg, atopy), utilization, and asthma medication factors. This is the first study to show that treatment with a specific class of nonrespiratory medications, proton pump inhibitors, increases the risk of subsequent asthma emergency hospital care. Strengths of this study include the large number of asthmatic patients managed by a single integrated health care system, objective documentation of medication dispensing, ability to document a dose-response relationship, and the control for potential confounding demographic, socioeconomic, prior utilization, and asthma severity and treatment factors.

The most plausible mechanism for the association is that proton pump inhibitor dispensing identifies patients with substantial gastroesophageal reflux disease that may aggravate their asthma and increase their risk of severe episodes. This mechanism is supported by the findings of Shireman et al, who reported that gastroesophageal reflux disease independently increased the risk of subsequent hospitalizations for asthma. Proton pump inhibitors are generally used in our health care program according to specific gastroesophageal reflux disease guidelines when $H_2$ blocker treatment fails. Because $H_2$ blocker treatment alone was not associated with an increased risk, the current findings suggest that it is patients with more severe gastroesophageal reflux disease who are at risk. An alternative explanation is that reflux is suspected in patients with more severe asthma, and trials of proton pump inhibitor therapy are undertaken in such patients who fail to improve with $H_2$ blockers. Additional studies of asthmatic patients with potential gastroesophageal reflux disease should be performed to more rigorously assess the incidence of reflux under-treatment and its effect on asthma control. Although a direct adverse effect of proton pump inhibitors on the course of asthma

### Table 1 Characteristics of the study sample ($n = 6310$)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999 emergency department</td>
<td>379 (6.0)</td>
</tr>
<tr>
<td>visit/hospitalization</td>
<td></td>
</tr>
<tr>
<td>2000 emergency department</td>
<td>240 (3.8)</td>
</tr>
<tr>
<td>visit/hospitalization</td>
<td></td>
</tr>
<tr>
<td>Female sex</td>
<td>4038 (64.0)</td>
</tr>
<tr>
<td>Subsidized insured</td>
<td>151 (2.4)</td>
</tr>
<tr>
<td>Acid gastrointestinal</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>1094 (17.3)</td>
</tr>
<tr>
<td>Histamine$_2$ blockers only</td>
<td>787 (12.5)</td>
</tr>
<tr>
<td>Proton pump inhibitors</td>
<td>307 (4.9)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>1481 (23.5)</td>
</tr>
<tr>
<td>Beta-blockers</td>
<td>266 (4.2)*</td>
</tr>
<tr>
<td>ACE-Inhibitors</td>
<td>595 (9.4)</td>
</tr>
<tr>
<td>Other</td>
<td>708 (11.2)</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>305 (4.8)</td>
</tr>
<tr>
<td>Insulin</td>
<td>125 (2.0)</td>
</tr>
<tr>
<td>Oral agents</td>
<td>254 (4.0)</td>
</tr>
<tr>
<td>Depression</td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>1333 (21.1)</td>
</tr>
<tr>
<td>SSRIs</td>
<td>787 (13.9)</td>
</tr>
<tr>
<td>Tricyclics</td>
<td>357 (5.7)</td>
</tr>
<tr>
<td>Other</td>
<td>481 (7.6)</td>
</tr>
</tbody>
</table>

ACE-Inhibitors = angiotensin-converting enzyme inhibitors; SSRIs = Selective serotonin reuptake inhibitors. *221 selective, 48 non-selective, 3 both
cannot be excluded from the present data, there are no prior studies or likely mechanisms to support this hypothesis.

Another explanation for the association identified in this study is unmeasured confounding. Cigarette smoking has been shown to be a risk factor for emergency hospital care for asthma in one study\(^23\) and has also been associated with gastroesophageal reflux disease diagnosis or symptoms in some\(^24-28\) but not other\(^29-31\) populations. Because cigarette smoking could not be captured in this administrative data-base study, it cannot be excluded as a confounder of the relationship between proton pump inhibitor treatment and asthma emergency hospital care demonstrated in this study.

This study did account for another potential confounder, oral corticosteroid use, which has been shown to be an independent risk factor for asthma emergency hospital care\(^6,7,13,19\) and has been associated with gastroesophageal reflux disease.\(^32\) Obesity has been associated with both asthma\(^33,34\) and gastroesophageal reflux disease,\(^25,28\) but it would not qualify as a confounder in this study, because it has not been shown to be independently related to the outcome (emergency hospital care for asthma).\(^35\)

Risk stratification is used to identify patients at increased risk of subsequent adverse outcomes who would be candidates for targeted intervention.\(^7,13,36,37\) Whatever the mechanism of the association in the current study, these findings suggest that proton pump inhibitor treatment is an independent marker of increased risk that can be used in population management to identify patients at increased risk of subsequent asthma emergency hospital care. However, the population attributable fraction does appear to be quite small (3.8%).

The current study does have some potential limitations. Dispensing of medication does not automatically equate to use of medication, although pharmacy records are recommended as a method of assessing patient adherence.\(^38,39\) The use of proton pump inhibitors could have been for peptic ulcer disease, rather than gastroesophageal reflux disease, but gastroesophageal reflux disease is much more common than peptic ulcer disease.\(^40\) This study did not capture unscheduled outpatient visits other than emergency department visits, although emergency department visits would be the most expensive and urgent type of outpatient unscheduled asthma visits.

An important additional potential limitation of this study is that the comorbid conditions were identified based on pharmacy record-defined treatment, rather than using administrative diagnostic coding. This would be especially relevant for gastroesophageal reflux, for which a positive association was identified. Unfortunately, the diagnostic coding information available in our administrative data system in 1999 was not sufficiently accurate for the current purpose, especially compared with computerized pharmacy records. In this regard, many other health care systems also likely have computerized pharmacy records that are more

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>Exposed 2000 emergency department visit or hospitalization (%)</th>
<th>Unexposed 2000 emergency department visit or hospitalization (%)</th>
<th>Odds ratio (95 % CI)*</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid gastrointestinal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>55/1094 (5.0)</td>
<td>185/5216 (3.6)</td>
<td>1.3 (1.0–1.8)</td>
<td>0.07</td>
</tr>
<tr>
<td>(H_2) blockers only</td>
<td>35/787 (4.5)</td>
<td>205/5523 (3.7)</td>
<td>1.1 (0.8–1.6)</td>
<td>0.54</td>
</tr>
<tr>
<td>PPIs</td>
<td>20/307 (6.5)</td>
<td>220/6003 (3.7)</td>
<td>1.7† (1.1–2.8)</td>
<td>0.03</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>61/1481 (4.1)</td>
<td>179/4829 (3.7)</td>
<td>1.3 (1.0–1.7)</td>
<td>0.39</td>
</tr>
<tr>
<td>Beta-blockers</td>
<td>14/266 (5.3)</td>
<td>226/6044 (3.7)</td>
<td>1.5 (0.9–2.6)</td>
<td>0.16</td>
</tr>
<tr>
<td>ACE inhibitors</td>
<td>19/595 (3.2)</td>
<td>221/5715 (3.9)</td>
<td>0.9 (0.5–1.4)</td>
<td>0.50</td>
</tr>
<tr>
<td>Other</td>
<td>31/708 (4.4)</td>
<td>209/5602 (3.7)</td>
<td>1.2 (0.8–1.8)</td>
<td>0.38</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>8/305 (2.6)</td>
<td>232/6005 (3.9)</td>
<td>0.6 (0.3–1.2)</td>
<td>0.15</td>
</tr>
<tr>
<td>Insulin</td>
<td>4/125 (3.2)</td>
<td>236/6185 (3.8)</td>
<td>0.7 (0.3–2.0)</td>
<td>0.55</td>
</tr>
<tr>
<td>Oral agents</td>
<td>8/254 (3.2)</td>
<td>232/6056 (3.8)</td>
<td>0.7 (0.4–1.5)</td>
<td>0.38</td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any</td>
<td>52/1333 (3.9)</td>
<td>188/4977 (3.8)</td>
<td>0.9 (0.7–1.3)</td>
<td>0.72</td>
</tr>
<tr>
<td>SSRI</td>
<td>36/878 (4.1)</td>
<td>204/5432 (3.8)</td>
<td>1.0 (0.7–1.4)</td>
<td>0.92</td>
</tr>
<tr>
<td>Tricyclics</td>
<td>12/357 (3.4)</td>
<td>228/5953 (3.8)</td>
<td>0.8 (0.4–1.4)</td>
<td>0.39</td>
</tr>
<tr>
<td>Other</td>
<td>24/481 (5.0)</td>
<td>216/5829 (3.7)</td>
<td>1.3 (0.9–2.1)</td>
<td>0.19</td>
</tr>
</tbody>
</table>

\(H_2\) = histamine receptor 2; PPIs = proton pump inhibitors; ACE-inhibitors = angiotensin-converting enzyme inhibitors; SSRI = selective serotonin reuptake inhibitors.

*Adjusted for sex, subsidized insurance, asthma emergency hospital care in 1999, and number of beta agonist canisters, inhaled corticosteroid canisters and oral corticosteroid dispensings in 1999.

†Above and beyond the significant effects of 1999 emergency hospital care (odds ratio = 3.6, 95% CI 2.5–5.2), oral corticosteroid use (odds ratio = 1.1, 95% CI 1.0–1.2) and inhaled corticosteroid use (odds ratio 0.6, 95% CI 0.4–0.9).
accurate than their administrative diagnostic coding information, and the current findings should be very applicable to those health care systems.

References

Being a physician is rewarding but also challenging in the complex health care system. As physicians, we are continually trying to deliver more effective and higher quality care to our patients. With improvement in mind, a list of precepts has been generated as a tool to remind all of us in clinical medicine about the exemplary characteristics, behaviors, and attitudes that are expected as the norm in this profession. The list is organized into four categories: promotion of relationships with patients, principles of the effective clinician, growth and improvement, and values to guide one’s career in medicine. The list is envisioned as an instrument that may be helpful to medical learners and physicians by promoting reflection about ways to consistently perform at a high level while more fully appreciating the joy of practicing medicine. The list of precepts may also be useful to medical educators who wish to successfully mold the physicians of tomorrow.

© 2005 Elsevier Inc. All rights reserved.

The practice of medicine is a journey. As the pace of the travel increases, it is often difficult to focus on anything other than the completion of tasks. Outcomes, such as the number of patients seen, are often given the highest priority. Attention to process and matters such as the relationships between patients and their physicians are frequently overlooked. Resentment and burnout among physicians have become prevalent.1-3

The Hippocratic oath stresses the sharing of precepts to “pupils who have signed the covenant and have taken an oath according to the medical law”.4 Contemporary governing bodies, like the American Board of Internal Medicine and the Accreditation Council for Graduate Medical Education, also emphasize the attainment of competence, if not excellence, for physicians in a myriad of domains—including professionalism and communication skills.5

Although much has been written to guide the professional behaviors and attitudes of medical learners and practicing physicians,5-9 these documents are lengthy and not convenient for daily reference or reflection. Moreover, authors have focused on specific aspects of professionalism or ethics rather than addressing precepts that guide a patient-centered approach to patient care or that emphasize the importance of a fulfilling professional life. With this in mind, a user-friendly list of 52 precepts is presented herein. Our culture emphasizes “achieving your full potential” and “doing things better today than yesterday.” Not surprisingly, “self help books” are routinely found atop best-seller lists.10 The following manuscript is offered as a self-help piece for medical students and physicians. It is hoped that these precepts will allow each of us to provide better care for our patients and to be more fulfilled, more compassionate, and more balanced physicians.

The list of precepts that follows has been elaborated with input from well-regarded physicians, teachers, and mentors from the Johns Hopkins University School of Medicine. The list has been elaborated to capture the behaviors, at-
tributes, and attitudes of respected physicians. It may serve as a stimulus to foster self-assessment and may allow for the identification of areas for self-improvement. The precepts have been developed with the hope that they will be applicable to a wide range of individuals—from first-year medical students to seasoned physicians. The teachings cover a broad array of subjects including how physicians should relate to patients, refine their clinical skills, strive for personal growth, and appreciate the joy of practicing medicine. As you look through the list, it is likely that there may be some lessons that don’t inspire you at this particular point in time. This collection is offered as a list that has been helpful to us as physicians and educators through different stages and phases of our careers.

Development of the list

Literature search

Previous publications that have been written about becoming an excellent physician consist almost exclusively of opinion pieces wherein experts share their insights. In many of these, authors have focused on one aspect of medicine, such as communication skills or diagnostic considerations. Although our initial intent was to write an evidence-based review related to this topic, the real absence of significant empiric work in this area precluded this possibility.

Brainstorming

The authors met regularly over the course of 6 months to identify and distill the precepts that would be proposed to be included on the final list of the 52 precepts. Although a variety of techniques were used at the meetings, the authors attempted to reflect on teachings from their mentors and to document the “pearls” that were emphasized by their colleagues that seemed to be most relevant. The authors also engaged in dialogue about the precepts that seemed to be most appreciated by learners with whom they had worked in the prior year. Based on these discussions, a list of 46 precepts was developed.

Feedback from award-winning clinicians and educators

The list of 46 precepts was then sent to 15 respected faculty colleagues at the Johns Hopkins University School of Medicine. These faculty members were selected because they had won awards recognizing their excellence as clinicians, teachers, or both. These physicians provided feedback and suggestions. Based on this feedback, 6 precepts were eliminated, 9 precepts were modified, and 12 were added to give rise to the final list of 52 precepts.

The list of precepts

The 52 precepts that made the final list are shown in Table 1. The list is organized into 4 categories that are closely interrelated, yet represent distinct paradigms. The precepts conveyed under each of the 4 categories—promotion of relationships with patients, principles of the effective clinician, growth and improvement, and values to guide one’s career in medicine—are not listed in any particular order.

Precepts assigned to the domain of promotion of relationships with patients include ideas that relate to the formation of genuine, respectful partnerships between physicians and patients that allow physicians to work with patients toward common goals. Principles of the effective clinician consist of tenets that relate to the consistent performance as a skilled, honorable, and thorough physician. Suggestions under the heading of growth and improvement relate to the importance of physician self-assessment and the commitment to developing both personally and professionally. The final domain, values to guide one’s career in medicine, represents a group of overriding principles that emphasize that the care of patients is a privilege and that to be worthy of the public’s trust, physicians must pay attention to their own needs as well as those of their patients.

Ways to use the list

Below we discuss a few applications of the list; however, we recognize that readers will use the precepts in a way that best suits their own needs and preferences.

Medical trainees and practicing clinicians

The list can be used as a self-assessment tool to promote readers to reflect on their own skills, attitudes, and behaviors within the various domains and to identify areas for improvement. Medical trainees (students, residents, and fellows) could advise their teachers that they are working on specific items and request that faculty observe them in these areas and give them feedback on their performance within that area. Medical trainees who are early in their training or those wishing to improve themselves in all areas can select a different precept to work on each week and can think of creative ways to focus on that precept during the week (e.g., a daily reminder about the precept on one’s PDA or at the top of one’s “to do list”).

Educators

Physician teachers who work with medical learners can use the list of precepts to frame aspects of clinical teaching. Explicit role modeling of lessons felt to be relevant for particular learners with clear communication about why and how they should be performed may produce valuable teaching moments. The precepts could also be incorporated into feedback sessions
Table 1  A list of the 52 precepts that medical trainees and physicians should consider regularly

<table>
<thead>
<tr>
<th>Promotion of relationships with patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Greet patients by their names, tell them your name and your role in their care</td>
</tr>
<tr>
<td>2. Smile</td>
</tr>
<tr>
<td>3. Sit down when talking to patients</td>
</tr>
<tr>
<td>4. Listen</td>
</tr>
<tr>
<td>5. Be wholly present when interacting with patients and avoid unnecessary interruptions</td>
</tr>
<tr>
<td>6. Learn who your patients are and consider sharing something about yourself with them</td>
</tr>
<tr>
<td>7. Show the utmost respect for all patients</td>
</tr>
<tr>
<td>8. Be humanistic, compassionate and caring</td>
</tr>
<tr>
<td>9. Even if it is a struggle to think positively of a patient, always speak of them in a positive way; this will influence your thinking positively</td>
</tr>
<tr>
<td>10. If you are feeling negative emotions towards a patient, try to understand why you are feeling this way</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Principles of the effective clinician</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. The history and physical examination are not like a biopsy fixed in formalin, but are dynamic entities that should be revisited frequently</td>
</tr>
<tr>
<td>12. A patient’s history should not be “aspirated”; it should instead be “built” purposefully with effective communication skills</td>
</tr>
<tr>
<td>13. Be curious – seek to find out exactly how and why events occurred and do not accept diagnoses and conclusions made by others</td>
</tr>
<tr>
<td>14. Recognize the patient as teacher</td>
</tr>
<tr>
<td>15. Elaborate a differential diagnosis that is as broad as the history and physical examination dictate</td>
</tr>
<tr>
<td>16. After forming a diagnostic hypothesis, focus on any symptoms or signs that are either atypical or incompatible with the diagnosis; these must be explained and not ignored</td>
</tr>
<tr>
<td>17. Always consider and exclude catastrophic treatable diseases</td>
</tr>
<tr>
<td>18. Continually strive to improve your diagnostic skills by mentally committing to a specific answer or conclusion before definitive testing</td>
</tr>
<tr>
<td>19. Watching patients walk is a critical component of the physical examination, particularly if their level of function is compromised</td>
</tr>
<tr>
<td>20. Look at the sacrum and heels of any patient who is bed-bound</td>
</tr>
<tr>
<td>21. Think about and plan for how to best deliver the information before telling important news to patients about their health</td>
</tr>
<tr>
<td>22. Explain medical concepts in simple language: avoid medical jargon and make sure that the patient understands</td>
</tr>
<tr>
<td>23. Teach patients what they need to know to make an informed decision</td>
</tr>
<tr>
<td>24. Strive to become a healer</td>
</tr>
<tr>
<td>25. Solicit help when you are stumped or at a loss in caring for a patient</td>
</tr>
<tr>
<td>26. Review your patient’s drug list and require explicit justification for every medication</td>
</tr>
<tr>
<td>27. Remember that the ill patient is not at his best</td>
</tr>
<tr>
<td>28. Do not discuss patients in public places (eg, elevators)</td>
</tr>
<tr>
<td>29. Appreciate the contributions of all members of the health care team</td>
</tr>
<tr>
<td>30. Try to be as organized as possible – be prepared and be thorough yet efficient</td>
</tr>
<tr>
<td>31. Focused reading to answer specific clinical questions is more nourishing leafing through a current issue of a medical journal</td>
</tr>
<tr>
<td>32. Know that much practice, reading, and years of hard work are essential parts of becoming an excellent physician</td>
</tr>
<tr>
<td>33. When you have made a mistake in the care of a patient, follow these steps: (a) admit it, (b) inform the patient, (c) if possible, initiate reparation, (d) institute a mechanism whereby you will not repeat the error, (e) attempt to establish a mechanism whereby others in the system cannot make the error, (f) forgive yourself</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Growth and improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>34. Strive to achieve personal awareness and an understanding of your beliefs, values, and attitudes</td>
</tr>
<tr>
<td>35. Recognize and acknowledge powerful experiences</td>
</tr>
<tr>
<td>36. Seek out and embrace helping relationships</td>
</tr>
<tr>
<td>37. Make time for reflection</td>
</tr>
<tr>
<td>38. Observe other physicians carefully and learn from role models</td>
</tr>
<tr>
<td>39. Realize that people are watching you closely – strive to be a role model for others</td>
</tr>
<tr>
<td>40. Be creative and innovative</td>
</tr>
<tr>
<td>41. Try to look into an accurate mirror</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Values to guide one’s career in medicine</th>
</tr>
</thead>
<tbody>
<tr>
<td>42. Avoid being cynical</td>
</tr>
<tr>
<td>43. Understand that medicine is a public trust</td>
</tr>
<tr>
<td>44. Be humble</td>
</tr>
<tr>
<td>45. Be ethical in all of your work as it relates to the profession of medicine</td>
</tr>
<tr>
<td>46. Aspire to become a great teacher</td>
</tr>
<tr>
<td>47. Stand up for what you believe in</td>
</tr>
<tr>
<td>48. Aim for a comfortable balance between your personal and professional lives</td>
</tr>
<tr>
<td>49. Try your best</td>
</tr>
<tr>
<td>50. Continually search for meaning in your work in medicine</td>
</tr>
<tr>
<td>51. Celebrating successes may help to avoid burnout</td>
</tr>
<tr>
<td>52. Be thankful and happy that you are in medicine</td>
</tr>
</tbody>
</table>
and evaluation forms. The One-Minute Manager\textsuperscript{19} has taught us “to catch people doing things right” and it is likely that medical learners will be motivated if their physician teachers notice and compliment them when they are following the precepts. Physician teachers may also wish to attach one of the precepts to each of the 52 playing cards in a deck and have learners pick a card that they will then have to work on for a couple of days—the “pick a card, any card” technique. The learner would then report back to the physician teacher (and perhaps their colleagues) about the ways that they have practiced and reflected on that particular precept.

Naturally, the precepts may also be incorporated into more formal evaluation systems. Clerkship directors, residency program directors, and physician teachers could transform the list into an evaluation tool, either to be completed by the learner (self-assessment) or others who have observed the learner. Reliability and validity testing will be needed if the list of precepts is to be converted into an assessment instrument. This type of medical education research effort may become particularly pertinent as the Accreditation Council for Graduate Medical Education hopes to find a way to comprehensively assess the competencies of interpersonal and communication skills and professionalism.\textsuperscript{5}

Further development of the list

Whereas some principles guiding our professional behavior are sacrosanct, we understand that other precepts guiding behavior and attitudes evolve over time. It has been pointed out, for example, that differences in the views of the world and of the medical profession are readily seen when comparing the American Board of Internal Medicine’s Charter on Medical Professionalism and the Hippocratic Oath.\textsuperscript{20} We surmise that if such an inventory had been compiled 10 years ago or if the list were to be redone in 10 years, it would likely look different from the one presented in this article. We therefore recognize that the list will benefit from periodic updating, with revisiting of the significance of each item. We welcome comments and suggestions from clinicians, educators, and users of the list (by e-mail) about additional precepts that should be included in future iterations or precepts that do not appear useful.

Conclusions

This list of precepts may be helpful to remind all of us in clinical medicine about the exemplary characteristics, behaviors, and attitudes that are expected as the norm in our profession. It is hoped that the list will promote reflection on the part of medical trainees and physicians about ways in which they can become more effective clinicians, more connected with patients, and more appreciative of the joys and privilege of practicing medicine.

Acknowledgment

The authors are indebted to Ms. Cheri Smith for her assistance in searching the literature.

The authors are grateful for the feedback and suggestions from Drs. Shehzad Basaria, Lisa Beck, Nisha Chandra-Strobos, Michael Choi, Thomas Finucane, Bruce Leff, Rachel Levine, John Mann, David Pearse, Darcy Reed, Eric Seifter, and Donna Windish.

REFERENCES

Medical professionalism and the generation gap
Lawrence G. Smith, MD
Mount Sinai School of Medicine

In recent years, medical professionalism has been scrutinized by physicians, educators, medical literature, and the media. The result of this examination is a generally accepted consensus that professionalism is decreasing in medicine due to a failure to satisfy patient and societal expectations as well as a loss of the medical profession’s dedication to its core values. This seeming deterioration has placed increased pressure on physician educators to measure professionalism among physicians-in-training.1

Criticism regarding professionalism in medicine has often focused on younger physicians, members of a generation that appear to many older physicians as uniquely unprofessional.2,3 This younger generation, with its focus on personal lifestyle and balance, appears to lack the intrinsic virtues necessary for the medical profession. The conflict between generations accentuates the “crisis of professionalism” and has the potential to divide the profession along generational lines, creating many unintended and negative consequences.4,5 However, in focusing on generational loss of virtue, the current discussion has overlooked a key element to professionalism—the transformation of “lay person” to physician.

In the past decade, professionalism has been a topic of hundreds of articles in medical literature (Figure). A possible explanation for this explosion of articles is that the corporatization of medicine and the resulting consumerism of patients has disrupted the contract implicit in the meaning of a “trusted profession.” Another possible explanation is the term “professionalism,” in its current usage, is a meaningless catchphrase and therefore defies satisfactory description. Most of these articles define professionalism as a set of virtues, including altruism, honesty, compassion, and integrity, then create behavioral definitions under each of these virtues that are quantifiable in physicians. In addition to the medical literature’s attempt to define and evaluate professionalism, the American Board of Internal Medicine Foundation, American College of Physicians Foundation, and European Federation of Internal Medicine developed Medical Professionalism in the New Millennium: A Physician Charter, a statement that outlines physicians’ responsibilities to both patients and society.6 Although valuable in the debate, these attempts to define professionalism as a set of virtues, obligations, and behaviors fall short of capturing its essence.

The core of professionalism is the personal transformation of self that takes place in stages during the early years of medical training and practice. Once “lay persons,” medical students redefine themselves as physicians, accepting that they now interact with all of society in a new and different manner.7,8 Accepting this role colors all of one’s perceptions and opinions, setting standards for behavior. Once this transformation occurs, it is impossible to believe being a physician is “just a job.” With this role comes respect, privilege, and trust. The tradition in society is to bestow the title of “doctor” not only on individuals while working in the ambulatory or outpatient setting; rather, the title is given to the persons themselves, believing that the
The transformation from “lay person” to physician has occurred.9

Students today come from a unique generation. This generation has grown up culturally skeptical and technologically savvy and values free time and life balance. As a class, students enter training later in life and include a higher percentage of women. They come predominantly from higher-income families, and will incur record-setting amounts of educational debt.10-12 Led by women, this generation of students will work fewer hours and demand flexible employment opportunities.

Today’s students are potentially as competent and professional as the physicians that have come before. Vying for few positions in US medical schools, today’s students are smart and energetic. This generation of students, however, is criticized for shunning primary care and choosing specialties which provide positive lifestyle factors. As a result, this generation appears to members of other generations as placing personal priorities above those of the patient. In addition, a decreasing number of students appear enthusiastic about the possibility of being someone’s doctor.13-17 This is an important area of concern because this younger generation—skeptical of “total commitment”—may resent the personal transformation to physician.

The conflict in the medical workplace that triggered this recent dialogue on professionalism is between the Baby Boomer Generation and Generation X. Baby boomers define professionalism predominantly in terms of hours worked and “complete” dedication to the job. Dedicated to life balance, Generation Xers do not aspire to be like baby boomers. They believe baby boomers are hypocritical and susceptible to early burnout. In fact, having been raised by absentee, workaholic baby boomers, their priorities are very different from their parents. Their focus on caring for themselves and their families is a positive attribute of Generation X. Baby boomers—creating a value system based on their own life ethic—have confused work ethic with

**Table 1** Generational profiles

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>traditional</td>
<td>personal satisfaction</td>
<td>self-reliant</td>
<td>modern traditional</td>
</tr>
<tr>
<td>Ethic</td>
<td>rapidly declining</td>
<td>dominant</td>
<td>small group</td>
<td>large</td>
</tr>
<tr>
<td>Gender Role</td>
<td>respect, loyalty</td>
<td>ambitious, political</td>
<td>progressive, cynical</td>
<td>loyal, conservative</td>
</tr>
<tr>
<td>Work</td>
<td>classic gender roles</td>
<td>mixing gender roles</td>
<td>unclear</td>
<td>gone</td>
</tr>
<tr>
<td>Heroes</td>
<td>respect the system</td>
<td>respect experience</td>
<td>respect expertise</td>
<td>work to live</td>
</tr>
<tr>
<td>Seminal Events</td>
<td>strong heroes</td>
<td>likes to work</td>
<td>work to live</td>
<td>anti-heroes</td>
</tr>
<tr>
<td>Upbringing</td>
<td>Depression, WWII</td>
<td>some heroes</td>
<td>no heroes</td>
<td>9–11</td>
</tr>
<tr>
<td>Reward</td>
<td>traditional family</td>
<td>Vietnam, BCP</td>
<td>weak USA</td>
<td>protective parents</td>
</tr>
<tr>
<td></td>
<td>a job well done</td>
<td>traditional family</td>
<td>absenteesim parents</td>
<td>work</td>
</tr>
</tbody>
</table>

professionalism. Still in charge of the medical system, baby boomer physicians have continued to enforce a workplace that demands long hours, total dedication to work, rigid approaches to patient care, and disdain for anyone who does not accept their “rules of life.” Generation X physicians are not eager to join this work environment. In addition, Generation Xers—skeptical of organizations and hesitant to make total commitments—appear to be afraid of truly embracing physicianhood and the personal transformation that is critical to professional development.

The recent enactment of work hour regulations by the Accreditation Council for Graduate Medical Education has created a typical generational conflict. Baby boomer physicians, who “thrived” in the old system, blame residents and students for these new regulations. They fail to acknowledge that society is deeply concerned about the harmful effects of long work hours and fatigue on making life and death decisions. The new regulations have the potential to accentuate the professionalism rift as older physicians blame residents for being less dedicated. Likewise, some less motivated young physicians use these rules to justify less than professional dedication to their well being. The profession has created a typical generational conflict. Baby boomer physicians have continued to enforce a workplace that demands long hours, total dedication to work, rigid approaches to patient care, and disdain for anyone who does not accept their “rules of life.” Generation X physicians are not eager to join this work environment. In addition, Generation Xers—skeptical of organizations and hesitant to make total commitments—appear to be afraid of truly embracing physicianhood and the personal transformation that is critical to professional development.

The medical workplace of the future must embrace and encourage the new generation of physicians. Simultaneously, it must allow and encourage the next generation of physicians to grow to be true professionals’ accepting society’s role of “doctor.” The future environment must be patient focused but have flexible work hours and flexible practice design (Table 3). This workplace must recognize that physician well being and balance in life is a valid and important concern and does not negate the attainment of professionalism. It must reward excellence, not endurance. These systems, which by nature will become more discontinuous, must promote seamless team care so that patients never sense a loss of the professional dedication to their well being. The profession has a right to expect excellence and total commitment to medicine but should also allow for structures that encourage balance in life. Finally, this environment should foster the joy of being a physician.

The ultimate challenge that all physicians face, regardless of generation, is to flexibly and respectfully redefine excellence and professionalism in terms that are both generationally diverse and appropriate. Physician leaders need to build bridges instead of barriers. Established physicians need to stop defining perfection as being “just like ourselves” and realize that encouraging professional excellence in ways that are culturally and generationally diverse is the only hope for the future of the medical profession. Let us never allow a medical culture to exist where young physicians are afraid of falling in love with being a doctor.18,19

<table>
<thead>
<tr>
<th>Table 2 Essential attributes of the “physician”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embrace being a physician</td>
</tr>
<tr>
<td>Caring and altruistic</td>
</tr>
<tr>
<td>Honesty, integrity</td>
</tr>
<tr>
<td>Team player</td>
</tr>
<tr>
<td>Strive for excellence</td>
</tr>
<tr>
<td>Accept the duty for serving patients and society</td>
</tr>
<tr>
<td>Courage, heroism</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 Attributes of the “future environment”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient focused</td>
</tr>
<tr>
<td>Flexible work hours</td>
</tr>
<tr>
<td>Prioritize physician well-being and life balance</td>
</tr>
<tr>
<td>Reward excellence, not endurance</td>
</tr>
<tr>
<td>Promote seamless team care</td>
</tr>
<tr>
<td>Expect excellence and total commitment doing work</td>
</tr>
<tr>
<td>Foster joy of being a doctor</td>
</tr>
</tbody>
</table>
References

LETTERS

Severe hypertriglyceridemia in a patient with lupus

To the Editor:

A 26-year-old woman was referred in September 2002 because of the new onset of severe hypertriglyceridemia associated with recurrent pancreatitis. She had a past history of systemic lupus erythematosus (SLE) and had been treated in 1988 and 1996 for lupus nephritis. We excluded hyperglycemia, pregnancy, and excess dietary fat or alcohol as secondary causes of hypertriglyceridemia. The patient’s only medication was Bezafibrate.

Clinical examination was normal and blood tests showed serum creatinine 0.72 mg/dL, albumin 4.8 g/dL, and antidsDNA 75.4 µg/mL (0-30). The fasting triglycerides were 2407 mg/dL and the cholesterol was 278 mg/dL. Agarose gel lipid electrophoresis showed a Fredrickson type I pattern suggesting defective lipolysis. Following heparin release, the patient’s plasma had no detectable lipolytic activity and small aliquots of pre-heparin plasma from the patient significantly reduced lipolysis in samples from normal controls. We suspected the presence of an autoantibody either to lipoprotein lipase or apolipoprotein CII and commenced immunosuppression.

We prescribed Azathioprine in October 2002 but withheld steroids for fear of exacerbating hypertriglyceridemia. Because of the patient’s persistent hypertriglyceridemia and recurrent pancreatitis, we performed 6 plasma exchanges spread over 2 weeks in November 2002. We substituted Cyclophosphamide for Azathioprine and commenced high-dose steroids. The patient’s triglycerides normalized and there were no further episodes of pancreatitis. In November 2003 hypertriglyceridemia and pancreatitis recurred despite ongoing immunosuppression. We intensified immunosuppression and performed 5 plasma exchanges in December 2003. Hypertriglyceridemia persisted and was associated with recurrent pancreatitis. In March 2004 we administered 4 doses (375 mg/m²) of Rituximab at weekly intervals. We continued immunosuppression with Cyclophosphamide and Prednisone. Plasma triglycerides decreased substantially following Rituximab and were normal when last tested in October 2004 (Figure 1). The patient reported no further attacks of acute pancreatitis but chronic abdominal pain consistent with chronic pancreatitis had set in.

Severe hypertriglyceridemia (>1000mg/dL) of any etiology can result in acute pancreatitis. Abnormal lipoprotein metabolism is common in SLE with moderate hypertriglyceridemia, HDL hypocholesterolemia and small dense LDL particles being the commonest abnormalities. Severe hypertriglyceridemia may occur in predisposed patients exposed to secondary metabolic stressors such as diabetes or steroids. Only a few patients with severe hypertriglyceridemia secondary to autoantibodies against lipolytic enzymes or their cofactors have been reported.

A single plasma exchange lowers triglycerides rapidly but temporarily. We performed a series of plasma exchanges, as our primary aim was removing the pathogenic antibody. The first series of plasma exchanges was very successful, but we were unable to induce a second remission following relapse. We speculate that new antibody production exceeded removal by plasma exchange. Because the patient’s clinical condition was dire and plasma exchange and conventional immunosuppression had failed, we treated her with Rituximab. This chimeric murine/human monoclonal anti-CD20 antibody causes a selective transient depletion of the CD20+ B-cell subpopulation. Rituximab was originally developed to treat certain lymphomas; but due to its powerful and long-lasting effect on B-cells, it is increasingly being used in autoimmune disorders. Acute pancreatitis in SLE is often attributed to lupus disease activity but may occasionally be linked to other causes. Identifying a remediable cause is the key to successful therapy.

Acknowledgements

We thank Miss B. Ratanjee for expert technical assistance.

Dirk J. Blom, MBChB (UCT), M.Med (UCT), FCP (SA);
A. David Marais, MBChB (UCT), FCP (SA)
Division of Lipidology
Department of Medicine
University of Cape Town
doi:10.1016/j.amjmed.2004.10.023
Absence of effect of folic acid flour fortification on anticonvulsant drug levels

To the Editor:

There is a bidirectional interaction between folic acid and phenytoin. Long-term phenytoin therapy can result in folate deficiency and hyperhomocysteinemia, while supplementation with folic acid may lower serum phenytoin, possibly leading to poorer seizure control. We questioned whether serum phenytoin concentrations changed in relation to mandatory folic acid fortification of Canadian flour, which has provided about 0.2 mg per day of additional synthetic folic acid since March 1997, with complete fortification achieved by March 1998.

We performed a population-based study of all phenytoin drug concentrations measured by a large Canadian provincial laboratory (BC Biomedical Laboratories, British Columbia) between January 1995 and October 2003. Valproic acid, another anticonvulsant, was chosen as a control during the same time period because it is not known to affect, or be affected by, folate metabolism.

An interventional ARIMA time series model assessed the effect of fortification on the subsequent monthly mean concentrations of phenytoin. The effect of folic acid fortification was represented by a ramp function, beginning in March 1997, with full fortification by March 1998. Individual age and sex were included as covariates in the model.

A 2-sided P value of 0.05 was set as the limit of statistical significance for all analyses, using SAS version 8.2. All data were anonymized, and permission to conduct this study was obtained from the Research Ethics Board of Sunnybrook and Women’s College Health Sciences Centre and BC Biomedical Laboratories.

There were 39,636 phenytoin measures performed in 7103 persons (mean [SD] age 54.3 [21.1] years; 40.0% female). No significant change in phenytoin drug levels was observed in relation to initial (P = 0.72) or complete (P = 0.65) food fortification (Figure 1). A total of 37,778 serum valproic acid measures were taken in 8086 participants (42.4 [19.9] years; 42.4% female). Again, initial (P = 0.52) or complete (P = 0.71) folic acid food fortification was also not associated with a significant change in mean monthly valproic acid concentrations (Figure 1).

In a large outpatient population of persons who underwent testing, no appreciable change in phenytoin drug concentrations was observed in relation to universal folic acid food fortification.

As a limitation, we did not determine which persons in this study were folate deficient or taking folic acid tablet supplements, nor did we document the rate of seizures over...
time. As a strength, these novel data were derived from a large sample of persons who underwent anticonvulsant drug testing in real clinical practice during a period in which few were aware that folic acid flour fortification had begun.3

We previously documented more than a 65% rise in serum folate concentrations among more than 15,000 older women living in British Columbia and Ontario.3 The current study findings may provide some degree of assurance—especially in countries like Canada, the U.S., and Chile with universal folic acid fortification in place—that higher folic acid intake has not impacted adversely on anticonvulsant metabolism, as reflected in aggregate serum levels.

Women taking anticonvulsants and planning a pregnancy are encouraged to take at least 1 mg of folic acid daily.5 In previous small studies, more than 1 mg of daily folic acid supplementation in nonpregnant folate-deficient patients was associated with a lowering of serum phenytoin concentrations.6 The current study findings do not address whether free phenytoin levels are altered at higher doses of folic acid intake. Large-scale clinical studies of persons on long-term phenytoin are needed to assess folate deficiency and determine whether higher doses of folic acid affect phenytoin metabolism.

Acknowledgment

We thank BC Biomedical Laboratories for contributing data to this study. The Spina Bifida and Hydrocephalus Association of Canada and the physicians of Ontario through the Physicians’ Services Incorporated Foundation supported this study.

Figure 1  Mean monthly serum phenytoin (upper) and valproic acid (lower) concentrations in relation to folic acid food fortification.

References

Dysautonomia in Gulf War syndrome and in fibromyalgia

To the Editor:

Fibromyalgia (FM), chronic fatigue syndrome, and Gulf War syndrome (GWS) have overlapping clinical features. Basically, they are characterized by diffuse musculoskeletal pain, chronic fatigue, and cognitive impairment.1

We read with interest the article by Haley et al on the blunted circadian variation of autonomic regulation in veterans with Gulf War syndrome.2 Their methods and their results are very similar to our findings in patients with fibromyalgia.3

Haley et al analyzed the spectral power of heart rate variability in a group of GWS veterans. They found that the high-frequency spectral power (a surrogate of parasympathetic activity) was blunted at night in GWS veterans when compared to healthy veterans who served as controls.

The parasympathetic and sympathetic branches of the autonomic nervous systems have antagonistic effects on the sinus node of the heart. So blunted parasympathetic activity reflects, on the other hand, sympathetic overexpression.

In a subgroup of their patient population (15 patients and 11 controls), Haley et al also studied peroneal nerve sympathetic traffic by means of microneurography. There was a tendency for GWS veterans to have increased values. This tendency did not reach statistical significance, perhaps due to the small sample size of this subgroup.2

We used similar heart rate variability analysis to study the pathogenesis of FM. We found changes consistent with the small sample size of this subgroup.

We proposed that such autonomic dysfunction explains all FM symptoms.9 FM core features (chronic widespread pain and widespread allodynia) can be explained through the pathogenesis known as sympathetically maintained pain. We based this proposal on 3 arguments derived from controlled clinical studies: 1) patients with FM display signs of sympathetic hyperactivity, 2) FM pain is submissive to sympatholytic interventions, and 3) pain is rekindled by norepinephrine injections.10

So autonomic dysfunction may be the unifying pathogenesis behind these overlapping chronic multisystem illnesses. From a research perspective, GWS has the advantage of displaying a uniform triggering event both in time and place. This fact may allow researchers to uncover the nature of the triggering event that leads to autonomic dysfunction.

Manuel Martinez-Lavin, MD
Chief, Rheumatology Department

References


Letter regarding: A 70-year-old African-American man with a history of fever, chills, and malaise

To the Editor:

In the interesting case presented by Nivatpumin for “Images of Osler,” the legend for Figure 1: Lead II rhythm strip reads, “. . . day of admission, demonstrating first degree atrioventricular block (middle)”; however, the referred rhythm strip shows progressive prolongation of the PR interval that can be more clearly seen on the last 3 QRS complexes in the strip (Figure 1). As such, it represents a second degree atrioventricular (AV) block Mobitz type I or Wenckebach, one of the many conduction system complications associated with infective endocarditis. Due to the position of their respective valve annulus relative to the conduction system structures, extension of mitral valve endocarditis, being closer to the AV node, can present with first degree AV block and second degree AV block Mobitz I, while extension of aortic valve endocarditis, being closer to the His-Purkinje system, can present with second degree

Manuel Martinez-Lavin, MD
Chief, Rheumatology Department

Antonio G. Hermosillo, MD
Chief, Electrophysiology Department
National Institute of Cardiology Mexico
do:10.1016/j.amjmed.2004.11.023
AV block Mobitz II, AV dissociation and bundle branch block.

Despite significant advances in medical science and technology over the last few decades, infective endocarditis is still associated with considerable morbimortality. Continuous ECG monitoring or a simple daily ECG during the first 2 weeks of treatment together with a high index of suspicion for complications will help physicians caring for patients with endocarditis.

Victor J. Castro, MD
McAllen Heart Hospital
University of Texas Health Science Center

Figure 1 Lead II rhythm strips 5 months before admission, demonstrating normal sinus rhythm with normal PR interval (top); day of admission, demonstrating first degree arterioventricular block (middle); and hospital day 5, showing third degree heart block with prolongation of the QRS complex.
The purpose of the new Medical Humanities section in
The American Journal of Medicine is to create a forum for
writers on medicine and the arts, including literature. We
invite our readers to submit articles about topics in these
areas, including narratives about individual experiences as
patient or physician. The goal is to add a new dimension to
the Green Journal that is not strictly scientific and quanti-
fiable, but nevertheless essential to good medical practice.

Becoming a better clinician is a complicated project
involving reading and reflection outside the clinic, an
activity that takes up spare time and for which there is no
reimbursement. What is the benefit to the individual prac-
titioner of joining a conversation about other doctors, other
times, other cultures, on the last pages of The American
Journal of Medicine? Quantifiable benefit is lacking here, as
no double-blind experiment has yet been done, nor has it
been demonstrated that reading a patient’s story, for exam-
ple, makes you a more empathetic physician. What we are
left with is common sense telling us that information about
medicine derived from art and literature must be useful to
medical professionals. Relevant here is the fact that numer-
ous doctors have been writers; these are clearly of particular
interest. Many names ring a bell: Dannie Abse, Robert
Bridges, Ethan Canin, Robert Coles, Alex Comfort, Robin
Cook, Michael Crichton, A.J. Cronin, Harvey Cushing, Sir
Arthur Conan Doyle, and Perri Klass. These are only some
of the 173 twentieth-century medical doctors writing in
English, including MDs from medicine, psychiatry, surgery,
neurology, research, administration/public health, pediat-
rics, and pathology. This factual study of physician-writers
affirms a certain relationship between art and medicine that
finds expression in poetry and prose.¹

Two of the best-loved and most quoted physician-writers
explain this connection in their own lives. First, William
Carlos Williams (1883–1963), for 40 years a practitioner in
Rutherford, New Jersey, writes in his autobiography about
his energizing contact with patients: “As I say, often after I
have gone into my office harassed by personal perplexities
of whatever sort, fatigued physically and mentally, after 2
hours of intense application to the work, I came out at the
finish completely rested (and I mean rested), ready to smile
and to laugh as if the day were just starting.”² Reading a
story or poem has been known to produce a similar catharsis.

Second, the witty Russian doctor-writer Anton Chekhov
(1860–1904) draws the parallel between art and medicine in
terms of sexual excitement. Writing to his good friend the
critic Suvorin, he jokes: “Medicine is my lawful wedded
wife, and literature my mistress. When one gets on my
nerves, I spend the night with the other. This may be
somewhat disorganized, but then again it’s not as boring,
and anyway, neither one loses anything by my duplicity. If
I didn’t have medicine, I’d never devote my spare time and
thoughts to literature. I lack discipline.”³ Stimulated by their contacts with a range of pa-
tients, these two doctor-writers communicate their excite-
ment, frustration, or sorrow in stories that call for a re-
sponse: yes, I have felt anger toward a patient; yes, I have
been a patient and experienced that kind of mortification;
yes, I have been surprised by a patient and have learned
from that patient. Williams’ and Chekhov’s doctor stories
trigger a pleasant twinge of that recognition which is the
essence of our enjoyment of art.

References

1. Daniel C. Bryant. A roster of twentieth-century physicians writing in
York: New Directions; 1951. 357.
3. Heim MH, Karlinsky S. Anton Chekhov’s Life and Thought: Selected