COMMENTS AND RESPONSES

Quality of Care in Patients with Diabetes

TO THE EDITOR: We applaud Greenfield and colleagues for demonstrating the confounding effects of physician-level clustering on quality assessments across provider groups (1). We disagree, however, with one of the study's major inferences: that endocrinologists and generalists do not significantly differ in the quality of patient diabetes care after adjustment for case mix and physician-level clustering.

Physician-level clustering limited the power of this study to detect major differences in care between endocrinologists and generalists. Still, the maximal likelihood estimations for most of the process and outcome measures suggested that endocrinologists offered favorable care. Physician-level clustering did not qualitatively change maximal likelihood estimates on any of the measures; it simply widened the confidence intervals. The failure to achieve "statistical significance" due to high variances should not necessarily minimize the importance of a consistent point estimate that suggests meaningful differences in care quality.

There is face validity to the notion that endocrinologists assess hemoglobin A_{1c} levels, screen for urine proteinuria, and ask patients for self-measured blood glucose levels more frequently than generalists. Greenfield and colleagues' Figure, which displays the proportion of patients achieving hemoglobin A_{1c} levels less than 0.10 (10%), further supports meaningful specialty differences. Seven endocrinology sites compared with only two generalist sites statistically performed above the mean.

Despite advancements in case-mix and physician-clustering techniques, we question the rationale for comparisons between specialists and generalists. In the end, aren't we comparing apples and oranges? Generalists and specialists operate under different contexts in our health care system. In our opinion, the most enlightening aspect of Greenfield and colleagues' analysis was the observed within-specialty variation in care quality. Why do some endocrinologists' practices achieve 90% success in reducing patients' hemoglobin A_{1c} levels to less than 0.10 (10 %) while others achieve only 50% success? We look forward to future evaluations that address this question, particularly those that appropriately adjust for physician-level clustering.

Parthiv J. Mahadevia, MD
Seth Himelhoch, MD
Joel Braunstein, MD
Johns Hopkins Hospital
Baltimore, MD 21287

Reference

1. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med. 2002;136:111-21. [PMID: 11790062]

TO THE EDITOR: The Board of Directors of the American Association of Clinical Endocrinologists (AACE) unanimously expresses its concern about the study by Greenfield and colleagues on profiling care provided by different groups of physicians (1). Their study has multiple methodologic flaws. We are disappointed that *Annals of Internal Medicine* distributed a Summary for Patients (2) that uncritically accepted the paper's conclusions, assumed extrapolation to physicians' practices in general, and distributed oversimplified facts to patients and the medical community.

There are several problematic issues. First, classification of sites as generalist or specialist was in error. Substantial co-management occurred, and since practice site and physician group and cluster were commingled to an unknown degree for multiphysician sites, internal validity is a serious issue. Second, the American Diabetes Association's Provider Recognition Program (PRP), from which the data used by Greenfield and colleagues were derived, was not designed as a research study. A convenience sample rather than a random sample was used. Only two sites were audited, and there was no attempt to make the patients selected representative. Third, Greenfield and colleagues made no distinction between type 1 and type 2 diabetes mellitus. Fourth, clinical measures used as a basis of comparison were not uniform, nor were sequential data collected to determine whether physician effect occurred. Fifth, Greenfield and colleagues' study had multiple methodologic problems in research design, handling of sample size, and imputation of missing data (3, 4).

Of importance, the original uncorrected data showed significant benefit from endocrinologist care in hemoglobin A_{1c} level, patient satisfaction, and other measures. Many statistically significant differences persisted after correction for patient characteristics. Only when the data were given a second round of "adjustment" for physician clustering did most of the differences become nonsignificant. This could have been related to a loss in statistical power caused by the model itself.

Previous studies have demonstrated that endocrinologists provide superior care for patients with diabetes (5, 6). Members of the AACE firmly support measures to improve the quality of life in such persons. We applaud quality improvement efforts and appropriate assessment tools, and we look forward to well-designed studies that will clarify our efforts to treat this serious disorder.

Rhoda H. Cobin, MD
American Association of Clinical Endocrinologists
Jacksonville, FL 32204

Helena W. Rodbard, MD
American College of Endocrinology
Jacksonville, FL 32204

Note: An expanded version of this letter is available at www.annals.org.

References

1. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med. 2002;136:111-21. [PMID: 11790062]

2. Summaries for patients. Comparing the quality of diabetes care by generalists and specialists. Ann Intern Med. 2002;136:I42. [PMID: 11928735]

3. Campbell DT, Stanley JC. Experimental and Quasi-experimental Designs for Research. Boston: Houghton Mifflin; 1963.

4. Cook TD, Campbell DT. Design and Analysis Issues for Field Setting. Boston: Houghton Mifflin; 1979.

5. Levetan CS, Passaro MD, Jablonski KA, Ratner RE. Effect of physician specialty on outcomes in diabetic ketoacidosis. Diabetes Care. 1999;22:1790-5. [PMID: 10546009]

6. Hellman R, Regan J, Rosen H. Effect of intensive treatment of diabetes on the risk of death or renal failure in NIDDM and IDDM. Diabetes Care. 1997;20:258-64. [PMID: 9051368]

TO THE EDITOR: Our knowledge of the PRP and our experience as practitioners at sites applying for PRP recognition lead us to question the conclusions of the article by Greenfield and colleagues (1). The authors classified sites in the PRP pilot as endocrinologist or generalist. At least three sites, represented by two of us, were multispecialty groups of internists, family physicians, and endocrinologists but were classified as single-specialty groups. Care delivered with frequent formal and informal consultation between generalists and specialists cannot be classified as generalist or specialist care. Incorrect classification of groups is a major concern in a paper that bases its major findings on the premise that groups were either all endocrinologists or all generalists.

A second concern relates to the analysis of glycated hemoglobin results. Some sites submitted hemoglobin A_{1c} values while others submitted total glycated hemoglobin values. Adjustments were made to convert glycated hemoglobin results into hemoglobin A_{1c} results, but many variations among the analytic techniques used by the sites persisted even after such adjustment. This nonrandom variation in reporting of glycated hemoglobin results represents a significant potential confounder to the quality of outcome measures reported.

The PRP was never intended to be a research study. Sites believed they were contributing data as part of a quality improvement program. Quality checks of data collected did not occur during the pilot. There were two audits. Most sites relied on untrained office personnel to do medical record abstraction. Therefore, the data presented by Greenfield and colleagues are not adequate to support the conclusion of no difference in outcomes between generalists and specialists.

Sam Weir, MD
Carolina Permanente Medical Group

Robert Stone, MBA
American Healthways, Inc.
Nashville, TN 37215

Philip Levy, MD
Phoenix Endocrinology Clinic
Phoenix, AZ 85006

Stephen F. Hodgson, MD
Mayo Clinic
Rochester, MN 55905

Barbara Fleming, MD, PhD
Centers for Medicare & Medicaid Services
Silver Spring, MD 20904

Reference

1. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med. 2002;136:111-21. [PMID: 11790062]

IN RESPONSE: Mahadevia and associates underscore the importance of understanding the central point of our paper, that is, that variation between physicians (physician-level clustering) can limit the power of studies comparing *any* groups of physicians (for example, fee-for-service vs. managed care settings or high-quality vs. low-quality practice settings). All studies attempting such comparisons must be designed with not only sufficient numbers of patients per physician but also sufficient numbers of physicians per comparison group to ensure adequate power. The analytic methods we used were specifically chosen to investigate the *separate* effects of adjustment for patient characteristics and then physician clustering on observed differences, holding differences constant (1). The point estimates therefore did not change. It was the variability *between physicians* in each group rather than differences in case mix that accounted for the lack of statistical significance between groups. A recent study (2) showed similar findings when comparisons of groups of hospitals were adjusted for within-group variability.

Cobin and colleagues and Weir and associates raise substantially overlapping issues. We believe the issue of specialty misclassification to be inaccurate for the following reasons. First, as stated in our paper, each practice site self-reported its specialty and practices were specifically instructed to choose patients for whom they provided the principal diabetes care over the previous year. As a quality check, each site subsequently reaffirmed its specialty designation. Second, if there had been substantial misclassification, unadjusted differences in our Table 1 would have been smaller. Third, to determine the exact impact of site misclassification on our results, we reclassified each site as if it had been originally misclassified, first individually, then two and three sites at a time. The 29 individual, 406 paired, and 3654 three-way analyses of all possible misclassifications supported our original findings.

Cobin and Weir and their colleagues also confuse misclassification with co-management. As in other quality-of-care efforts (for example, the National Committee for Quality Assurance's Health Plan Employer Data and Information Set), co-management redounds to the credit of the principal caregiver, since coordination and referrals are the role of the responsible physician. We used no fructosamine values. For the two sites reporting total glycated hemoglobin, all values were corrected to hemoglobin A_{1c} level according to the instructions of the manufacturer of the test kits used. As explained to PRP participants, a random audit of 5% of all practices was to be performed to evaluate data quality. The observed high interrater reliability in the 5% sample of practices in the pilot program has been replicated in the subsequent 4 years of the program with over 300 participating practices ($\kappa = 0.87$).

Secondary data (for example, medical records and administrative data) are commonly and appropriately used for a variety of research purposes. The characteristics of the PRP database (its reliability, design, and sampling strategy) are more than adequate for the methodologic purpose to which we applied it. The average of 67 patients per practice allowed considerable precision of estimates of practice performance. However, not every site had 67 patients; the resulting total of 1750 patients reflects differences in the number by practice. No clinical performance measures were imputed, as noted in our paper. Only the measure of patient satisfaction (favoring endocrinologists) was imputed. The PRP measures committee concluded that the performance measures should not differ for type 1 compared with type 2 diabetes mellitus. We did, however, control specialty comparisons for the use of insulin.

Cobin and associates also confuse the methodologic conclusion of

our study (that is, the need for more appropriately designed studies comparing groups of physicians) with a more generalized conclusion about specialty differences. The highly selected nature of the PRP, which made it a conservative and appropriate resource for investigating physician clustering and patient case mix, would compromise any generalized conclusion about the two specialty groups. We made no such conclusions. We hope that the PRP and programs like it ultimately serve to improve diabetes care delivered by all provider groups.

Sheldon Greenfield, MD
Sherrie H. Kaplan, PhD, MPH
Tufts University School of Medicine
Boston, MA 02111

Richard Kahn, PhD
American Diabetes Association
Alexandria, VA 22311

References

1. Localio AR, Berlin JA, Ten Have, TR, Kimmel SE, Adjustments for center in multicenter studies: an overview. Ann Intern Med. 2001;135:112-23. [PMID: 11453711]

2. Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented internet health care report card: the risk of quality ratings based on mortality data. JAMA. 2002;287:1277-87. [PMID: 11886319]

Editors' Note: Drs. Cobin and Rodbard expressed disappointment in the Summary for Patients (1) that accompanied Greenfield and colleagues' article (2). The Editors respectfully disagree with their assertions that the summary uncritically accepted the paper's conclusions, extrapolated to physician's practices in general, and oversimplified facts. Contrary to our correspondents' view, the summary drew no conclusions about which type of physicians provide the best diabetes care. Rather, it concluded that without analyses that consider patient factors and physician-level clustering, one might come to erroneous conclusions about the care provided by different types of doctors. The summary did not overgeneralize but clearly cautions readers that the findings in the 29 study sites might not apply to other settings. Last, we do not feel that the summary oversimplified the facts. The Editors try to write the summaries at a reading level that permits them to be accessible to a large proportion of the public. In this case, we presented in plain language a difficult methodologic issue about comparisons among physician groups. We feel that the summary clearly conveys the paper's essential message that comparing care among groups of physicians is difficult to do because of the many patient, doctor, and practice factors that influence such comparisons.

References

1. Summaries for patients. Comparing the quality of diabetes care by generalists and specialists. Ann Intern Med. 2002;136:I42. [PMID: 11928735]

2. **Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL.** Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med. 2002;136:111-21. [PMID: 11790062]

Measuring Patient Adherence

TO THE EDITOR: In their editorial comment (1) on our article (2), Turner and Hecht provided perspective regarding adherence measurement in research and clinical practice. However, we disagree with several of their comments. The composite adherence score (CAS) is not simply a "supplemented Medication Event Monitoring System." The rigorous approach to replacing electronic bottle cap measurements in creation of CAS, the variety of prevalent reasons to use CAS beyond missing electronic measurements, and the fact that nearly 40% of CAS values were derived from calibrated pill count or interview measurements demonstrate the complex composition of an improved measure that uses several data sources. Improving adherence measurement requires evaluation of why data are missing and consideration of the quality of available information. As Turner and Hecht point out, the interaction between adherence-enhancing devices (for example, pillboxes) and electronic methods makes this process even more complex. Therefore, explicit consideration of such devices in computing adherence is crucial.

Missing data are problematic in measuring adherence. We found that this is an important issue not only for self-report measurements but also for pill count and electronic measurements. Missing proportions for each type of measurement were 15%, 6%, and 12%, respectively, for study periods for which CAS was computed. In contrast to most studies that measure adherence cross-sectionally or for brief periods, this level of missing values demonstrates the difficulty in measuring this elusive behavior over a longer period. However, since clinicians strive for long-term adherence in the treatment of HIV disease and other chronic illnesses, these are precisely the types of adherence measurements that are needed.

Last, we do not share the optimism of Turner and Hecht concerning self-report adherence measurement. Our data add to the mountain of evidence decrying the insensitivity of self-reported adherence. We see little justification to rely on such methods in the research setting when other methods exist that are more closely linked to important clinical outcomes.

Honghu Liu, PhD
Andrew Kaplan, MD
Neil Wenger, MD
University of California, Los Angeles
Los Angeles, CA 90095

References

1. Turner BJ, Hecht FM. Improving on a coin toss to predict patient adherence to medication [Editorial]. Ann Intern Med. 2001;134:1004-6. [PMID: 11352702]

2. Liu H, Golin CE, Miller LG, Hays RD, Beck CK, Sanandaji, S, et al. A comparison study of multiple measures of adherence to HIV protease inhibitors. Ann Intern Med. 2001;134:968-77. [PMID: 11352698]

IN RESPONSE: We appreciate the thoughtful comments of Dr. Liu and his coauthors. We understand that the authors feel that their adherence score reflects a composite of several approaches to measure adherence. However, the authors first used electronic monitor data and then filled in missing data points with pill count and, finally, interview data. Our editorial aimed to help the reader understand that electronic monitor data served as the centerpiece of this approach. If a researcher had nearly complete adherence data on his or her study participants from electronic monitors, Liu and colleagues' approach would require minimal additional data from pill counts or patient interviews. Other potential weaknesses of electronic monitors, such as patients "decanting" multiple doses at once, are not addressed by Liu and colleagues' approach. Liu and colleagues' letter reiterates their preference for using electronic monitors

to measure adherence and their reluctance to trust patient self-report. In our editorial, we described a study that observed a significant relationship between self-reported adherence to antiretroviral therapy and viral load (1). We believe that carefully collected self-report data still offer promise in evaluating patient adherence to medications. However, strenuous efforts are needed to reduce the known biases of self-report data, such as social desirability, time frame covered by the questions, type of interviewer, language barriers, and poor recall. In many clinical settings, self-report is the best measure we have. In research settings, self-report remains a far less costly, complex, and intrusive method of determining how patients take their medications. Vitolins and coworkers (2) recently reviewed the strengths and weaknesses of multiple adherence measures but concluded that the best option is to use self-report with other "objective" measures. They acknowledged that it is necessary to make the best of a difficult situation but did not dismiss self-report. Similarly, we hope that researchers will continue to investigate ways to improve and supplement self-report as a measure of adherence.

Barbara J. Turner, MD
University of Pennsylvania
Philadelphia, PA 19104

Frederick M. Hecht, MD
San Francisco General Hospital
San Francisco, CA 94410

References

1. Bangsberg DR, Hecht FM, Charlebois ED, Zolopa AR, Holodniy M, Sheiner L, et al. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. AIDS. 2000;14:357-66. [PMID: 10770537]
2. Vitolins MZ, Rand CS, Rapp SR, Ribisl PM, Sevick MA. Measuring adherence to behavioral and medical interventions. Control Clin Trials. 2000;21(5 Suppl):188S-94S. [PMID: 11018574]

Blue Light and Milk

TO THE EDITOR: I am an enthusiastic reader of "On Being a Doctor," which is usually the first section I turn to in *Annals*. These essays have consistently described experiences and insights that were interesting, original, and thought provoking. My disappointment in the essay "Blue Light and Milk" (1) was therefore acute. This memoir struck me as sentimental and self-indulgent and the language as both trite and pretentious. I have no doubt that Dr. Morowitz is a warm, loving, and totally admirable father, but his reminiscences would have been more appropriate in a parenting magazine than in *Annals of Internal Medicine*.

Alice M. Godfrey, MD
Aptos, CA 95003

Reference

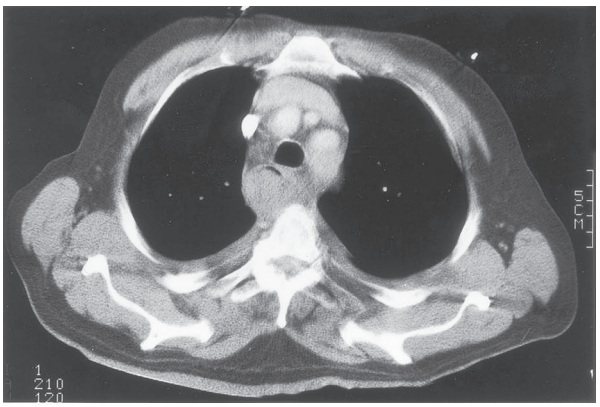
1. Morowitz DA. Blue light and milk. Ann Intern Med. 2001;134:1148-9. [PMID: 1141205]

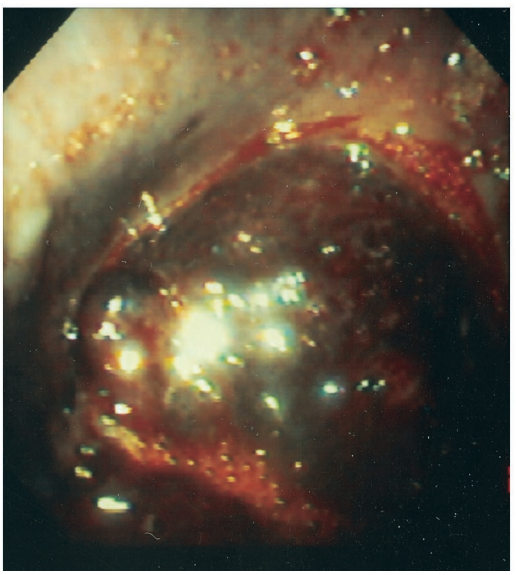
RESEARCH LETTERS

Spontaneous Intramural Hematoma of the Esophagus

TO THE EDITOR: Spontaneous intramural hematoma of the esophagus is an uncommon cause of acute odynophagia. It can result in chest pain, dysphagia, and hematemesis. We report a case of a man with hemophilia and partial factor VIII deficiency who presented with sudden onset of acute odynophagia and substernal chest pain radiating to the anterior chest and shoulder. The patient did not report a history of emesis, trauma, ingestion of nonsteroidal anti-inflammatory drugs, or gastrointestinal bleeding episodes. His vital signs were stable, and results of liver and kidney function tests were normal. His initial prothrombin time was 44 seconds; international normalized ratio was 1.2. Computed tomography of the chest (**Figure**, *top*) and subsequent endoscopy showed proximal esophageal obstruction due to intramural hematoma. The patient was treated with rVIII factor infusion. Repeated computed tomography (**Figure**, *bottom*) and endoscopy after 7 days showed resolution of the hematoma. Symptoms also resolved. Biopsy of the esophagus showed no

Figure. **Spontaneous intramural hematoma of the esophagus before (*top*) and after (*bottom*) therapy.**





malignant disease but indicated extensive hemorrhage. The patient was subsequently discharged from the hospital in stable condition.

Spontaneous intramural hematoma of the esophagus, especially in persons with hemophilia, is probably due to spontaneous bleeding within the submucous layer of the esophagus. A clinical history of acute odynophagia is suggestive. Imaging studies such as computed tomography and magnetic resonance imaging, oral contrast studies such as double-barrel esophagus and mucosal strip signs, and endoscopy are useful diagnostic tests. Differential diagnoses are the Mallory–Weiss syndrome, the Boerhaave syndrome, acute myocardial infarction, and dissection of thoracic aorta. No previous emesis is present in patients with spontaneous intramural hematoma of the esophagus, which is a major differentiating factor.

Ajay Chawla, MD
Michael H. Blume, MD
Jerald Insel, MD
Good Samaritan Hospital
Baltimore, MD 21239

References

1. Oldenburger D, Gundlach WJ. Intramural esophageal hematoma in a hemophiliac. An unusual cause of gastrointestinal bleeding. JAMA. 1977;237:800. [PMID: 299897]
2. Kerr WF. Spontaneous intramural rupture and intramural haematoma of the oesophagus. Thorax. 1980;35:890-7. [PMID: 6973833]
3. Shay SS, Berendson RA, Johnson LF. Esophageal hematoma. Four new cases, a review, and proposed etiology. Dig Dis Sci. 1981;26:1019-24. [PMID: 7028429]
4. Hiller N, Zagal I, Hadas-Halpern I. Spontaneous intramural hematoma of the esophagus. Am J Gastroenterol. 1999;94:2282-4. [PMID: 10445563]
5. Ackert JJ, Sherman A, Lustbader IJ, McCauley DI. Spontaneous intramural hematoma of the esophagus. Am J Gastroenterol. 1989;84:1325-8. [PMID: 2801687]

Severe Rhabdomyolysis Related to Cerivastatin without Gemfibrozil

TO THE EDITOR: The hydroxymethylglutaryl coenzyme A reductase inhibitors ("statins") decrease serum cholesterol levels and morbidity and mortality from coronary heart disease. Although these drugs are well tolerated, serious adverse effects may occur. All members of this class may cause rhabdomyolysis, particularly when combined with gemfibrozil.

We report six cases of severe rhabdomyolysis in patients taking cerivastatin without concomitant gemfibrozil. The patients were symptomatic (creatine phosphokinase levels >5000) and were hospitalized with no other clear causes of rhabdomyolysis (**Table**). The cases occurred among approximately 3000 patients exposed to the drug for about 9 months. All patients were taking 0.4 mg of the drug per day. Three cases occurred during the first month of exposure.

We practice in the Harris County Hospital District, a public health care system in Houston, Texas, that has a defined formulary and its own pharmacies and serves a largely uninsured population. To minimize costs of statins but still encourage their use, each year we have asked pharmaceutical companies for bids and then switched patients to the least expensive therapeutically equivalent statin. Thus, about 1400 persons taking fluvastatin and 1700 patients taking simvastatin were switched to cerivastatin in March 2000. About 800 patients were receiving the "second-line" statin, atorvastatin, until March 2000, when the switch to higher-dose simvastatin began.

We reviewed all adverse drug reaction reports for Harris County Hospital District from 1998 through 2000, along with the discharge diagnosis database maintained by the internal medicine service of the largest hospital in the system. The total number of persons prescribed statins was similar in 1998, 1999, and 2000. No cases of severe rhabdomyolysis occurred in patients taking other statins without gemfibrozil. Thus, the incidence of rhabdomyolysis in patients taking cerivastatin as their only lipid-lowering drug is striking compared with the incidence with other statins in the same population.

David J. Hyman, MD, MPH
Andrea Henry, PharmD
Addison Taylor, MD
Harris County Hospital District
Houston, TX 77030

Table. **Rhabdomyolysis in Patients Taking Cerivastatin***

Patient	Age, y	Ethnicity	Sex	Medical History	Medication	Peak CPK Level
1	61	Black	F	Diabetes mellitus, asthma, peptic ulcer disease, hypertensive neuropathy, gastric ulcer, psychological treatment, diverticula, hysterectomy	Azithromycin Z pak (9 weeks previously), risperidone, ianoprazole, vitamin E, conjugated estrogen, trezadone, amantadine, quetiapine, naproxen, albuterol, salmeterol, fluticasone (inhalation)	30 916
2	47	Hispanic	M	Smoker, unstable angina, acute MI, hypertensive neuropathy	Quinapril, nitroglycerin, raniditine, aspirin, clopidogrel, metroprolol	48 580
3	62	Black	F	Hypertensive neuropathy, intermediate coronary syndrome, atherosclerosis	Ibuprofen, $FeSO_4$, captopril	26 150
4	46	Black	M	Hypertensive neuropathy, asthma, intermediate coronary syndrome	Lanoprazole, celecoxib, sublingual nitroglycerin, felodipine, albuterol, atenolol, cyclobenzaprine, HCTZ	8410
5	56	Hispanic	F	Hypertensive neuropathy, diabetes, CVA, hypothyroidism, cellulitis	Insulin 70/30, aspirin, HCTZ, quinapril, felodipine, clopidogrel, levothyroxine, lanoprazole	34 420
6	72	Black	M	Carotid stenosis, stent–CVA, hypertensive neuropathy, prostatectomy, colon polyps	Amlodipine, enteric-coated aspirin, clopidogrel, famotidine, HCTZ, propoxyphene/acetaminophen, insulin	30 000

* CPK = creatine phosphokinase; CVA = cerebrovascular accident; F = female; HCTZ = hydrochlorothiazide; M = male; MI = myocardial infarction.

Gabapentin Toxicity Requiring Intubation in a Patient Receiving Long-Term Hemodialysis

TO THE EDITOR: Gabapentin, or 1-(aminomethyl) cyclohexaneacetic acid, is an antiepileptic medication used increasingly to treat neuropathies. It is highly lipid and water soluble and is excreted by the kidneys. Its half-life is 5 to 7 hours in healthy persons; in patients with renal failure, its half-life is 132 hours during nonhemodialysis periods and 3.8 hours during hemodialysis (1, 2). Thus far, massive gabapentin overdose has been associated with few adverse effects (3–5).

A 46-year-old woman with end-stage renal disease was admitted somnolent and hypoxic after taking multiple doses of gabapentin over 2 days without intervening hemodialysis. The patient was tremulous and hyperreflexic, with a Glasgow Coma Scale score of 8. Oxygen saturation was 80% on room air, and the patient was intubated. Gabapentin level was 22.6 μg/mL (normal range, 2.2 to 6.1 μg/mL) (2). Hemodialysis resulted in rapid clinical improvement and extubation. Three months later, the patient presented with similar symptoms after two extra doses of gabapentin and again required intubation. Hemodialysis was again associated with rapid improvement in mental status and rapid extubation.

Previous reports of gabapentin toxicity have emphasized the benign effects of elevated levels. In one report, a woman with end-stage renal disease and a gabapentin level of 85 μg/mL was only mildly confused (3). Another woman with an initial gabapentin level of 62 μg/mL was lethargic and improved quickly (4). A man who took massive doses of both valproate and gabapentin developed shock and coma that resolved with hemodialysis; symptoms were attributed to a valproate level of 1300 μg/mL (5). In contrast, our patient twice developed life-threatening symptoms that were temporally related to increased ingestion of gabapentin and resolved with hemodialysis. She was not taking valproate or other medications.

This case suggests that toxicity should be considered when mental status changes occur in patients who are taking gabapentin and have end-stage renal disease. Moreover, patients with end-stage renal disease should be cautioned about taking extra doses of this medication.

Heather Jones, MD
Elvira Aguila, MD
Harrison W. Farber, MD
Boston Medical Center
Boston, MA 02118

References

1. Wong MO, Eldon MA, Keane WF, Turck D, Bockbrader HN, Underwood BA, et al. Disposition of gabapentin in anuric subjects on hemodialysis. J Clin Pharmacol. 1995;35:622-6. [PMID: 7665723]

2. Blum RA, Comstock TJ, Sica DA, Schultz RW, Keller E, Reetze P, et al. Pharmacokinetics of gabapentin in subjects with various degrees of renal function. Clin Pharmacol Ther. 1994;56:154-9. [PMID: 8062491]

3. Verma A, St Clair EW, Radtke RA. A case of sustained massive gabapentin overdose without serious side effects. Ther Drug Monit. 1999;21:615-7. [PMID: 10604821]

4. Fischer JH, Barr AN, Rogers SL, Fischer PA, Trudeau VL. Lack of serious toxicity following gabapentin overdose. Neurology. 1994;44:982-3. [PMID: 8190316]

5. Fernandez MC, Walter FG, Kloster JC, Do SM, Brady LA, Villarin A, et al. Hemodialysis and hemoperfusion for treatment of valproic acid and gabapentin poisoning. Vet Human Toxicol. 1996;38:438-43. [PMID: 8948077]

Celecoxib-Induced Cholestatic Hepatotoxicity in a Patient with Cirrhosis

TO THE EDITOR: A 49-year-old man with alcoholic cirrhosis of the liver developed jaundice, fatigue, and choluria after he began taking celecoxib, 200 mg/d, for minor musculoskeletal pain. He took the drug for 15 days before presentation. The patient had a history of heavy alcohol consumption and had developed ascites 6 months earlier. At that time, chronic hepatic disease had been diagnosed, and viral serologic characteristics compatible with previous contact with hepatitis B virus had been detected. The patient had received althiazide, spironolactone, oxazepam, tiapride, lactitol, vitamins, and minerals and remained clinically well (anicteric, with no ascites or edema) during the subsequent months. He and his family reported no consumption of alcohol, changes in medication, or ingestion of toxic products since the previous visit.

At observation, the patient was markedly jaundiced and fatigued and had discrete hepatomegaly, with no other abnormalities. Laboratory tests showed a normal leukocyte formula, an aspartate aminotransferase level of 1.1 μkat/L (66 U/L), an alanine aminotransferase level of 0.8 μkat/L (49 U/L), a total bilirubin level of 540.3 μmol/L (31.6 mg/dL), a direct bilirubin level of 454.9 μmol/L (26.6 mg/dL), an alkaline phosphatase level of 3.42 μkat/L, γ-glutamyltransferase level of 0.92 μkat/L, an albumin level of 2.6 g/L, and a prothrombin rate of 58%. Imaging studies excluded space-occupying lesions or biliary duct dilation. The viral investigation confirmed the aspects already known. The biopsy showed aspects of cirrhosis, mononuclear infiltration of the portal triads, mild lobular activity, and marked hepatocellular cholestasis with the formation of bile plugs. No Mallory bodies or steatosis was seen. With suspension of celecoxib, bilirubin levels began to decrease. The patient was discharged on day 20 with a total bilirubin level of 290 μmol/L (17 mg/dL). After 3 months, he had a total bilirubin level of 102 μmol/L (6 mg/dL). One year later the patient is well, with a total bilirubin level of 44.5 μmol/L (2.6 mg/dL).

The patient's history strongly suggests celecoxib-induced cholestasis. Similar cases have been reported (1, 2), but allergy features were present, unlike in our patient. Further study will probably be necessary to understand the safety profile of celecoxib in patients with low hepatic reserve (3).

Paulo Alegria, MD
Luís Lebre, MD
Cristina Chagas, MD
Hospital de Santa Cruz
Carnaxide, Portugal

References

1. Galan MV, Gordon SC, Silverman AL. Celecoxib-induced cholestatic hepatitis [Letter]. Ann Intern Med. 2001;134:254. [PMID: 11177350]

2. Carrillo-Jimenez R, Nurnberger M. Celecoxib-induced acute pancreatitis and hepatitis: a case report. Arch Intern Med. 2000;160:553-4. [PMID: 10695699]

3. Davies NM, McLachlan AJ, Day RO, Williams KM. Clinical pharmacokinetics and pharmacodynamics of celecoxib: a selective cyclo-oxygenase-2 inhibitor. Clin Pharmacokinet. 2000;38:225-42. [PMID: 10749518]

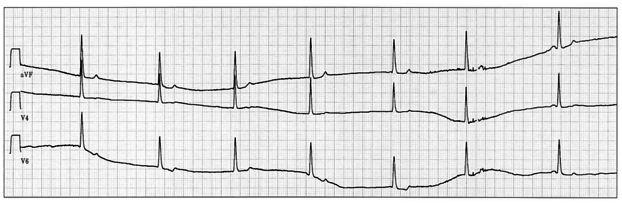
Adenosine and Aggrenox: A Hazardous Combination

Aggrenox (Boehringer Ingelheim, Ingelheim, Germany), a combination of low-dose aspirin and extended-release dipyridamole, is a new, effective medication for secondary stroke prevention (1). Patients taking Aggrenox may be referred for adenosine nuclear perfusion heart scans or could be considered for termination of supraventricular tachycardia with intravenous adenosine. Because dipyridamole inhibits cellular reuptake and breakdown of adenosine, the extended-release formulation may exaggerate the negative chronotropic and dromotropic effects of injected adenosine. Warning letters were recently published in the cardiology literature about the possible hazard associated with adenosine stress tests in patients taking Aggrenox (2, 3), but to our knowledge the actual occurrence of this type of drug interaction has not yet been reported.

A 79-year-old woman with diabetes mellitus, hypertensive heart disease, and a recent episode of transient ischemic attack was referred for adenosine nuclear perfusion heart scan because of atypical chest pain. The patient did not have a history of the sick sinus syndrome or atrioventricular block. At the time of the study, the patient's cardioactive medications included amlodipine, lisinopril, and twice-daily Aggrenox. The last Aggrenox dose was taken 12 hours before the heart scan. Intravenous adenosine was administered through a peripheral line at a rate of 140 μg/kg of body weight per minute. At 2 minutes into adenosine infusion, the patient became profoundly bradycardic, dizzy, and presyncopal. The electrocardiogram demonstrated prolonged sinus arrest with a slow, slightly irregular atrioventricular junctional escape rhythm at a rate of 36 beats per minute (**Figure**). The test was aborted. Intravenous theophylline was not given because the patient remained responsive. Two minutes after discontinuation of adenosine, a sinus rate of 60 beats per minute returned. A dobutamine nuclear heart scan was performed the same day without complication.

This case describes profound symptomatic bradycardia associated with adenosine administration in a patient taking Aggrenox. Adenosine itself may cause sinus bradycardia or atrioventricular block, but the incidence of clinically important bradycardia in patients who do not have the sick sinus syndrome or baseline conduction abnormality is very low. In a study of 1351 consecutive patients who underwent adenosine nuclear studies with two different infusion protocols, the adenosine infusion did not have to be suspended for sinus bradycardia or sinus arrest in any case (4). In another recent report of 600 consecutive patients, of whom 52% were 66 years of age or older, bradycardia requiring intervention or discontinuation of the adenosine infusion did not occur (5). At our institution, 1209 adenosine nuclear heart scans were performed during the past 3 years, but this was the first time that a study had to be aborted because of severe symptomatic bradycardia. We believe that in the presented case, prolonged sinus arrest was almost certainly due to the use of adenosine in a patient who was also taking Aggrenox.

Figure. **Electrocardiogram (leads aVF, V_4, and V_6; paper speed, 25 mm/s) recorded 2 minutes after initiation of adenosine drip in a patient taking Aggrenox.**



The rhythm strip demonstrates sinus arrest with a slow and slightly irregular atrioventricular junctional escape rhythm at a rate of 36 beats per minute.

All clinicians who order or perform stress tests or who treat patients with supraventricular tachycardia should be reminded of the potentially hazardous nature of the drug interaction between adenosine and Aggrenox. In patients taking Aggrenox, pharmacologic stress testing should be performed with dipyridamole or dobutamine. Adenosine heart scan should be done only after Aggrenox has been discontinued for several days. Similarly, the use of intravenous adenosine should be discouraged in patients with paroxysmal supraventricular tachycardia who are taking Aggrenox. Alternative treatments (for example, intravenous verapamil) should be considered. Intravenous theophylline, a competitive adenosine antagonist, is the drug of choice for sustained, profound bradycardia provoked by the combination of adenosine and Aggrenox.

Laszlo Littmann, MD
Justin D. Anderson, MD
Michael H. Monroe, MD
Carolinas Medical Center
Charlotte, NC 28232

References

1. Pettigrew LC. Antithrombotic drugs for secondary stroke prophylaxis. Pharmacotherapy. 2001;21:452-63. [PMID: 11310519]
2. Bokhari S. Warning to physicians performing pharmacologic stress tests [Letter]. J Nucl Cardiol. 2000;7:546. [PMID: 11083205]
3. Bergmann SR. Aggrenox and stress testing [Letter]. Circulation. 2001;103: E50. [PMID: 11222491]
4. O'Keefe JH Jr, Bateman TM, Handlin LR, Barnhart CS. Four- versus 6-minute infusion protocol for adenosine thallium-201 single photon emission computed tomography imaging. Am Heart J. 1995;129:482-7. [PMID: 7872175]
5. Alkoutami GS, Reeves WC, Movahed A. The frequency of atrioventricular block during adenosine stress testing in young, middle-aged, young-old, and old-old adults. Am J Geriatr Cardiol. 2001;10:159-61. [PMID: 11360841]

Quality of Care in Patients with Diabetes

TO THE EDITOR: The Board of Directors of the American Association of Clinical Endocrinologists (AACE) and the Trustees of the American College of Endocrinology (ACE) express their unanimous concern about conclusions reached by, and generalizations stemming from, the study by Greenfield and colleagues on profiling care by different groups of physicians (1). Their findings are of indeterminate validity because of multiple methodologic flaws. Although we applaud the authors for their enhanced statistical approach to assessing the quality of medical care received by persons with diabetes, they may have inadvertently introduced enough noise into the data to bias their results.

We are further disappointed that *Annals of Internal Medicine* distributed a Summary for Patients (2) that uncritically accepted the conclusions of the paper; assumed that these conclusions may be extrapolated to physician practices in general; and distributed the

findings as oversimplified "facts" to patients, payers, and the medical community at large. High stakes require strong evidence, and we do not see the present evidence as sufficiently strong to warrant the wording in the summary and the message that most consumers will likely take from it.

There are several problematic issues on which the paper's conclusions are based. First, classification of sites as generalist or specialist was in error. Some sites had data submitted from both endocrinologists and generalists, likely resulting in an unknown amount of group-effect dilution. A potentially influential proportion of multiphysician groups (for example, managed care plans) had patients who were at least partly co-managed by primary care physicians and endocrinologists. This type of classification error is serious and can result in fatally flawed data for group comparisons. Essentially, practice site and physician group (or "cluster") were commingled to an unknown degree for multiphysician sites. This issue is a serious threat to internal validity, the justifiable conclusions about the sample itself. Support for the lack of commingling of generalists and specialists beyond the investigators' verification of specialty was necessary, at least for the multiphysician sites.

Second, the American Diabetes Association's Provider Recognition Program (PRP), from which the data were derived, was not designed as a research study. A convenience sample of physicians and groups, rather than a random sample from a defined population, was used. Lack of randomization and the authors' lack of knowledge about actual selection of patients within each practice introduce many possible biases. Only two sites were audited, and it is impossible to know whether representative samples were used at audited or other sites. Since these were volunteer groups applying for certification or, in some cases, sites actively recruited for the program, statistical support rather than logic alone should have been used to justify the appropriateness of each group's representativeness within its classification as either generalist or specialist.

Third, no distinction was made between type 1 and type 2 diabetes mellitus. This is a serious clinical issue because excellence in outcomes may be much more difficult to achieve in patients with the former disorder. Given that this patient characteristic was not addressed in the initial case-mix adjustment, outcomes between groups cannot be justifiably compared.

Fourth, the clinical measures used as a basis of comparison were mixed. Some groups used glycosylated hemoglobin as an index of glycemic control, while others used fructosamine. Furthermore, annual monitoring does not provide the physician with the information needed to adjust the patient's therapy. Therefore, it provides at best a very crude "accountability measure" but is not a measure of the quality of care. In fact, the criteria chosen for "satisfactory" levels of care, hemoglobin A_{1c} levels of 0.10 (10%) and 0.08 (8%), are both well above the treatment goals of 0.65 (6.5%) that the ACE established in 2001. If these "satisfactory" levels were not the goals of both groups of physicians, the comparison of groups at these levels is inappropriate.

Fifth, Greenfield and colleagues' research design is a preexperimental static-group comparison (3, 4). For this type of design, generalizability (that is, external validity) is not warranted without evidence that the researchers' covariates removed the substantively important nontreatment (in this case, "specialist") differences between groups, especially when evidence showed non–treatment-related group differences. Such evidence should extend beyond a logical explanation through statistical testing. Greenfield and colleagues performed no such testing, so no generalizability beyond the sample is warranted on the basis of research design alone.

Sixth, tracing Greenfield and colleagues' sample size from the start of the study to the final analytic sample is difficult at best given the provided information. For example, 29 sites averaging 67 patients per site yields 1943 patients. Surveys were completed by 1258 patients (73.5%), yielding a before-survey sample of 1712 patients. What happened to the other 231 patients? Furthermore, how did the 1712 patients increase to 1750 in the tables? In short, sample size issues need more explanation.

Seventh, imputation of missing data on the 1750 patients in the tables needs more explanation than is given. The paper implies that data were imputed for 28.1% (1 − [1258/1750]) of patients. This degree of imputation usually decreases group differences when data from the combined groups are used. If imputation were done separately for each group, other issues would be encountered. Imputation issues are not satisfactorily explained in the paper. In addition, Greenfield and colleagues' Figure seems to show three outlying endocrinologist sites. This might be due to the issue of multiphysician sites, the issue of volunteerism, the self-reporting of specialty, or even clustered (for example, within-abstractor) data. Greenfield and colleagues do not provide enough information about this situation to classify it or to determine whether it is a problem.

Eighth, it is important to note that the original, uncorrected data showed a significant advantage from endocrinologist care in terms of hemoglobin A_{1c} level, patient satisfaction, and several other measures. When the data were corrected for case mix to compensate for differences in severity, duration, and other patient characteristics, there were still several statistically significant differences between generalists and specialists. Only when the data were given a second round of "adjustment" or "correction" for physician clustering did most of the differences become nonsignificant. This loss of statistically significant group differences may not have been due to "proper" clustering adjustment; rather, it might have been related to a loss in statistical power caused by the model itself, resulting in an insensitive study.

In view of the multiple methodologic problems, analysis of Greenfield and colleagues' data by an independent third party would be advisable. The authors should be requested to make their data available so that others can apply alternative statistical methods and evaluate the effect of case-mix adjustment and corrections for clustering of patients with physicians.

Previous studies have demonstrated that endocrinologists provide superior care for patients with diabetes (5–7), resulting in more favorable short-term and long-term outcomes. Endocrinologists serve as key leaders in the team effort for diabetes management in many different patient care settings. The AACE's Medical Guidelines for the Management of Diabetes Mellitus (8) serve as an important benchmark for achieving optimum care and reducing diabetic complications. Members of the AACE firmly support all efforts to improve the quality of life of persons with diabetes and applaud all efforts aimed at quality improvement and its appropriate assessment. We look forward to well-designed studies that will clarify our efforts to treat this serious disorder.

Rhoda H. Cobin, MD
American Association of Clinical Endocrinologists

Jacksonville, FL 32204

Helena W. Rodbard, MD
American College of Endocrinology
Jacksonville, FL 32204

References

1. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med. 2002;136:111-21. [PMID: 11790062]

2. Summaries for patients. Comparing the quality of diabetes care by generalists and specialists. Ann Intern Med. 2002;136:I42. [PMID: 11928735]

3. Campbell DT, Stanley JC. Experimental and Quasi-experimental Designs for Research. Boston: Houghton Mifflin; 1963.

4. Cook TD, Campbell DT. Design and Analysis Issues for Field Setting. Boston: Houghton Mifflin; 1979.

5. Levetan, CS, Salas JR, Wilets IF, Zumoff B, Impact of endocrine and diabetes team consultation on hospital length of stay for patients with diabetes. Am J Med. 1995;99: 22-8. [PMID: 7598138]

6. Levetan CS, Passaro MD, Jablonski KA, Ratner RE. Effect of physician specialty on outcomes in diabetic ketoacidosis. Diabetes Care. 1999;22:1790-5. [PMID: 10546009]

7. Hellman R, Regan J, Rosen H. Effect of intensive treatment of diabetes on the risk of death or renal failure in NIDDM and IDDM. Diabetes Care. 1997;20:258-64. [PMID: 9051368]

8. The American Association of Clinical Endocrinologists medical guidelines for the management of diabetes mellitus: the AACE system of intensive diabetes self-management. 2000 update. Endocrine Practice. 2000;6:42-84.