

LETTERS

Scoring the Quality of Clinical Trials

To the Editor: Dr Jüni and colleagues¹ compared 25 checklists from systematic reviews. We agree that readers should be critical of the heterogeneity of the content and results of checklists. Therefore, empirical studies in this field are useful. However, by using the same collection of checklists as Moher et al,² Jüni et al portray an unfair representation of the scientific development of research groups. Our list,³ which Jüni et al included in their analysis, was developed in 1990 and published in 1991. Thereafter, however, we have changed and hopefully improved our checklist, according to the new insights provided by Moher et al⁴ and others. This has resulted in an updated version of our checklist, which has been published in the method guidelines for systematic reviews within the Cochrane Back Review Group.⁵ The updated checklist has already been used in several protocols and reviews in the module of the Back Review Group, as well as in related journal articles.

We agree that, based on current knowledge, the use of checklists is problematic. However, we argue that replications of studies like that of Jüni et al should evaluate the latest versions of checklists used by the research groups at issue, not lists from the relatively early years of systematic review methodology. Otherwise the results of such empirical studies will be of predictably limited value.

Willem J. J. Assendelft, MD, PhD
University of Amsterdam
Amsterdam, the Netherlands

Bart W. Koes, PhD
Erasmus University
Rotterdam, the Netherlands

Maurits W. van Tulder, PhD
Lex M. Bouter, PhD

Institute for Research in Extramural Medicine, Vrije Universiteit
Amsterdam

1. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054-1060.
2. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials*. 1995;16:62-73.
3. Koes BW, Assendelft WJJ, van der Heijden GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilization for back and neck pain: a blinded review. *BMJ*. 1991;303:1298-1303.
4. Moher D, Jadad JR, Tugwell P. Assessing the quality of randomized controlled trials: current issues and future directions. *Int J Technol Assess Health Care*. 1996;12:195-208.
5. van Tulder MW, Assendelft WJJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for spinal disorders. *Spine*. 1997;22:2323-2330.

To the Editor: Dr Jüni and colleagues¹ illustrated the problem of rating the quality of clinical trials with summary scores for study quality. We have 3 comments. First, in 1 of the 25 scales tested, the original authors refused to define a cut point

for quality.² This was ignored by Jüni et al, who depicted a point the original authors merely used for illustration of a graph as a cut point. Within the context of acupuncture in chronic pain, this graph showed the increasing proportion of negative study results as the quality score increased. The graph encouraged a flexible definition of cut points.

Second, Jüni et al wrote, "The importance of individual quality domains and, possibly, the direction of potential biases associated with these domains will thus vary according to the context. . . ." However, in taking the 25 scales out of their highly variable contexts and applying them to deep vein thrombosis, they ignore their own advice.

Jüni et al also claim that "[r]elevant methodological aspects should be assessed individually and their influence on effect sizes explored." This leads to the conclusion that a relevant methodological aspect could have no influence on the effect sizes. In our opinion, such an aspect would be relevant if it is part of an a priori idea about what constitutes scientific proof in a given research context. When given a specific problem, scientists draw up a methodological recipe that, if followed, will lead to valid answers. The a priori ideas about which recipes may lead to meaningful answers are grounded in theory. However, theories themselves may be incomplete or wrong. In this case, how can we use meta-regression to test our theories? It is similarly unclear how one should deal with design components that are not associated with effect size. How many meta-regressions should indicate that, for example, intention-to-treat analysis does not influence effect sizes before we drop this aspect from our recipe?

Gerben ter Riet, MD, PhD
Pieter Leffers
Maurice Zeegers
Maastricht University
Maastricht, the Netherlands

1. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054-1060.
2. ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol*. 1990;43:1191-1199.

GUIDELINES FOR LETTERS. Letters discussing a recent *JAMA* article should be received within 4 weeks of the article's publication and should not exceed 400 words of text and 5 references. Letters reporting original research should not exceed 500 words and 6 references. All letters should include a word count. Letters must not duplicate other material published or submitted for publication. Letters will be published at the discretion of the editors as space permits and are subject to editing and abridgment. A signed statement for authorship criteria and responsibility, financial disclosure, copyright transfer, and acknowledgment is required for publication. Letters not meeting these specifications are generally not considered. Letters will not be returned unless specifically requested. Also see Instructions for Authors (January 5, 2000). Letters may be submitted by surface mail: Letters Editor, *JAMA*, 515 N State St, Chicago, IL 60610; e-mail: JAMA-letters@ama-assn.org; or fax (please also send a hard copy via surface mail): (312) 464-5824.

Letters Section Editors: Phil B. Fontanarosa, MD, Deputy Editor; Stephen J. Lurie, MD, PhD, Fishbein Fellow.

To the Editor: As a coauthor of the original meta-analysis on low-molecular-weight heparin (LMWH),¹ I have often referred to it when speaking or writing about the role of subjectivity in research.² I am therefore not entirely surprised by the finding of Dr Jüni and colleagues³ that varying judgments on the quality of studies can lead to diametrically opposite conclusions. Nevertheless, the solution to look at individual elements of quality—albeit only those that are judged important—will always remain an ad hoc decision of a meta-analyst. Moreover, the interpretation is 1-sided: we will only believe that quality has an influence when the higher-quality studies find a lower effect. The undisputed advantage of individual scoring is that one can discuss (and disagree) more clearly whether some element of quality is important in a particular set of trials.

Jüni et al conclude that different quality scores measure different things. Much like the decision by the individual which quality element to scrutinize, one might tell the reader what type of quality score one proposes. Our original quality score was concordant with scores that predominantly reflected internal validity and study design.^{1,3} For medical decision making, it is the quality of the research that counts. Other quality scores have other uses (eg, to judge readability). Of acute interest is the observation that one of the quality scores that maximally reverses our findings is a byproduct of a meta-analysis of homeopathy with a positive outcome,² which is scientifically impossible. Are we to infer that quality criteria that maximally suit homeopathy will also show a maximal benefit of LMWH?

Jüni et al show a strong influence of open outcome assessment in heparin trials, which they judge to be important. Their final verdict, based on their own selection of the best trials, is in the same direction as ours: the difference between the heparins is not one of clear therapeutic superiority. The same was concluded in a more recent meta-analysis that included more trials⁴ and in other recent overviews.⁵ This follow-up might be the criterion standard to decide which quality scores perform best. In this follow-up, additional knowledge about the mecha-

nism of action of diverse heparins has also helped decide which trial results are more credible.^{2,5} Rightly, Jüni et al do not offer any absolute standard to judge the quality of studies. Judgment remains particular to particular trials and particular situations. In that respect, randomized trials are not different than other study designs.

Jan P. Vandenbroucke, MD, PhD
Leiden University Medical Center
Leiden, the Netherlands

1. Nurmohamed MT, Rosendaal FR, Büller HR, et al. Low-molecular-weight heparin versus standard heparin in general and orthopaedic surgery: a meta-analysis. *Lancet*. 1992;340:152-156.
2. Vandenbroucke JP. Medical journals and the shaping of medical knowledge. *Lancet*. 1998;352:2001-2006.
3. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282:1054-1060.
4. Koch A, Bouges S, Ziegler S, Dinkel H, Dauares JP, Victor N. Low-molecular-weight heparin and unfractionated heparin in thrombosis prophylaxis after major surgical intervention: update of previous meta-analysis. *Br J Surg*. 1997;332:1330-1335.
5. Thomas DP. Does low-molecular-weight heparin cause less bleeding? *Thromb Haemost*. 1997;78:1422-1425.

In Reply: In our study, we repeated Nurmohamed and colleagues¹ meta-analysis of 17 trials of LMWH for prevention of deep vein thrombosis, using 25 scales to identify high-quality trials. We included all widely used scales but none produced summary scores that were significantly associated with treatment effects. Dr ter Riet and colleagues argue that our results could be explained by the inclusion of scales that were developed for specific contexts. Dr Assendelft and colleagues imply that our findings would have been different if we had used updated versions of scales. They argue that we should have used the 1997 version of their scale,² which gives more weight to the internal validity of trials, rather than the version published in 1991.³ The revised scale is recommended by the Cochrane Back Review Group,² but only 2 of the 8 reviews published by this group in the Cochrane Library have actually used this scale. This observation, and the large number of published scales, underscore the uncertainty that exists in this area.

We addressed ter Riet and colleagues' criticism by excluding the 11 scales that were only used in particular contexts. Results were not materially altered (data available from the authors). We examined Assendelft and colleagues' claim by using both the old and the updated versions^{2,3} of their scale (TABLE). Although classification of some trials changed, point estimates from high-quality trials were similar to the overall estimate, and summary scores were not significantly associated with treatment effects. This is not surprising considering that even scales that exclusively assess items related to the internal validity of trials⁴ failed to detect an association with treatment effects. Quality scales, including the updated version of Assendelft and colleagues' scale, consistently failed to show an association with treatment effects. However, when assessing key components of methodological quality individually, we found that lack of blinding of outcome assessments influenced effect size, with the effect of LMWH on average being exaggerated by 35% ($P < .05$). Clearly, when scoring the qual-

Table. Combined Results for All 17 Trials Comparing Low-Molecular-Weight Heparin With Standard Heparin*

	No. of Trials	Deep Vein Thrombosis		P Value
		RR (95% CI)	Ratio of RRs (95% CI)	
All trials	17	0.79 (0.65-0.95)
Stratified by quality†				
Original scale ⁴				
High quality	12	0.74 (0.61-0.91)	1.00 (Referent)	.12
Low quality	5	1.13 (0.70-1.82)	1.53 (0.90-2.59)	
Revised scale ³				
High quality	10	0.85 (0.68-1.08)	1.00 (Referent)	.28
Low quality	7	0.68 (0.50-0.93)	0.80 (0.54-1.19)	

*Results from sensitivity analyses dividing trials in high-quality and low-quality strata, using the original and a revised version of Assendelft et al's quality scale. RR indicates relative risk; CI, confidence interval; and ellipses, data not available.

†Threshold for high quality set at 50% of maximum score, as proposed by the authors of the scale.^{2,3}

ity of the 17 trials with scales, the points attributed to blinding of outcome assessment (the factor that matters in this particular meta-analysis¹) are diluted by other aspects of methodological or reporting quality that are of limited importance in this situation. As pointed out by ter Riet and colleagues, not all aspects that are relevant on theoretical grounds will matter in a given context.

We agree with Dr Vandembroucke that scales should not be used to identify trials of low or high quality in a given meta-analysis. Rather, the relevant methodological aspects should be identified a priori and checked individually. Oxman⁵ has drawn an analogy with aviation to illustrate the importance of checklists. Before takeoff, pilots use checklists to ensure that all vital systems of the aircraft work properly. Unlike many meta-analysts, they do not rely on scales and decide to take off if a summary score of 50% or more is reached. Our empirical study and theoretical considerations⁴ indicate that what would certainly be disastrous in aviation is also a hazardous approach for assessing the quality of trials for systematic reviews and meta-analyses.

Peter Jüni, MD
Matthias Egger, MD, MSc
MRC HSRC, Department of Social Medicine
Bristol, England

1. Nurmohamed MT, Rosendaal FR, Büller HR, et al. Low-molecular-weight heparin versus standard heparin in general and orthopaedic surgery: a meta-analysis. *Lancet*. 1992;340:152-156.
2. van Tulder MW, Assendelft WJJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for spinal disorders. *Spine*. 1997;22:2323-2330.
3. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ*. 1991;303:1298-1303.
4. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG. *Systematic Reviews in Health Care: Meta-Analysis in Context*. London, England: BMJ Books. In press.
5. Oxman AD. Checklists for review articles. *BMJ*. 1994;309:648-651.

Sources of Bias in the Economic Analysis of New Drugs

To the Editor: Mr Friedberg and colleagues¹ found that studies sponsored by drug companies were less likely to report unfavorable results than studies funded by nonprofit organizations. The authors posited several alternative explanations for their findings, including publication bias, unconscious bias because of relationships between the investigator and the pharmaceutical company, early screening of studies by drug firms, and direct company involvement with the study design.

As noted in the accompanying editorial,² the conflict of interest hypothesis would be strengthened by a study design comparing several economic evaluations, from different funding sources, of a single agent. However, the design of the study by Friedberg et al allows alternative explanations to be considered. First, many pharmaceutical companies conduct internal economic evaluations throughout the development of a drug product. Because reimbursement for a drug product is often dependent on positive economic findings,³ otherwise promis-

ing products may be terminated during development by the drug company. Hence, drugs that are marketed are less likely to receive an unfavorable economic evaluation, regardless of the funding source. Second, pharmaceutical companies fund the majority of the clinical and economic studies conducted during drug development. Because marketed products have demonstrated positive clinical findings during development and because these findings often lead to positive economic benefits, it should not be surprising that industry-sponsored economic analyses are more frequently favorable. Third, because drug companies are commercial enterprises, it is appropriate for them to allocate their resources to those activities that promise the best return. It is in their interest to subject operating hypotheses to rigorous economic analyses by funding external studies. None of these explanations are associated with scientific bias in terms of an external researcher's conflict of interest.

Although it is easy to allege scientific bias arising from a conflict of interest,⁴ proving such a bias is difficult. Indeed, Friedberg et al acknowledged that they did not identify bias in individual studies. Rather than supporting a conflict of interest hypothesis, this study may simply indicate that pharmaceutical firms are exhibiting rational economic behavior in their development of new products.

Jean Paul Gagnon, PhD
for the members of the Health Outcomes Committee
Pharmaceutical Research and Manufacturers of America
Washington, DC

1. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA*. 1999; 282:1453-1457.
2. Krimsky S. Conflict of interest and cost-effectiveness analysis. *JAMA*. 1999; 282:1474-1475.
3. Langley PC. Meeting the information needs of drug purchasers: the evolution of formulary submission guidelines. *Clin Ther*. 1999;21:768-787.
4. Rothman KJ. Conflict of interest: the new McCarthyism in science. *JAMA*. 1993; 269:2782-2784.

To the Editor: Mr Friedberg and colleagues¹ have analyzed a sample of published pharmacoeconomic studies and found that "pharmaceutical company-sponsored studies were less likely than nonprofit sponsored studies to report unfavorable qualitative conclusions." It is important to notice that they identified a publication bias, not a study bias. In fact they fail to find biases in individual studies, meaning that the peer review process work reasonably well and that individual published studies are generally reliable. However, it is also important to encourage publication of negative results. We recently published a study with negative results² despite reservations from the referees. We felt that they would have been more enthusiastic about positive results. Clinicians may also fear that reporting negative cost-effectiveness results will reinforce nonmedical limitations to their autonomy to prescribe.

It is obviously difficult to compel companies to support and publish studies showing that their products are not cost-effective. Because pharmacoeconomic studies are not the only factor in medical decision making, lack of positive data is not

necessarily an indication that the drug has no value. The solution probably lies in a more systematic use of pharmaco-economic data for treatment guidelines or for decisions about reimbursement and pricing, in both private and public health care settings. This would not guarantee that industry-sponsored negative results will be reported more frequently, but would allow more appropriate interpretation of unreported results.

Claude Le Pen, PhD
University Paris-Dauphine
Paris, France

1. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analysis of new drugs used in oncology. *JAMA*. 1999; 282:1453-1457.
2. Woronoff-Lemsi MC, Arveux P, Limat S, Morel P, Le Pen C, Cahn JY. Erythropoietin and preoperative autologous blood donation in the prevention of hepatitis C infection: necessity or luxury? *Transfusion*. 1999;39:933-937.

To the Editor: Mr Friedberg and colleagues¹ take a systematic approach to identifying the characteristics of published cost studies. However, relying on study sponsorship as the dominant criterion for evaluating potential conflict of interest may be too simplistic. A more thorough review of methods, data inputs, assumptions, and publication sources and their impact on study conclusions is warranted. For example, prospective, randomized clinical studies (single center or multicenter) using actual use and cost data may be less biased than studies developed using decision analysis models. Additionally, some of the studies included in the analysis were not published in peer-reviewed journals. To combine heterogeneous samples in drawing conclusions is improper both in clinical research and the evaluation of conflict of interest.

Furthermore, in more than 10% of cases in which authors were contacted regarding potential conflicts, authors and/or publishers failed to report industry relationships. Therefore, it can be difficult for readers to fully understand potential sources of conflict of interest unless disclosure and reporting policies are standardized and enforced.

In evaluating the potential for biases, it may be more prudent for readers to evaluate clinical and economic literature across several criteria including sponsorship, author affiliation, journal type, study methods and assumptions and the extent to which authors' conclusions and statements are supported by the data presented. Rather than focusing solely on pharmaceutical industry sponsorship, perhaps the best advice should be *caveat lector* (let the reader beware).

Mary M. Prendergast, MBA
Fujisawa Healthcare Inc
Deerfield, Ill

1. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA*. 1999; 282:1453-1457.

To the Editor: The article by Mr Friedberg and colleagues¹ includes a statement that the sponsor had a contractual right to review and comment on manuscripts and abstracts prior to submission. However, there is no indication of whether such

review was made, whether any comments were offered and, if so, whether they influenced the final manuscript. Disclosure should extend beyond the information that a review is possible and include the facts concerning the review, if one took place, and the influence, if any, that the review had on the manuscript. A separate question is whether investigators should sign contracts that allow sponsors the right to review and comment on data prior to peer review for publication.

Herbert Rakatansky, MD
Council on Ethical and Judicial Affairs
American Medical Association
Chicago, Ill

1. Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA*. 1999; 282:1453-1457.

In Reply: We agree with Dr Gagnon that it is only rational for pharmaceutical firms to support positive economic analyses, although that does not preclude conflict of interest. It should also be noted that the perspective of our study is from the scientific community, not from the pharmaceutical industry. We found that fewer studies with negative results are sponsored by pharmaceutical firms, which leads to a bias in the available economic literature for these drugs. As Dr Le Pen notes, pharmaco-economic research (and medical research in general) would benefit from the publication of studies with both positive and negative results.

Ms Prendergast points out that the issues of bias are complex and include many factors in addition to sponsorship source, including methods, assumptions, and publication source. We hope that pharmaceutical companies with outcomes programs will support efforts to expand databases of cost-effectiveness studies. This would allow for multivariate analyses related to factors influencing study outcome.

In response to Dr Rakatansky's concern, it is typical for pharmaceutical companies to request a review period to protect the proprietary interests of their products. In this case, the sponsor had 60 days to review our manuscript and offered no comments or suggestions.

Mark Friedberg, BA
Northwestern University
Chicago, Ill
Bernie Saffran, PhD
Swarthmore College
Swarthmore, Pa
Tammy Stinson, MS
Charles Bennett, MD, PhD
Lakeside Veterans Affairs Medical Center
Chicago

Bisphosphonate Therapy and Vascular Calcification

To the Editor: The 3-year, randomized controlled trial reported by Dr Harris and colleagues¹ demonstrated that rise-

©2000 American Medical Association. All rights reserved.

dronate increased bone mineral density and decreased fracture incidence in postmenopausal women with established osteoporosis. Although the overall safety profile was similar in the placebo and risedronate groups, serious adverse events were more common in the treatment group.

Risedronate inhibits osteoclast-mediated bone resorption.² Vascular calcifications have similarities to bone, and macrophages in the vessel wall have both osteoclastic and osteoblastic potential.³ It is possible that subtle changes in vascular calcification occur during bisphosphonate treatment for postmenopausal osteoporosis.⁴

Postmenopausal women are at higher risk for cardiovascular disease.⁵ Changes in vascular wall calcification may render some plaques more prone to rupture and lead to an increased risk of cardiovascular events in this population. Perhaps the authors can provide information on the proportion of cardiovascular events in the treatment vs placebo groups. If there was a higher cumulative incidence of cardiovascular events in the treatment group, this would cast doubt on the long-term safety of bisphosphonate therapy in postmenopausal women.

Mark R. Goldstein, MD
West Chester, Pa

1. Harris ST, Watts NB, Genant HK, et al, for the Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. *JAMA*. 1999;282:1344-1352.
2. *Physicians' Desk Reference*. 53rd ed. Montvale, NJ: Medical Economics Books; 1999:2531-2533.
3. Demer LL. A skeleton in the atherosclerosis closet. *Circulation*. 1995;92:2029-2032.
4. Goldstein MR. Long-term therapy for postmenopausal osteoporosis: stronger bones but weaker arteries? *Circulation*. 1999;100:446-477.
5. Mosca L, Manson JE, Sutherland SE, et al. Cardiovascular disease in women: a statement for healthcare professionals from the American Heart Association. *Circulation*. 1997;96:2468-2482.

In Reply: Serious events reported in this study were defined in accordance with regulatory guidelines.¹ As noted in Table 3 of our article, the overall incidence of serious adverse events reported was similar between the placebo (27%) and 5-mg risedronate (29%) groups.

We performed a safety analysis focusing on major cardiovascular events reported during the study. Based on this analysis, the most common events occurred at similar incidences between treatment groups. Cardiovascular events reported at a cumulative incidence of more than 1% were right heart failure (placebo, 1.6%; risedronate, 1.8%), myocardial infarction (placebo, 1.6%; risedronate, 1.5%), and cerebrovascular accident (placebo, 1.0%; risedronate, 1.4%).

In addition, the study's sponsor (Procter & Gamble, Cincinnati, Ohio) has reviewed cardiovascular events reported in a larger database from trials evaluating risedronate in the treatment or prevention of postmenopausal and corticosteroid-induced osteoporosis. This analysis included 4878 patients treated with placebo and 4846 patients treated with risedronate for up to 3 years. Patients also received elemental calcium, 1 g/d, in most studies; cholecalciferol was provided if base-

line levels were less than the lower limit of normal. The cumulative incidence of all cardiovascular events reported was similar between groups, further supporting the conclusion that risedronate treatment does not increase the occurrence of cardiovascular events.

Dr Goldstein points to potential effects of bisphosphonates on inhibiting macrophages in blood vessels with resulting changes in vascular wall calcification. The concentrations of risedronate required to directly inhibit serum macrophages are not achieved with clinical dosages.² While our study was not designed to examine the effects of risedronate on vessel wall calcification, the risedronate clinical adverse event database does not suggest that risedronate therapy increases the risk of cardiovascular events.

Steven T. Harris, MD
University of California, San Francisco

1. IND safety reports. In: *Code of Federal Regulations: Food and Drugs*. 1996: 82-84. 21 CFR §312.32.
2. Luckman SP, Coxon FP, Ebetino FH, Russell RGG, Rogers MJ. Heterocycle-containing bisphosphonates cause apoptosis and inhibit bone resorption by preventing protein prenylation: evidence from structure-activity relationships in J774 macrophages. *J Bone Miner Res*. 1998;13:1668-1678.

Vitamin D Deficiency in Women With Hip Fracture

To the Editor: The study by Dr LeBoff and colleagues¹ further confirms the evidence that many older people have vitamin D deficiencies and are likely to benefit by its correction.^{2,3} However, their data do not prove that such deficiency is the primary, or even a major, factor in hip fractures in the elderly.

The large age difference between the hip fracture patients and elective control patients presents problems with evaluation for which statistical analysis cannot readily adjust. The authors properly excluded patients with obvious comorbid medical conditions, but there are important comorbidities that are difficult to ascertain, especially in a presurgical setting. Many of these can be important risk factors for falls (eg, minor gait or balance difficulties, diminished vision, weaker musculature, arthritic stiffness and pain, and mild dementia) and all such conditions were surely of higher prevalence in the (mean age) 78-year-old patients in the fracture group than in the 64- and 70-year-old control patients. The presence of such disabilities tends to be confirmed by the markedly reduced patterns of prior physical activity in the fracture patients—1.5 hours weekly vs 16 and 20 hours in control patients. Such inactivity is in itself a potential risk factor for falls and fractures.

One might also consider the implications of the much lower body fat content of the fracture patients, a mean of 32% compared with 42% and 47% in controls. The lower amounts of body fat, as well as the lower serum albumin levels, further suggest higher prevalence of premorbid status in these older women, with an associated potential for falls. There have also been some suggestions that external hip pads help to prevent

fractures and subcutaneous hip-buttock fat “pads” appear similarly protective.

Gerson T. Lesser, MD
The Jewish Home and Hospital
New York, NY

1. LeBoff MS, Kohlmeier L, Hurwitz S, Franklin J, Wright J, Glowacki J. Occult vitamin D deficiency in postmenopausal US women with acute hip fracture. *JAMA*. 1999;281:1505-1511.
2. Gloth FM III, Tobin JD. Vitamin D deficiency in older people. *J Am Geriatr Soc*. 1995;43:822-828.
3. Utiger RD. The need for more vitamin D. *N Engl J Med*. 1998;338:828-829.
4. Meyer HE, Tverdal A, Falch JA. Body height, body mass index, and fatal hip fractures: 16 years follow-up of 674,000 Norwegian men and women. *Epidemiology*. 1995;6:299-305.

In Reply: Vitamin D deficiency has been documented in older patients in nursing homes and residential communities in the United States, but there are few data available on vitamin D levels in women with acute osteoporotic fractures. In our study we found significant differences in vitamin D levels in the women with hip fractures and no known secondary comorbid medical conditions. In contrast, women admitted for elective joint replacement did not show vitamin D deficiency. We rigorously excluded patients with a multitude of risk factors that might predispose to osteoporosis, as indicated in Table 1 of our article. As shown in Figure 1 of our article, we excluded 190 women with comorbid medical conditions or those taking medications that could affect bone. In contrast to the vitamin D deficiency among women with hip fractures, we found that 17% of ambulatory women with osteoporosis by bone density criteria in an osteoporosis program had vitamin D deficiency.¹ While many women fall without adverse skeletal effects, fractures are more likely to occur in those who have a low bone density.

As stated by Dr Lesser, physical activity was less frequent in the women with hip fractures. Physical inactivity is a risk factor for osteoporosis and predisposes to reduced strength. Furthermore, the reduced physical activity may be associated with decreased sun exposure and reduced activation of vitamin D in the skin. Poor health and frailty are risk factors for osteoporosis, although we rigorously excluded patients who had known comorbid medical conditions.² We did not evaluate gait difficulties after the hip fracture. With reference to the percentage of total body fat content in the women with hip fractures, their mean [SD] percentage of body fat (32.1 [10.6]; Table 2) was in the normal range (33.8 [9]) for postmenopausal women aged 70 to 79 years.³ Thus, although our study showed that women with hip fractures had less body fat, the body fat content of the women with hip fractures was consistent with other studies of body fat in postmenopausal women.⁴ Furthermore, the mean albumin level was in the low normal range (Table 3) as indicated in our article. The lower albumin levels in the women with hip fractures may reflect poorer nutritional status in general and is a predictor of vitamin D deficiency.

We concur with Lesser that external hip pads may reduce fractures. In the population of community-dwelling women, however, many would need to wear hip pads to reduce the incidence of hip fracture. Because vitamin D deficiency is pre-

ventable and vitamin D repletion reduces the risk of hip fracture, increased public awareness of good vitamin D nutrition in postmenopausal women is of paramount importance.

Meryl S. LeBoff, MD
Shelley Hurwitz, PhD
Jennifer Franklin
John Wright
Julie Glowacki, PhD
Brigham and Women's Hospital
Boston, Mass
Lynn Kohlmeier, MD
Washington Medical Group
Fremont, Calif

1. Haden ST, Fuleihan GE, Angell JE, Cotran NM, LeBoff MS. Calcidiol and PTH levels in women attending an osteoporosis program. *Calcif Tissue Int*. 1999;64:275-279.
2. Cummings SR, Nevitt MC, Browner WS, et al. Risk factors for hip fracture in white women. *N Engl J Med*. 1995;332:767-773.
3. Svendsen OL, Hassager C, Christiansen C. Age- and menopause-associated variations in body composition and fat distribution in healthy women as measured by dual-energy x-ray absorptiometry. *Metabolism*. 1995;44:369-373.
4. Harris SS, Dawson-Hughes B. Weight, body composition, and bone density in postmenopausal women. *Calcif Tissue Int*. 1996;59:428-432.

Tobacco Dependence Curricula in Medical Schools

To the Editor: Dr Ferry and colleagues¹ present results of a national survey of the content of tobacco dependence curricula in US medical schools. Their data show that training of US medical graduates in the treatment of tobacco dependence is inadequate, and they report that clinical curricula involving smoking cessation techniques with patients and evaluation of student performance are seriously deficient.

In 1991, the Liaison Committee on Medical Education formally incorporated performance-based assessment into the accreditation standards of medical schools.² We believe, based on the overwhelming research evidence,³ that a critical part of the medical school curriculum on smoking cessation should be performance assessment of the clinical techniques. Key elements to guide the development of performance-based assessment are (1) integration of 2 or more new basic learned capabilities, (2) observed behaviors, (3) relevant clinical tasks, and (4) problem content at an appropriate level for medical students.⁴ We feel that clinical techniques in smoking cessation fit these criteria.

The curricula on tobacco dependence and clinical treatment at the Michigan State University College of Human Medicine include all reported content areas¹ with more than 10 hours of required time from years 1 through 3 of medical school. The main focus is a required assessment of smoking cessation counseling techniques during the family practice clerkship in the third year, which is standardized across 6 community sites. Students receive 10% of their final examination grade in the clerkship based on the performance assessment, and those who fail the assessment must remediate it or fail the clerkship. Our 1998 cohort of 93 students was videotaped during a 10-minute coun-

seling session with simulated patients. Two independent observers rated the sessions using a 10-point checklist based on the Agency for Health Care Policy and Research (now called the Agency for Healthcare Research and Quality [AHRQ]) criteria. Student scores ranged from 6.5 to 10. Only 10% of the scores were below 8, with a high level of interobserver agreement ($r = 0.847$).

Our data show that students can readily learn cessation counseling skills and such performance can be reliably evaluated.

William Wadland, MD, MS
Carole Keefe, PhD
Margaret Thompson, MD
Mary Noel, PhD, RD
Michigan State University
College of Human Medicine
East Lansing

1. Ferry HF, Grissino LM, Runfola PS. Tobacco dependence curricula in US undergraduate medical education. *JAMA*. 1999;282:825-829.
2. Liaison Committee on Medical Education. *Functions and Structure of a Medical School*. Chicago, Ill: Council on Medical Education, American Medical Association; 1991:14.
3. US Department of Health and Human Services. *Smoking Cessation, Clinical Practice Guideline No. 18*. Rockville, Md: US Dept of Health and Human Services, Agency for Health Care Policy and Research; 1996. AHCPR publication 96-0692.
4. Mavis BE, Henry RC, Ogle KS, Hoppe RB. The emperor's new clothes: the OSCE reassessed. *Acad Med*. 1996;71:447-453.

In Reply: While conducting our survey, we learned from faculty at several medical schools that clinical training and evaluation of nicotine dependence is most commonly incorporated in the primary care clerkships, as described by Dr Wadland and colleagues. Objective structured clinical examinations (OSCEs) are increasingly used in the third and fourth years to evaluate mastery of interviewing, physical diagnosis, and counseling skills. We know of a few medical schools that currently incorporate nicotine dependence treatment into their OSCE panel of cases, similar to those at the Michigan State University College of Human Medicine. Their encouraging outcome is consistent with other studies that show that medical students can develop skills in counseling nicotine-dependent patients.^{1,2} Unfortunately, schools with such commitments to nicotine dependence education remain in the minority.

Since our 1997 survey, other schools may have improved their nicotine dependence curricula. We are soliciting any teaching material, (eg, curriculum goals and objectives, OSCE case descriptions, videotapes, teaching manuals, handouts, interactive computer training program, role playing, problem-based learning, and evaluation tools) to create a syllabus of the best examples for distribution to all medical school faculty, and would be happy to receive any such material.

The release of the revised AHRQ clinical practice guidelines for smoking cessation in 2000 will be an opportunity for medical schools to reevaluate their basic and clinical science curriculum for the content areas suggested as requisite for all physicians.³ If the clinical training skills recommended by the AHRQ panel are followed, the next survey of medical school curricula would ideally find no medical school reporting less

than 1 hour of nicotine dependence training in the clinical years. In our survey, this minimal level was not met in 46.6% of the third-year curricula and in 79.3% of fourth-year curricula.

The most costly and lethal health behavior in the United States is still overlooked in most medical school curricula. The simple inclusion of didactic presentations, without evidence that students can demonstrate skills and application of the knowledge, is not the answer. Curricula like those described by Wadland and colleagues demonstrate the effective training physicians will need for nicotine dependence interventions in the 21st century.

Linda Hyder Ferry, MD, MPH
Linda M. Grissino, MD, MPH
Loma Linda University School of Medicine
Loma Linda, Calif

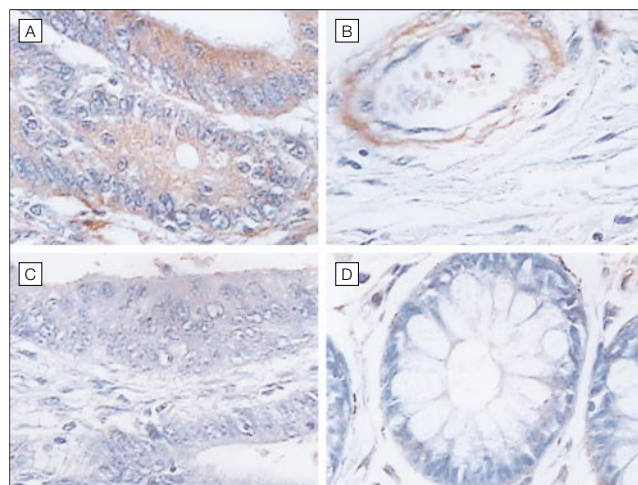
1. Coultas DB, Klecan DA, Whitten RM, et al. Training medical students in smoking-cessation counseling. *Acad Med*. 1994;69:S48-S50.
2. Allen SS, Bland CH, Dawson SJ. A mini-workshop to train medical students to use a patient-centered approach to smoking cessation. *Am J Prev Med*. 1990;6:28-33.
3. US Department of Health and Human Services. *Smoking Cessation, Clinical Practice Guideline No. 18*. Rockville, Md: US Dept of Health and Human Services, Agency for Health Care Policy and Research; 1996. AHCPR publication 96-0692.

CORRECTIONS

Omission: In the Preliminary Communication entitled "The Relationship Between Cyclooxygenase-2 Expression and Colorectal Cancer" in the October 6, 1999, issue of THE JOURNAL (1999;282:1254-1257), Ms Theresa Keane was omitted from the acknowledgment. The authors thank Ms Keane for her technical assistance.

Incorrect Color Reproduction: In the same Preliminary Communication, the colors of the 4 photomicrographs in Figure 1 on page 1255 were incorrectly reproduced. The correct image is shown below.

Figure 1. Cells Stained for Cyclooxygenase-2 (COX-2)



A, COX-2 expression in colon cancer epithelial cells; B, COX-2 expression in tumor vascular endothelium and stroma; C, inhibition of COX staining in colon cancer epithelial cells after preabsorption of anti-COX-2 antibody with COX-2 polypeptide; and D, normal colonic epithelium cells. COX-2 staining appears brown. Following antibody incubation, color was developed by immersion of the sections in diaminobenzidine tetrahydrochloride/hydrogen peroxide solution for 2 minutes. Sections were counterstained with hematoxylin (original magnification $\times 40$).