Sertraline for Treatment of Depression in Acute Coronary Syndromes

To the Editor: Dr Glassman and colleagues1 found that sertraline treatment was more efficacious than placebo in treating depression in this population. No information, however, is provided regarding procedures to ensure the blindness of ratings on the depression rating scales. Because those who rated depression also monitored adverse events, such as nausea and diarrhea, it cannot be assumed that they remained blinded to treatment assignment. However, approaches exist to protect blindness of ratings on the depression rating scales. Thus, sertraline may produce an increase, a decrease, or no difference in cardiac hospitalization rates. A larger trial, with more meaningful end points, is necessary to assess the safety of this intervention.

The authors found that sertraline was more efficacious than placebo in treating depression in this population. No information, however, is provided regarding procedures to ensure the blindness of ratings on the depression rating scales. Because those who rated depression also monitored adverse events, such as nausea and diarrhea, it cannot be assumed that they remained blinded to treatment assignment. However, approaches exist to protect blindness of ratings on the antidepressant trials. The authors did centralize the rating of multiple-gated acquisition scans, a computerized and relatively objective measure, to protect blindness of raters and to reduce interrater variability, but did not take similar precautions for the measurement of depression. Independent raters, ie, those not involved in other assessments or the clinical care of participants, can administer depression rating scales centrally by telephone. These assessments can be recorded to allow for reliability assessments. This approach is essential for reducing bias. Finally, the high rate of patients who discontinued the trial (28.5% in the sertraline group, and 25.1% in the placebo group) and the use of “last observation carried forward” analyses also threaten the validity of the results.

François Lespérance, MD
Department of Psychiatry
University of Montreal
Montreal, Quebec
Nancy Frasure-Smith, PhD
Department of Psychiatry
McGill University
Montreal


©2002 American Medical Association. All rights reserved.
blinding due to adverse events. We disagree that independent raters, rather than treating physicians, are better to ensure blindness. Such concerns are unrealistic in this population. The average patient receiving placebo was receiving 11 other medications as well, and thus the physician was often unable to determine which medication was causing a patient’s complaint.

Lespérance and Frasure-Smith are also concerned about the lower rate of adverse medical events in the patients treated with sertraline. We agree that our data do not prove that treatment of depression with sertraline reduces the associated cardiovascular event risk in acute coronary syndromes. Demonstration of risk reduction would have profound implications for clinical practice and should be based on a properly powered, large clinical trial.

Dr Steinfeld would consider an LVEF change of greater than 2% clinically significant, while we selected 5%. He also prefers to use absolute risk rather than our use of relative risk, but still concludes that “the increase in risk with sertraline, if any, is likely to be small.” We agree with his computations and his conclusions.

Alexander H. Glassman, MD
Department of Clinical Psychopharmacology
New York State Psychiatric Institute
New York
Christopher M. O’Connor, MD
Department of Cardiology
Duke University Medical Center
Durham, NC

**Counting Deaths Due to Medical Errors**

To the Editor: In their Controversies article about medical errors, Leape and colleagues praise the safety record of anesthesia: “Everyone . . . agrees that the current practice of anesthesia provides an outstanding example of how a high level of safety can be achieved in health care. Anesthesia is the only system in health care that begins to approach the vaunted ‘six sigma’ level of perfection that other industries strive for. Mortality from elective anesthesia has declined 10-fold in the past several decades as the result of a concerted effort to improve safety.” Similarly, the Institute of Medicine (IOM) report To Err Is Human: Building a Safer Health System stated: “Studies . . . indicate that, today, anesthesia mortality rates are about one death per 200,000-300,000 anesthetics administered, compared with two deaths per 10,000 anesthetics in the early 1980s.”

The improvement has actually been considerably less than those reports suggest. Based on 72,959 anesthetics administered over the past 10 years (1989-1999), Newland et al recently reported an incidence of cardiac arrest attributable to anesthesia of 0.69 per 10,000 anesthetics and a death rate due to anesthesia-attributable cardiac arrest of 0.55 per 10,000. Biboulet et al reported 11 cardiac arrests in 101,769 anesthetics administered during 1989-1993 (1.1 per 10,000) and an anesthesia-related mortality rate of 0.62 per 10,000 elective anesthetics. Similarly, Lagasse found anesthesia-related mortality rates of 1 per 12,641 anesthetics (0.79 per 10,000) based on 37,924 cases (1992-1994) and 1 per 13,322 (0.75 per 10,000) based on 146,548 anesthetics (1995-1999) at 2 medical centers.

From these studies, the current death rate related to anesthesia care appears to be about 1 in 15,000 anesthetics, a 3-fold improvement over the past several decades. If 20 million anesthetics are performed annually in the United States (a conservative estimate), the above evidence suggests that 1100 to 1600 anesthesia-related deaths occur in the United States every year.

Another important consideration is that improvement in anesthesia safety is not fully reflected in death rates that are unadjusted for risk. Older and sicker patients, who are more likely to experience adverse consequences of anesthesia, are increasingly receiving anesthesia, and more high-risk operations are being performed overall. Risk-adjusted indicators of anesthesia safety are needed to document improvement, which is probably substantially greater than that reflected in raw death rates.

Ronald A. Gabel, MD
Department of Anesthesiology
University of Rochester
Rochester, NY


**5.** Lagasse RS. Anesthesia safety: model or myth? Anesthesiology. In press.
In Reply:  Dr Gabel cautions that the reduction in overall anesthetic mortality may be less than we quoted for anesthesia in low risk patients, although the figure he cites, 1 in 15000, is also 90% lower than the rate found in the 1950s. 1 This progress in anesthesia mortality is even more remarkable, given that patients undergoing anesthesia today are older, have more comorbid conditions, and are receiving riskier and more complicated operations than before. Another measure of improvement is the dramatic decline in anesthetists' malpractice insurance rates.2 Anesthesia remains a powerful example of what can be accomplished by focusing on safety.

Dr Hayward’s letter focuses not on the topic of our article but on the number of preventable deaths. One of the 2 articles he cites as evidence is his study of 111 nonrepresentative patients, which used a nonrandom and variable number of reviews.3 Nonetheless, his rate of 6% preventable deaths is very close to the 8% rate in the HMPS analysis of 30000 deaths that yielded the 98000 figure quoted by the IOM.

His study reduced that number, however, because of reviewer judgments that only 0.5% would have expected to live 3 months or longer in good cognitive health. Leaving aside questions about the validity or reliability of conclusions based on this small number of patients, or of retrospective estimates of what life expectancy would have been for those patients who died, we reject that interpretation as inappropriate and unethical. Seriously ill and debilitated patients are entitled to our best efforts.

Other evidence suggests that the 100000 figure might be an underestimate. Physicians in the United States may systematically undercount injuries. All replications of the HMPS—in Australia, the United Kingdom, New Zealand, and Denmark—have found adverse event rates that are 2 to 4 times higher, and preventable death rates that are 50% to 100% higher.4,5 Autopsy studies consistently find undiagnosed disease as the cause of death in 10% to 20% of patients, of whom half could have been successfully treated.6 Applied to current US hospital mortality, this yields 40000 to 80000 preventable deaths annually from missed diagnoses alone. Studies of specific types of adverse events, such as adverse drug events, invariably document much higher rates. Outpatient adverse events are also considerable and were not included in the HMPS.8

But from a safety perspective, focusing attention on mortality and life expectancy sends the wrong message. Hospitals cannot be made safer by counting bodies. The transforming insight of the IOM report was that most errors are due to systems failures, not to individual carelessness or incompetence. We believe that identifying these systems failures and fixing them is the safety agenda, not counting deaths, and it is time to get on with it.

Lucien L. Leape, MD
Department of Health Policy and Management
Harvard School of Public Health
Boston, Mass

Donald M. Berwick, MD
Institute for Healthcare Improvement
Boston

David W. Bates, MD, MSc
Department of Medicine
Brigham and Women’s Hospital
Boston


Virolologic Outcomes of Complex Drug Regimens for Human Immunodeficiency Virus

To the Editor: Dr Hammer and colleagues1 reported that increasing to a more potent regimen (including 2 protease inhibitors [PIs] and efavirenz) improved virologic outcome at 24 weeks among patients with human immunodeficiency virus (HIV) who had not improved on a PI-based regimen. Despite these intensive, complex regimens (including up to 40 pills daily), only 47% of patients receiving 2 PIs and who were naive to efavirenz achieved an undetectable plasma HIV-RNA load. It would be helpful to know more about the factors, apart from regimen potency, that also contributed to successful therapy such as prior therapy, prior patient adherence, on-study tolerability, and on-study adherence. A potent but intolerable regimen is doomed to fail, especially for longer periods, as is the prescription of a complex regimen to a patient who has been unable to adhere to similar regimens in the past.2

The authors did not describe whether they enrolled patients failing a first regimen only or whether patients with prior

©2002 American Medical Association. All rights reserved.
nucleoside analog mono/dual therapy were enrolled. Inclusion of substantial numbers of patients with multiple prior regimens would likely reduce the capacity to detect differences of activity between study arms.

It is also unclear why these patients had failed their baseline regimens. Given that adherence and adverse effects are strong predictors of regimen outcome, it is possible that outcomes may have been different in those who had failed for previous nonadherence or adverse events rather than for suboptimal prescribing (eg, dual therapy).

It would also be helpful to know the relative contributions to virologic failure of baseline viral susceptibility, poor adherence, regimen complexity/pill burden, choice of nucleoside analogs, and adverse events. The latter seems especially likely given that only 157 (48%) of 324 patients randomized to dual-PI therapy completed 24 weeks of therapy (with very few developing a new acquired immunodeficiency syndrome–defining illness or dying). Indeed, 19% and 28% of patients had grade 3 or grade 4 clinical or laboratory adverse event, respectively. It is not stated what percentage of these events (or of grade 1-2 events for that matter) resulted in reduced adherence to therapy or withdrawal from the study.

If suboptimal adherence (due to toxicity and regimen complexity) was the main contributor to the high failure rate in this study, then the appropriate response would be simpler, better tolerated therapy. On the other hand, if failure was due mainly to lack of potency, then patients who fail PI-based therapy will require even more potent regimens.

Andrew Carr, MD  HIV, Immunology and Infectious Diseases Clinical Services Unit  St Vincent’s Hospital  Sydney, Australia

Financial Disclosure: Dr Carr has received consultancies and honoraria from Abbott, Bayer, Boehringer-Ingelheim, Bristol-Myers Squibb, GlaxoSmithKline, and Roche, and grant support from Abbott, Boehringer-Ingelheim, Bristol-Myers Squibb/ DuPont, GlaxoSmithKline, Pfizer-Agouron, Roche, and Schering-Plough.


In Reply: In response to Dr Carr, the patients in our study were highly experienced with antiviral drugs: 100% had prior PI exposure (an entry requirement) and of these 79% had exposure to 2 or more PIs; 100% were nucleoside analogue experienced; and 44% were nonnucleoside reverse transcriptase inhibitor (NNRTI) experienced. This high degree of drug experience was reflected in the baseline phenotypic drug susceptibility. As described in our article, of 139 randomly selected patients at baseline, 19% to 66% had viral strains with 50% inhibitory concentrations of more than 2.5-fold above the reference strain for at least 1 of the 4 drugs common to all study arms.

We did in fact collect data on patients’ adherence at baseline and found that this was not a significant predictor of virologic failure. We certainly agree, and noted in our article, that the pill burden was high and no doubt a challenge to adherence. This, however, was not the sole reason for virologic failure. In a univariate logistic regression analysis of the virologic, pharmacokinetic, and adherence contributors to virologic failure, the major contributors to virologic failure were baseline efavirenz susceptibility, efavirenz clearance, NNRTI experience, treatment discontinuation for toxicity, and grade 3/4 toxicity. Baseline HIV-RNA levels and on-study treatment adherence were statistically significant but less important predictors.

The complexity of virologic failure in our study does not change our basic conclusion that patients who received the dual-PI combination regimens had better virologic outcomes than those in the single PI arm. Combining maximal potency of regimens with simplicity, convenience, and lack of toxicity is certainly the appropriate goal for antiretroviral therapy.

Scott M. Hammer, MD  Department of Medicine  Columbia University College of Physicians and Surgeons  New York, NY

Florin Vaida, PhD  Kara K. Bennett, MS  Department of Biostatistics  Statistical and Data Analysis Center  Harvard School of Public Health  Boston, Mass

John W. Mellors, MD  Department of Medicine  University of Pittsburgh  Pittsburgh, Pa

Financial Disclosure: Dr Hammer is a consultant or site investigator for Boehringer Ingelheim, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline, Roche- Trimeris, Shionogi, Shire Biochem, Tibotec-Virco, and Triangle. Dr Mellors has received research grants from Bristol-Myers Squibb, Chiron, DuPont Pharmaceutical, GlaxoSmithKline, Merck and Co, and Abbott Laboratories, and serves as a consultant to Achillion, Agouron/Pfizer, Boehringer-Ingelheim, Bristol-Myers Squibb, DuPont Pharmaceutical, GlaxoSmithKline, Gilead Sciences, Intelligent Therapeutics Solutions, Merck and Co, Novirio, Pharmasset, Triangle Pharmaceuticals, Trimeris, Tibotec-Virco, and Visible Genetics.


Allocation Concealment in Clinical Trials

To the Editor: Dr Balk and colleagues1 found that randomized controlled trials (RCTs) with high scores for allocation concealment, double-blinding, and 22 other quality measures had similar results as low-quality trials, as defined by these indicators. However, the authors’ ambitious efforts may have diluted the quality of some their assessments and hindered comparability with prior studies.

Most notably, Balk et al defined allocation concealment differently than in the studies they apparently wanted to repli-
To the Editor: Dr Balk and colleagues1 found no consistent relationship between the size of treatment effects in RCTs and their methodological quality. However, previous studies found that the lack of an allocation system that prevents the investigator from knowing in advance which treatment the next person will get was consistently associated with larger treatment effects.2 Proper concealment of the allocation sequence thus appears to be crucial to prevent selection bias in RCTs.3 Balk et al argue that such quality effects may not be as important as previously thought.

We updated our meta-analysis3 of the relevant studies and found that the results of Balk et al differ from those from earlier studies, and introduce heterogeneity (P<.001) into a previously homogeneous meta-analysis (Figure). Nevertheless, taken together, these studies continue to show that inadequate concealment of allocation is associated with larger treatment effects (ratio of odds ratios, 0.75; 95% confidence interval [CI], 0.63-0.89). Similarly, they show that absence of blinding is associated with larger effects (ratio of odds ratios, 0.88; 95% CI, 0.78-0.99).

It is possible that the discordant results of Balk et al were due to their selection of meta-analyses that had significant between-study heterogeneity. However, differences in trial quality are only one of several possible sources of between-trial heterogeneity. Other sources, such as differences in study populations and interventions, publication bias, or language bias, may have introduced error variance and reduced the study's ability to detect the influence of trial quality.

Heterogeneity is not a prerequisite for detecting effects of trial quality in meta-analysis: we found, for example, a strong association of treatment effect estimates with lack of blinding in a meta-analysis with little between-trial heterogeneity.4 It is also possible that inconsistencies in the assessments of trial quality between the authors may have resulted in the misclassifi-

©2002 American Medical Association. All rights reserved.

Figure. Comparison of Treatment Effect Estimates From Trials With Inadequate or Unclear Allocation Concealment With Adequately Concealed Trials

<table>
<thead>
<tr>
<th>Source</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schulz et al, 1995</td>
<td>0.66 (0.59-0.73)</td>
</tr>
<tr>
<td>Moher et al, 1998</td>
<td>0.63 (0.45-0.88)</td>
</tr>
<tr>
<td>Kjaergard et al, 2001</td>
<td>0.60 (0.31-1.15)</td>
</tr>
<tr>
<td>Jüni et al, 2001</td>
<td>0.79 (0.70-0.89)</td>
</tr>
<tr>
<td>Balk et al, 2002</td>
<td>0.95 (0.83-1.10)</td>
</tr>
<tr>
<td>Combined</td>
<td>0.75 (0.63-0.89)</td>
</tr>
</tbody>
</table>

Odds ratios (ORs) lower than 1 represent trials that had inadequate or unclear allocation concealment overestimate effects compared with adequately concealed trials. CI indicates confidence interval. The trials, except for Balk et al,1 are referenced in Jüni et al.4

(Reprinted) JAMA, November 20, 2002—Vol 288, No. 19 2407
cation of some trials, thus leading to an underestimation of associations between treatment effects and methodological quality. Unfortunately, Balk et al do not report k values, which would allow an unbiased assessment of the degree of nonrandom agreement between the investigators.

We agree with Balk et al, and have argued previously, that it is problematic to adjust or “correct” the results from a single trial, or a meta-analysis of several trials, for trial quality. However, we are concerned that their study will be misinterpreted as showing that the standards for the conduct and reporting of RCTs, which we and many others have promoted in recent years, are not important. Theoretical considerations, underpinned by our reanalysis, provide arguments for the importance of allocation concealment. The fact that in empirical studies aspects of the design and conduct of trials appear to be more important in some situations but less important in others does not mean that these standards should be optional.

Peter Juni, MD
Matthias Egger, MD
Departments of Social and Preventive Medicine
and Rheumatology
University of Bern
Bern, Switzerland

In Reply: Dr Schulz and colleagues suggest that our results may differ from previous reports because definitions of quality measures, specifically allocation concealment, were not identical. Their point highlights a difficulty in applying quality measures and achieving consensus on precise definitions. We took great care to create definitions that were consistent with most prior applications of these quality measures and that were also unlikely to be misinterpreted by data extractors. We ended with very high agreement among data extractors.

Nevertheless, to address their concerns, we repeated our analysis by having a single data extractor (E.K.B.) use his definition for allocation concealment. Most changes in classification involved studies using envelopes. Studies in which pharmacists allocated treatments and most that used computers (as opposed to computer-generated lists) were coded as centrally randomized and thus were generally adequately concealed. The reclassification of studies and relative odds ratios for adequate vs inadequate allocation concealment and for adequate vs inadequate and unclear allocation concealment are shown in the Table. No significant differences favoring adequate concealment were found between adequately and inadequately concealed studies. Inadequately or unclearly concealed infectious disease trials actually showed a smaller treatment benefit than other infectious disease trials. This relationship was not statistically significant in the other medical domains or in the overall analysis.

In Reply: Drs Juni and Egger state that many causes of heterogeneity among the published studies may be associated with the magnitude of the treatment effect other than quality measures. Indeed, we discussed this point in our article. The relationship between study quality and treatment effect is complex. We have highlighted that there is heterogeneity in the association of quality measures with treatment effect among different medical domains. Our reanalysis reported herein also highlights how there can be heterogeneity due to minor differences in quality measure definitions. We believe that the meta-analysis of meta-analyses performed by Juni and Egger obscures understanding the reasons for discrepancies found among the studies. Calculating a single estimate of association in this situation can be misleading.

We agree that our results should not be interpreted to mean that proper blinding and randomization are not important. Trial quality measures, however, should not be used as a single universal determinant for choosing studies for inclusion in meta-analyses. Inadequate allocation concealment may inflate the treatment effect sometimes, but may deflate the treatment effect.

### Table. Studies With Adequate, Inadequate, and Unclear Allocation Concealment

<table>
<thead>
<tr>
<th>Allocation Concealment</th>
<th>Overall</th>
<th>Cardiovascular Disease</th>
<th>Infectious Disease</th>
<th>Pediatrics</th>
<th>Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequate</td>
<td>21</td>
<td>24</td>
<td>21</td>
<td>28</td>
<td>9</td>
</tr>
<tr>
<td>Inadequate</td>
<td>9</td>
<td>3</td>
<td>18</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Unclear</td>
<td>70</td>
<td>73</td>
<td>61</td>
<td>65</td>
<td>81</td>
</tr>
</tbody>
</table>

**Relative Odds Ratio (95% Confidence Interval)**

- Adequate vs inadequate (n = 29): 0.82 (0.43-1.57)
- Adequate vs inadequate and unclear (n = 276): 0.89 (0.72-1.09)
- Adequate vs unclear: 1.01 (0.85-1.21)
- Adequate vs inadequate: 0.61 (0.40-0.98)
- Adequate vs unclear: 1.07 (0.60-1.74)
- Adequate vs inadequate and unclear: 1.01 (0.58-1.70)

*A relative odds ratio of less than 1.00 suggests that the benefit of treatment appeared to be greater in high-quality than in low-quality trials. The relative odds ratio definition used here is the inverse of the definition of Juni et al. NA indicates not applicable (too few studies to analyze separately in each domain).

†Fifty-three additional studies coded as having adequate or inadequate concealment were excluded because of sparse numbers of eligible studies within each meta-analysis.
or have no relationship with it in other cases. Furthermore, even among experts, allocation concealment, as well as other quality measures, are subject to diverse interpretations.

Ethan M. Balk, MD, MPH
Peter A. L. Bonis, MD
Joseph Lau, MD
Division of Clinical Care Research
Tufts-New England Medical Center
Boston, Mass
John P. A. Ioannidis, MD
Department of Hygiene and Epidemiology
University of Ioannina School of Medicine
Ioannina, Greece


**CORRECTIONS**

Incorrect Number: In Appendix IA, Table 2 published in the September 4, 2002, issue of *THE JOURNAL* (2002;288:1143-1145), on page 1145 Baylor College of Medicine’s total enrollment should be 667 students.

Incorrect Affiliation: In the Medical News & Perspectives article entitled “Sewage Yields Clues to SV40 Transmission” published in the September 18, 2002, issue of *THE JOURNAL* (2002;288:1337-1338), Michael Carbone, MD, PhD, is erroneously referred to as “a University of Chicago pathologist.” Carbone is now associate professor in the Department of Pathology at Loyola University Chicago Stritch School of Medicine, Cardinal Bernardin Cancer Center, in Maywood, Ill. He was previously at the University of Chicago.

Incorrect Data: In the article entitled “Employer Drug Benefit Plans and Spending on Prescription Drugs” published in the October 9, 2002, issue of *THE JOURNAL* (2002;288:1733-1739), there was incorrect data in a table. On page 1736, in Table 2, the number (percentage) for patient characteristic aged 35 to 44 years stated “207473 (29.5)” but should have read “185025 (26.3).”

**CME ANNOUNCEMENT**

**Online CME to Begin in Mid-2003**

In mid-2003, online CME will be available for *JAMA/Archives Journals* and will offer many enhancements:

- Article-specific questions
- Hypertext links from questions to the relevant content
- Online CME questionnaire
- Printable CME certificates and ability to access total CME credits

We apologize for the interruption in CME and hope that you will enjoy the improved online features that will be available in mid-2003.