Nonsteroid drug selectivities for cyclo-oxygenase-1 rather than cyclo-oxygenase-2 are associated with human gastrointestinal toxicity: a full in vitro analysis. *Proc Natl Acad Sci USA* 1999;96:7563-7568.
2. Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti- inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. *JAMA* 2000;284:1247-1255.
3. Okie S. Missing data on Celebrex. Full study altered picture of drug. *Washington Post* 2001;5 Aug:A11.
4. Berg Hrachovec J, Mora M. Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. *JAMA* 2001;286:2398.
5. Wright JM, Perry TL, Bassett KL, Chambers KG. Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. *JAMA* 2001;286:2398-2399.
6. US Food and Drug Administration. Transcript of the arthritis advisory committee. www.fda.gov/ohrms/dockets/ac/01/transcripts/3677t1.rtf [Accessed December 10, 2001].
7. Lu HL. Statistical reviewer briefing document for the advisory committee. www.fda.gov/ohrms/dockets/ac/01/briefing/3677b1_04_stats.doc. [Accessed December 10, 2001].
8. Witter J. Medical officer review. www.fda.gov/ohrms/dockets/ac/01/briefing/3677b1_03_med.pdf. [Accessed December 10, 2001].
9. Silverstein F, Simon L, Faich G. Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. In reply. *JAMA* 2001;286:2399-2400.
10. Pharmacia earnings releases. Peapack, NJ, Pharmacia Corporation 2002. www.pharmacia.com/investor/earnings.asp. [Accessed February 25, 2002].
11. Gottlieb S. Researchers deny any attempt to mislead the public over *JAMA* article on arthritis drug. *BMJ* 2001;323:301.
12. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR Study Group. *N Engl J Med* 2000;343:1520-1528.

CLASS clarification: reaffirms the medical importance of the analyses and results

G Steven Geis

Pharmacia
5200 Old Orchard Rd
Skokie, IL 60077

G Steven Geis
group vice president, clinical research

Correspondence to:
G S Geis
george.s.geis@pharmacia.com

We appreciate the opportunity to correct the inaccuracies in the article by Jüni et al,[1] which appear to be based on an incomplete understanding of the CLASS design, data, and statistical issues. Contrary to the editorial's assertions, there were no post hoc changes in the study design or the outcome definitions. The CLASS authors reviewed all the data and decided the 6-month analyses were most appropriate for initial publication,[2] while the FDA chose 9-month data as most appropriate for a recent label change.[3] Despite differing medical judgment for the time interval that best reflected the data, and contrary to the *BMJ* editorial allegations, the conclusions were similar.

CLASS was a single study with the objective of comparing the rate of ulcer complications with celecoxib versus traditional NSAIDs. The need for two protocols was prespecified to ensure blinding of study medications. One protocol included celecoxib 400 mg BID and ibuprofen 800 mg TID; the other included celecoxib 400 mg BID and diclofenac 75 mg BID. Aside from this difference and the inclusion of quality of life measures in one protocol, the protocols were identical. The statistical plan stated that the data from the two protocols would be analyzed as a single study. Low-dose aspirin was allowed and the minimum expected duration of study participation was 6 months. Two important assumptions of the design were: 1) a constant rate of ulcer complications in the NSAID group[4] and 2) that about 11% of enrolled patients would use low-dose aspirin.[5]

The primary end point was ulcer complications (bleeding, perforation and outlet obstruction) verified by endoscopy or contrast *x* ray, but analysis of symptomatic ulcers was also prespecified. The protocols mandated study withdrawal if a patient was found to have a non-bleeding ulcer (ie, symptomatic ulcer).

The primary analysis was a comparison of ulcer complications (traditional definition)[6] with celecoxib versus the combined NSAID group (ibuprofen plus diclofenac) and only if the differences were statistically significant would comparisons of celecoxib to each of the individual NSAIDs be performed. This was specified in order to control the overall alpha-level. The plan included an analysis of the effects of risk factors for ulcer complications (eg, low-dose aspirin use) on the results of the different treatments.

Ulcer complication rates were not significantly different for celecoxib versus the NSAID group. However, the rate of the combined end point of symptomatic/complicated ulcers was significantly lower with celecoxib. Since the primary analysis was not significant, comparisons to the individual NSAIDs were not valid.

Once the study blind was broken, it was clear that important assumptions made in designing the trial had not proved true. NSAID ulcer complication rates decreased over time instead of remaining constant (figure 1). Moreover, the withdrawal rate of patients due to symptomatic ulcers was statistically greater in the NSAID group versus the celecoxib group, and the difference was most apparent after the first 6 months of the study (figure 2). Since symptomatic ulcers are precursors of ulcer complications, high-risk patients were being depleted from the NSAID group more quickly than from the celecoxib group. This differential withdrawal rate introduced study bias, which caused the analyses to become less valid with time.

As described by other authors of the *JAMA* paper,[7] after extensive review, the CLASS oversight committees judged the analyses of the 6-month data to be the most scientifically and clinically valid since: 1) 6 months was the minimum duration of study participation; 2) more than 50% of patients were continuing in the study at 6 months (median duration of exposure: 9 months for celecoxib and diclofenac; 6 months for ibuprofen) and 3) the impact of differ-

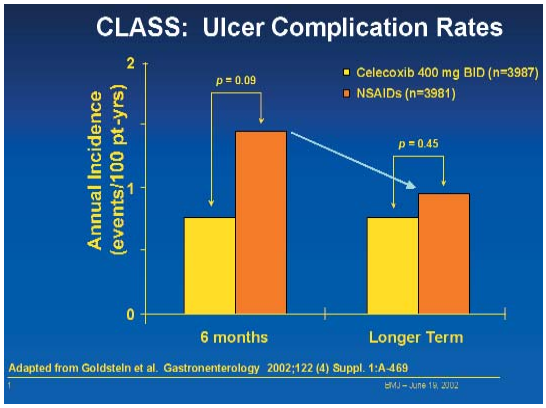

Figure 1 Ulcer complication rates expressed as annual incidence as calculated using the cumulative number of events and duration of treatment exposure in the celecoxib and NSAID groups over the first 6 months of the study (L) and over period of maximum treatment. Arrow notes the reduction in NSAID rate over time.

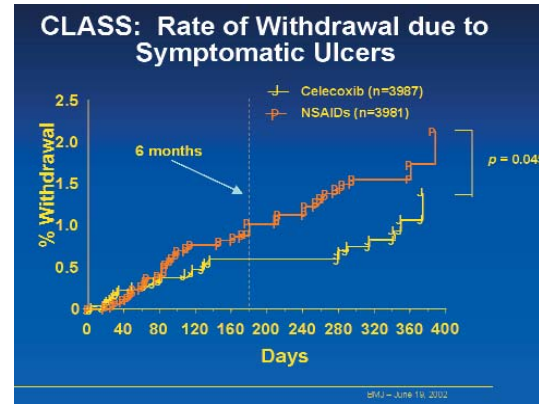

Figure 2 Kaplan-Meier estimates for withdrawals due to symptomatic ulcers. No withdrawal due to a symptomatic ulcer occurred after final time points.

ential withdrawal of patients at risk was most apparent beyond this time point. The CLASS oversight committees reported that "the data after 6 months were so confounded as to be difficult to interpret for assessing a drug-related causal GI toxicity."[7]

The second unmet assumption in the CLASS trial was that approximately 22% of patients in each treatment group took low-dose aspirin rather than the expected 11%. Analyses showed that low-dose aspirin use in CLASS was a risk factor for ulcer complications for celecoxib-treated patients but not for the NSAID-treated patients. Based on the 6-month analysis for the all-patient cohort, ulcer complication rates were not significantly different between the celecoxib and the NSAID groups. However, the cohort of patients not using aspirin showed a statistically significant lower rate of ulcer complications with celecoxib versus the NSAIDs. The FDA noted "... the use of aspirin ... may have obscured the ability to accurately compare the GI safety of Celebrex to other nonsteroidal anti-inflammatory drugs."[8]

The editorial describes the *JAMA* paper as "overoptimistic," using "post hoc changes to the protocol while omitting disappointing longer term data."[1] The *JAMA* manuscript clearly acknowledges that the primary end point of the study was not reached.[2] There were no post hoc changes to the protocol. The analyses of the longer-term data, although complicated by the differential dropouts, do not differ substantially from the 6-month analyses.[9]

We continue to stand behind the CLASS design, analyses and conclusions as stated in *JAMA* and invite discussions that will ensure an understanding of the facts and assist in clarifying the safety profile of celecoxib.

1. Jüni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-1288.
2. Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. *JAMA* 2000;284:1247-1255.
3. FDA-Approved Celebrex Approved Prescribing Information (June 7, 2002): http://www.fda.gov/cder/foi/label/1998/20998lbl.pdf.
4. MacDonald TM, Morant SV, Robinson GC, Shield MJ, McGilchrist MM, Murray FE, McDevitt DG. Association of upper gastrointestinal toxicity of nonsteroidal anti-inflammatory drugs with continued exposure: cohort study. *BMJ* 1997;315:1333-1337.
5. Celecoxib NDA, 1998.
6. Silverstein FE, Graham DU, Senior JR, Davies HW, Struthers BJ, Bittman RM, Geis GS. Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving non-steroidal anti-inflammatory drugs: a randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 1995;123:241-249.
7. Silverstein F, Simon, L. Faich G. Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. In reply. *JAMA* 2001;286:2399-2400.
8. FDA Talk Paper: http://www.fda.gov/bbs/topics/ANSWERS/2002/ANS01151.html.
9. Goldstein JL, for the CLASS investigators. Gastrointestinal (GI) event rates in the CLASS study: six months vs longer-term follow-up analyses. *Gastroenterology* 2002;122(4) Suppl 1:A-469.

The authors respond

Peter Jüni, Anne W S Rutjes, Paul A Dieppe

Geis characterizes CLASS as a single study, as was done in *JAMA*, which reported patients to be "randomly assigned on a 2:1:1 basis."[1] This description is misleading; there were two separate trials, with two separate patient recruitment and randomization procedures, and therefore requiring separate analyses to preserve randomization. Nonetheless, the two trials were combined by simply adding up numbers.[1]

The assumption underlying this approach is that allocation of patients to the two trials was ruled by chance alone. We tested this assumption