Letters

Website: **bmj.com**
Email: **letters@bmj.com**

Are selective COX 2 inhibitors superior to traditional NSAIDs?

Rofecoxib did not provide unequivocal benefit over traditional NSAIDs

EDITOR—In their editorial Jüni et al say that celecoxib is no safer than diclofenac or ibuprofen and that the CLASS authors "spun" their analysis to suggest otherwise.[1] They also state: "In contrast with the CLASS trial, the VIGOR trial, which was similar in design and outcomes, found an unequivocal benefit of another selective COX 2 inhibitor, rofecoxib, over traditional non-steroidal anti-inflammatory agents [NSAIDs]."[2]

I disagree. Although serious gastrointestinal adverse effects were less frequent in rofecoxib users than in patients with rheumatoid arthritis treated with naproxen (number needed to treat to prevent one serious upper gastrointestinal event 191; 95% confidence interval 114 to 586), rofecoxib was in fact less safe than naproxen. The published version of the VIGOR trial focused on the narrow outcome of serious gastrointestinal complications.

The US Food and Drug Administration took the unprecedented step of presenting its review of both the CLASS trial and the VIGOR trial on its website.[3] Review of the complete data presented there shows that when all serious adverse events are included—not just gastrointestinal events—patients treated with naproxen had fewer serious events. Among patients treated with rofecoxib, 9.3% experienced a serious adverse event compared with 7.8% of those treated with naproxen (relative risk 0.81; 0.62 to 0.97). When all serious adverse events are counted, the number needed to harm when rofecoxib is used compared with naproxen is 66 (36 to 332).

The increased risk of serious adverse events was due to an increase of serious adverse cardiovascular events, including a 300% greater risk of myocardial infarction in those treated with rofecoxib.

The VIGOR results, examined fully, show that at least one traditional non-steroidal anti-inflammatory drug—naproxen—is unequivocally safer than rofecoxib, albeit with an increased risk of adverse events limited to the gastrointestinal tract.

Brian R Budenholzer *director*
Clinical Enhancement and Development for Network Services Division, Group Health Cooperative, PO Box 204, Spokane, WA 99210-0204, USA
budenholzer.b@ghc.org

1 Jüni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-8. (1 June.)
2 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; 343:1520-8.
3 US Food and Drug Administration. www.fda.gov/ohrms/dockets/ac/01/briefing/3677b2_03_med.pdf (accesssed 9 July 2002)

❄Three other responses on bmj.com (bmj.com/cgi/eletters/324/7349/1287) made similar points.

Pharmacia's response to editorial

EDITOR—Contrary to the assertions of Jüni et al in their editorial,[1] the CLASS design, analyses, and outcome definitions were predefined. The CLASS authors reviewed all the data and decided that the six month analyses were most appropriate for initial publication while the Food and Drug Administration chose nine month data as most appropriate for a recent label change.[2 3] Despite differing medical judgment for the time interval that best reflected the data, and contrary to the allegations in the editorial, the conclusions were similar.

CLASS was a single study using two protocols to ensure treatment blinding, but analysis of the combined results was prespecified. The protocols were similar; however, one included celecoxib 400 mg twice daily and ibuprofen 800 mg thrice daily while the other used celecoxib 400 mg twice daily and diclofenac 75 mg twice daily. Low dose aspirin was allowed, and the minimum expected duration of participation in the study was six months. The primary end point was ulcer complications (bleeding, perforation, and outlet obstruction) verified by endoscopy or contrast radiography, but analysis of symptomatic ulcers was also prespecified. Early withdrawal for an uncomplicated ulcer—that is, symptomatic ulcer—was mandated in the protocol.

The primary analysis was a comparison of ulcer complication rates between celecoxib and the combined non-steroidal anti-inflammatory group (ibuprofen and diclofenac). To control the overall alpha-level, comparisons of celecoxib with each non-steroidal agent were allowed only if the primary analysis was statistically significant. Analyses of ulcer complication risk factors—for example, use of low dose aspirin—were also preplanned. Study assumptions included (*a*) constant complication rates for non-steroidal anti-inflammatory drugs[4] and (*b*) a rate of use of low dose aspirin of around 11%.

Ulcer complication rates were not significantly different for the two groups. 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 with the individual non-steroidal anti-inflammatory drugs were not valid.

Important design assumptions did not prove to be true. Ulcer complication rates with non-steroidal anti-inflammatory drugs decreased over time instead of remaining constant (figure). Those given non-steroidal anti-inflammatory drugs had a significantly greater withdrawal rate for symptomatic ulcers than those given celecoxib, which was most evident after the first six months (figure). Since symptomatic ulcers are precursors of ulcer complications, patients at high risk who were given non-steroidal anti-inflammatory drugs were being withdrawn more quickly than high risk patients given celecoxib. This differential withdrawal rate introduced study bias, which reduced statistical and medical validity of the analyses over time. Therefore, the CLASS oversight committees judged the six month data to be most valid and reported: "The data after six months were so confounded as to be difficult to interpret for assessing

---

Advice to authors

We prefer to receive all responses electronically, sent directly to our website. Processing your letter will be delayed unless it arrives in an electronic form.

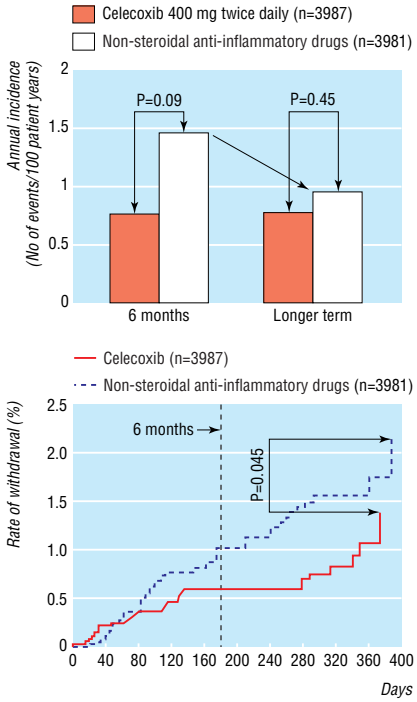
We are now posting all direct submissions to our website within 24 hours of receipt and our intention is to post all other electronic submissions there as well. All responses will be eligible for publication in the paper journal.

Responses should be under 400 words and relate to articles published in the preceding month. They should include ⩽5 references, in the Vancouver style, including one to the BMJ article to which they relate. We welcome illustrations.

Please supply each author's current appointment and full address, and a phone or fax number or email address for the corresponding author. We ask authors to declare any competing interest. Please send a stamped addressed envelope if you would like to know whether your letter has been accepted or rejected.

Letters will be edited and may be shortened.

bmj.com
letters@bmj.com



Top: Ulcer complication rates expressed as annual incidence calculated using cumulative number of events and duration of treatment in celecoxib and non-steroidal anti-inflammatory groups over the first six months of the study and period of maximum treatment.[7] Arrow shows reduction in rate with non-steroidal anti-inflammatory drugs over time. Bottom: Kaplan-Meier estimates of patients withdrawing because of symptomatic ulcers. No patient was withdrawn because of symptomatic ulcer after final time points

a drug-related causal gastrointestinal toxicity."[5]

About 22% of patients took low dose aspirin, which affected the results of treatment. Differences in treatment were not significant for the six month comparison of ulcer complication rates for the all patient cohort, but for the cohort of non-aspirin users, the rate was significantly lower for celecoxib than for non-steroidal anti-inflammatory drugs. The US Food and Drug Administration noted: "The use of aspirin … may have obscured the ability to accurately compare the gastrointestinal safety of Celebrex to other non-steroidal anti-inflammatory drugs."[6] The US label describes a fourfold increase in the nine month ulcer complication rate with celecoxib plus aspirin *v* celecoxib alone. For the rate of symptomatic/complicated ulcers, a near threefold increase for celecoxib plus aspirin *v* celecoxib alone is described.[3]

Jüni et al describes CLASS as "overoptimistic," using "post hoc changes."[1] However, the CLASS publication clearly acknowledges that the primary end point was not reached.[2] There were no post hoc protocol changes, and the analyses of the longer term data, although complicated by the differential withdrawal of patients, do not differ substantially from the six month analyses.[7]

Jüni et al misrepresented CLASS.[1] We continue to stand behind the study design, analyses, and conclusions.[2] Furthermore, we invite any discussions that will ensure an understanding of the facts and help in clarifying the safety profile of celecoxib.

G Steven Geis *group vice president, clinical research*
Pharmacia, 5200 Old Orchard Road, Skokie, IL 60077, USA
george.s.geis@pharmacia.com

1 Jüni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-8. (1 June.)
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;284:1247-55.
3 FDA approved Celebrex approved prescribing information (7 June 2002). Available at: www.fda.gov/cder/foi/label/1998/20998lbl.pdf (accessed 11 July 2002).
4 MacDonald TM, Morant SV, Robinson GC, Shield MJ, McGilchrist MM, Murray FE, et al. Association of upper gastrointestinal toxicity of non-steroidal anti-inflammatory drugs with continued exposure: cohort study. *BMJ* 1997;315:1333-7.
5 Silverstein F, Simon L, Faich G. Reporting of 6-month vs 12-month data in a clinical trial of celecoxib. In reply. *JAMA* 2001;286:2399-400.
6 FDA talk paper. Labeling changes for arthritis drug Celebrex. Available at: www.fda.gov/bbs/topics/ANSWERS/2002/ANS01151.html (accessed 11 July 2002).
7 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 (suppl 1):A-469.

Little is known about COX 2 inhibitors

EDITOR—The editorial by Jüni et al focuses attention not only on the reporting of clinical trials in peer reviewed journals but also on the interpretation of available evidence.[1] We agree more needs to be done to determine whether COX 2 inhibitors are superior to traditional non-steroidal anti-inflammatory drugs (NSAIDs) and have some suggestions of how that might be done.

Firstly, understanding of the cellular effects of the COX 2 inhibitors is evolving, and conclusions about the comparative safety of these agents based on in vitro data cited by Jüni et al may be premature. Although whole blood assays show that rofecoxib is more selective than celecoxib for COX 2,[2] such assays have been criticised for having limited clinical relevance.[3]

Furthermore, studies of cancer cell lines indicate that celecoxib has far greater antiproliferative effects than rofecoxib, implying that celecoxib has greater COX 2 selectivity.[4 5] Recent evidence also suggests, however, that celecoxib may possess unique and largely unknown COX independent characteristics over rofecoxib.[5] The clinical implications of such differences on the gastrointestinal and cardiovascular safety of these two agents are not known. In short, this is a new class of drugs about which comparatively little is known.

Secondly, meta-analysis of existing trial data cannot overcome the important design issues associated with the existing randomised trials of selective COX 2 inhibitors, including the choice of comparison drugs and outcomes. What is needed, and perhaps what should have been ordered by licensing bodies at the outset, is a study or set of studies designed to allow direct comparisons of COX 2 agents with appropriate existing alternative treatments.

Thirdly, regardless of the debate about phase III trials, there is an important role for phase IV pharmacosurveillance studies. While the clinical findings of well designed randomised trials are awaited, phase IV studies will shed light on the impact of COX 2 inhibitors in comparison with other agents in "real world" settings; provide much needed information on rare adverse events such as gastrointestinal haemorrhage, and acute myocardial infarction that may be associated with COX 2 inhibitors; and clarify the extent to which these agents are being used in patients who are most likely to benefit.

As noted by Jüni et al, billions of dollars are being spent on COX 2 inhibitors. It seems prudent to determine whether that investment is justified.

Muhammad Mamdani *scientist*
muhammad.mamdani@ices.on.ca

David N Juurlink *clinical pharmacologist*
Geoffrey M Anderson *senior scientist*
Institute for Clinical Evaluative Sciences, 2075 Bayview Avenue, G106, Toronto, Ontario, Canada M4N 3M5

1 Jüni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-8. (1 June.)
2 Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. *JAMA* 2000;284:1247-55.
3 Blain H, Boileau C, Lapicque F, et al. Limitation of the in vitro whole blood assay for predicting the COX selectivity of NSAIDs in clinical use. *Br J Clin Pharmacol* 2002;53:255-65.
4 Patti R, Gumired K, Reddanna P, Sutton LN, Phillips PC, Reddy CD. Overexpression of cyclooxygenase-2 (COX-2) in human primitive neuroectodermal tumors: effect of celecoxib and rofecoxib. *Cancer Lett* 2002;180:13-21.
5 Waskewich C, Blumenthal RD, Li H, Stein R, Goldenberg DM, Burton J. Celecoxib exhibits the greatest potency amongst cyclooxygenase (COX) inhibitors for growth inhibition of COX-2-negative hematopoietic and epithelial cell lines. *Cancer Res* 2002;62:2029-33.

Both the CLASS and VIGOR trials support the COX 2 hypothesis

EDITOR—With respect to the recent editorial by Jüni et al on the CLASS trial,[1] the statistical flaws in the trial's design have been well documented, and the manner in which the data were reported have been widely criticised.

Yet, the question raised in the title of this editorial seems far fetched. Clearly, the CLASS protocol was doomed at the outset by permitting aspirin use. Aspirin, after all, is the grandfather of all non-steroidal anti-inflammatory drugs (NSAIDs) with profound COX 1 inhibition. Thus, the Kaplan-Meier curves in figure 2 of the editorial do not reflect any true test of the COX 2 specific hypothesis with celecoxib.

Notwithstanding Jüni et al's protestations of the analytical presentations of the CLASS data, the incidence of ulcers with COX 2 selective treatment with celecoxib in the CLASS trial was considerably below the 2-4% per year.

The US Food and Drug Administration's labelling of celecoxib shows a Kaplan-Meier rate of complicated ulcers at nine months of 0.32% with celecoxib alone *v* 1.12% among patients taking aspirin with celecoxib. For rofecoxib the rate (at 10.5 months) was 0.52% *v* 1.22% for naproxen.[2]

Similarly, the rate of symptomatic ulcers for patients taking celecoxib was 0.78% *v* 2.19% for those taking apirin and celecoxib. For rofecoxib the labelling shows a rate of 1.80% *v* 3.87% for naproxen.

Presumably, the FDA has made the appropriate adjustments to the analysis.

It should be noted that CLASS enrolled a more elderly, at risk, population. In the CLASS trial 34.8% of patients taking celecoxib alone were aged 65 or over *v* 24.6% of patients taking rofecoxib in the VIGOR trial.

It would have been instructive to see what the rates were with ibuprofen and diclofenac. I agree with Jüni et al's recommendation for a meta-analysis. I think that within the CLASS trial the P value would be <0.05 if patients treated with celecoxib alone were compared with all patients who received a COX 1 inhibitor (aspirin, ibuprofen, diclofenac) using the Food and Drug Administration's methods.

Today, celecoxib, rofecoxib, and valdecoxib carry the traditional warning for non-steroidal anti-inflammatory drugs, implying that they are associated with ulcer rates of 2%-4% per year. I think that this represents mislabelling of the cruellest sort to patients and which is not supported by any scientific or clinical data. In the United States it enables managed care to deprive patients of the safer alternatives.

Richard R Stover *senior pharmaceutical research analyst*
Arnhold and S Bleichroeder, New York, NY 10105-4300, USA
rick.stover@asbinc.com

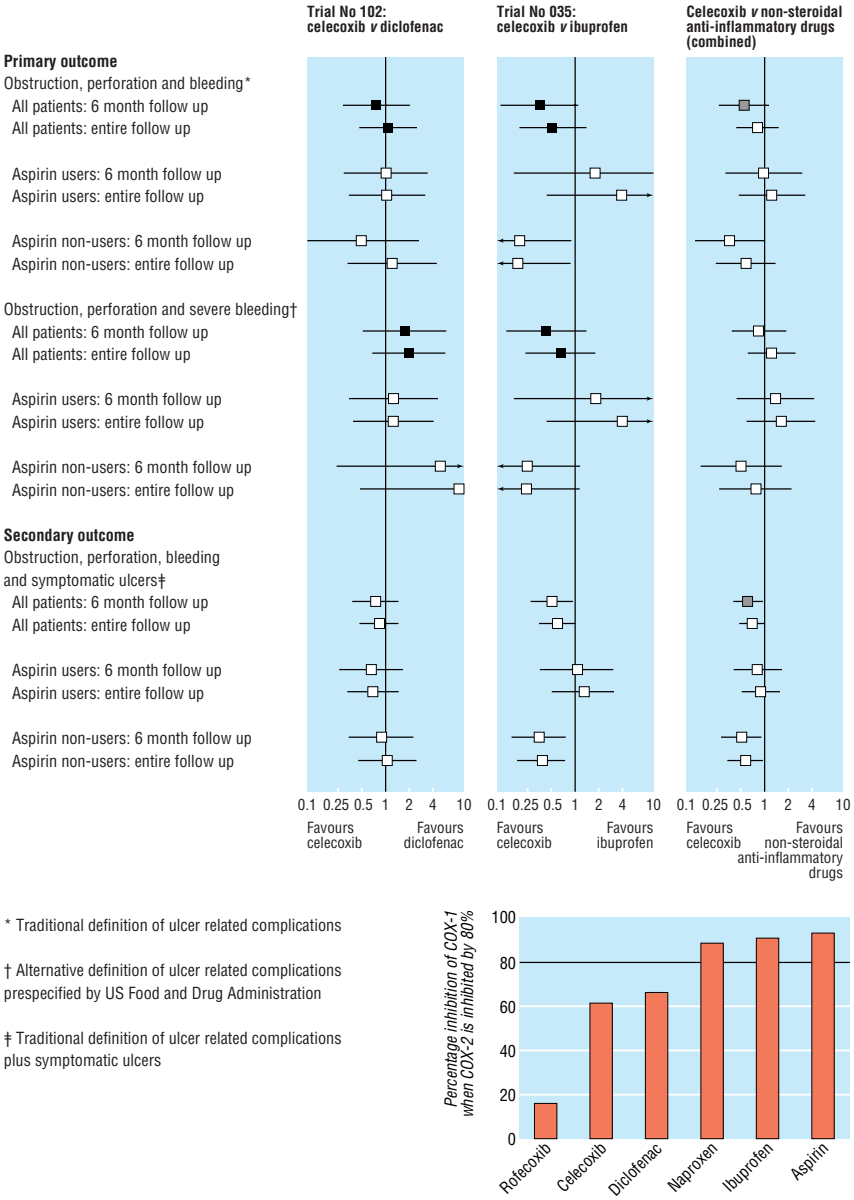
1 Jüni P, Rutjes AWS, Dieppe PA. Are selective COX 2 inhibitors superior to traditional non steroidal anti-inflammatory drugs? *BMJ* 2002;324:1287-8. (1 June.)
2 Celebrex labeling, revised 7 June 2002. Available at: www.fda.gov/cder/foi/label/2002/20998s009lbl.pdf (accessed 11 July 2002).

Authors' reply

Editor—We agree with Budenholzer, who criticises our statement on the "unequivocal benefit" of rofecoxib found in the VIGOR study.[1] We were referring to ulcer complications, and should have reported VIGOR to have found an "unequivocal gastrointestinal benefit" only. Rofecoxib's fivefold increase in myocardial infarctions observed in VIGOR is particularly worrying.[2] Incidentally, whereas patients in the CLASS study randomised to celecoxib had similar rates of myocardial infarction to those randomised to ibuprofen, they tended to experience more myocardial infarctions than patients randomised to diclofenac (relative risk 2.21, 95% confidence interval 0.74 to 8.94, $P=0.14$).[3]

According to Geis, CLASS was a single study, and the *JAMA* article reported patients to be "randomly assigned on a 2:1:1 basis."[4] This is misleading. There were two trials with separate patient recruitment and randomisation procedures requiring separate analyses to preserve randomisation. None the less, the two trials were combined by simply summing them.[4] This would be appropriate only if allocation of patients to trials was ruled by chance. However, there were highly significant differences at baseline between trials in patients' age, disease severity, ethnic group, and histories of intolerance to non-steroidal anti-inflammatory drugs and use of alcohol. The probability that these differences occurred by chance is $P<10^{-19}$.[3] Considering this and different follow up durations and COX 2 selectivities of comparator drugs,[5] simply summing results is inappropriate.

The figure (top) shows relative risks of comparisons discussed during the current debate. While there were some significant differences between celecoxib and ibuprofen, there were none between celecoxib and diclofenac in any of the 18 analyses. Relevant differences between trials were hidden and spurious statistical precision was implied through inappropriate pooling. Therefore, we find it remarkable that



Top: Relative risks and 95% confidence intervals calculated from proportions for primary and secondary outcomes in CLASS study according to separate and combined analyses of the two trials for different follow up durations and different groups of patients. Black squares indicate main outcome measures as prespecified in CLASS protocols, grey squares indicate main outcome measures as reported.[4] Identical event rates were assumed in celecoxib groups of both trials for secondary outcome because separate trial data were unavailable. Arrowheads indicate truncation of confidence intervals. Bottom: Analysis of percentage inhibition of COX 1 when COX 2 is inhibited by 80%. Horizontal line indicates equiactivity—80% inhibition of COX 1. Adapted from Warner et al[5]

Geis misuses CLASS's analytical two step procedure,[3] which should have been a safeguard against a type I error, as a justification for obscuring individual trial results and as a rationale to commit a type I error by inappropriately pooling secondary outcomes. Disturbingly, Pharmacia has also presented pooled results from different protocols with different comparator drugs for SUCCESS-1, the successor study to CLASS.[6]

The main argument for reporting only six month data from CLASS was that patients dropping out because of gastrointestinal adverse events/symptomatic ulcers were at increased risk of ulcer complications.[3] However, only 11 out of 44 patients with ulcer complications in CLASS developed gastrointestinal symptoms before an ulcer complication occurred (25%) and none had a symptomatic ulcer as a precursor.[3] Patients who experience gastrointestinal adverse events may be monitored and treated more carefully than patients without any gastrointestinal symptoms, with ulcer complications avoided particularly in patients with symptoms. It is therefore not surprising that gastrointestinal adverse events were associated with a significantly decreased risk of subsequent ulcer complications in CLASS (relative risk 0.28, 0.12 to 0.66, P = 0.0018).[3]

The wide confidence intervals of comparisons in the figure (top) indicate that CLASS trials were underpowered. We therefore fully agree with Mamdani et al's suggestions about future research. A meta-analysis of individual patient data will allow adequately powered explorative comparisons between COX 2 inhibitors and traditional non-steroidal anti-inflammatory drugs. Subsequently, an industry independent long term trial is needed, allocating patients to one of four agents: celecoxib, rofecoxib, diflofenac, or naproxen.

Stover suggests unfair testing of the COX 2 hypothesis in CLASS because of aspirin use in some patients. The figure (top) shows no consistent pattern distinguishing aspirin users from non-users, suggesting that CLASS comparisons were not confounded by aspirin use. According to the results of an in vitro analysis by Warner et al summarised in the figure (bottom), celecoxib and diclofenac actually have similar COX 2 selectivity. Therefore CLASS's failure to demonstrate celecoxib's superiority may have more to do with celecoxib's shortcomings as a COX 2 inhibitor. Not surprisingly, the US Food and Drug Administration refused to change celecoxib's labelling, which still states as of 7 June 2002 that "serious gastrointestinal toxicity such as bleeding, ulceration, and perforation can occur at any time."[7]

Peter Jüni *senior research fellow in clinical epidemiology*
Departments of Rheumatology, and Social and Preventive Medicine, University of Berne, 3010 Berne, Switzerland
peter.juni@insel.ch

Anne W S Rutjes *research fellow in clinical epidemiology*
Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, 1100 DE Amsterdam, Netherlands

Paul Dieppe *professor of health services research*
MRC Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol BS8 2PR

We thank Yoav Ben-Shlomo for helpful comments, and Ben Calnan for data management and entering.

1 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;343:1520-8, 2.
2 Mukherjee D, Nissen SE, Topol EJ. Risk of cardiovascular events associated with selective COX-2 inhibitors. *JAMA* 2001;286:954-9.
3 Jüni P, Rutjes AW, Dieppe PA. Pharmacia addresses June 1 editorial regarding CLASS study: authors' response. bmj.com 2002: bmj.com/cgi/content/full/324/7349/1287 23766 (accessed 12 July 2002).
4 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;284:1247-55.
5 Warner TD, Giuliano F, Vojnovic I, Bukasa A, Mitchell JA, Vane JR. 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 US* 1999;96:7563-8.
6 Goldstein J, Eisen G, Bensen W, Agrawal N, Singh G, Pavelka K, et al. SUCCESS in osteoarthritis (OA) trial: celecoxib significantly reduces the risk of upper gastrointestinal (UGI) hospitalizations compared to diclofenac and naproxen in 13,274 randomized patients with OA. *Abstract presented at EULAR 2001, European Congress of Rheumatology* Available at: www.eular.org/eular2001/AbstractsOnline.cfm [Accessed 24 June 2002].
7 Celebrex labeling, revised 7 June 2002. Available at: www.fda.gov/cder/foi/label/2002/20998s009lbl.pdf (accessed 28 June 2002)

NHS Direct audited

NHS Direct is value for money and improving

Editor—George selectively highlights negative points from the positive report by the National Audit Office on NHS Direct, drawing conclusions which merit further scrutiny.[1]

He notes no visible effect on demand for NHS services since NHS Direct started. Reducing demand was not its primary objective. It aimed to provide the public with a confidential, reliable, and consistent source of professional advice 24 hours a day with easy access to comprehensive health information.[2]

He later comments that the "inevitable consequences" of nurse telephone advice is "to fill a health system with people who do not need to be there," directly contradicting the overall unchanged demand he cited earlier. The key is appropriate use of of healthcare services. NHS Direct undoubtedly advises some callers to access NHS services who would not have done so in the absence of the telephone advice service. Equally, there is frequent redirection of callers from a previous intention to call their general practitioner or attend accident and emergency department to self care. The net numerical effect may be neutral, but the movement of callers above and below the waterline of the "iceberg of illness"[3] should increase appropriate use of services. Early evaluation did not address this, but such investigation is now under way.

Service use and awareness is not universal, but in a service that has only been nationally available for 18 months, this is not surprising. Specific campaigns targeted at hard to reach groups are now being developed.

Access to the content of NHS Direct Online is less limited than George suggests. It is available in touch screen format in 200 (expanding to 500 by 2004) health information points in the United Kingdom to be found in libraries, post offices, and health centres, so allowing use without internet access or computer skills. The content has also been used in recent pilot studies with digital television, which will potentially broaden internet access significantly. A popular part of the website is the self help guide, also available in hard copy.[4]

The sum of £45m represents around 0.1% of the total NHS budget. For this net outlay the NHS has delivered a service with over 7m consultations, with a safety record at least comparable to any other part of the NHS, a rate of satisfaction among users that has been consistently greater than 95%, and information management and availability far in advance of the rest of the NHS. It is now acting as a focus point for the integration of out of hours and emergency care around a single prioritisation and assessment process. Value for money? I think so.

Mike Sadler *medical director, NHS Direct Hampshire and Isle of Wight*
Strawberry Fields, Berrywood Business Village, Hampshire SO30 2UN
mike.sadler@hants-iow.nhsdirect.nhs.uk

1 George S. NHS Direct audited. *BMJ* 2002;324:558-9. (9 March.)
2 Department of Health. *NHS Direct—a new gateway to health care*. London: Stationery Office, 2002.
3 Hannay D. *The symptom iceberg*. London: Routledge and Kegan Paul, 1979
4 Banks I. *The NHS Direct self-help guide*. London: Stationery Office, 2001.

Audit neglected children

Editor—In its audit of NHS Direct the National Audit Office failed to examine how it serves children[1]; at least 20% of calls to NHS Direct are about the under 5s.

Perhaps if the children had made the calls themselves they would have been identified as sufficiently distinct a group to warrant at least some interest from the auditors. This audit is so superficial that the auditors failed to wonder whether adults might be calling about children and, if so, what specific systems were in place to respond, how appropriate the responses might be, and whether any concerns might emerge. What could their reasoning have been?

The reasoning was probably all too typical—that children, being small people, must have correspondingly small needs. The National Audit Office would do well to commission an audit specifically about children and NHS Direct. That might restore their credibility. It would certainly prove a service to children and their parents.

Neil J McLellan *consultant paediatrician*
Birmingham Children's Hospital, Birmingham B4 6NH
neil.mclellan@bhamchildrens.wmids.nhs.uk

1 George S. NHS Direct audited. *BMJ* 2002;324:558-9. (9 March.)

Telephone consultations in general practice should be tested

EDITOR—It is interesting to see evidence that NHS Direct is not value for money.[1] Having dealt with patients (both in surgery and at a general practice cooperative) who have been told by NHS Direct to see a general practitioner urgently but turn out to have what is clearly a trivial problem, this fits with my own experience.

It should not come as a surprise to anyone. Nurses' skills and training lie largely in managing and organising care for defined problems, not in sorting out unformulated ones: this is a different and complex skill for which general practitioners receive 10 years' training. Computer algorithms cannot replace this skill.

Although there is obviously a call for telephone advice, general practice has failed to meet this need. I recently assessed the impact of a telephone surgery run by general practitioners in a practice in east London. Around a quarter of the general practitioners' contacts were by telephone. This is similar to rates in the United States and other countries but much greater than is usual in the United Kingdom. The telephone was particularly suited to dealing with administrative and organisational queries and questions about drug regimens, side effects, and results, but a considerable proportion of new problems could also be dealt with by this method.

These findings are not new: several authors have reported similar experiences both in the United Kingdom and elsewhere. Yet few general practitioners in the United Kingdom encourage telephone consultation, and there has been no move by the government to support this as a way to solve problems of access. We need to understand more about the reasons for this lack of enthusiasm. Perhaps the negative attitude of the General Medical Council is partly responsible, or lack of training in telephone consultation skills. Perhaps both doctors and patients need guidelines on how to make the best use of the telephone.

It seems reasonable to wonder whether telephone advice from an experienced diagnostician with the benefit of a complete medical record might be more cost effective than NHS Direct. Research to test such hypotheses is needed.

Peter D Toon *senior lecturer in primary care*
Department of Primary Care and Population Sciences, University College, London N19 3UA
petertoon@aol.com

1 George S. NHS Direct audited. *BMJ* 2002;324:558-9. (9 March.)

Telephone triage in Western Australia is cheaper than NHS Direct

EDITOR—George's editorial commenting on the perceived lack of value of NHS Direct fails to compare the service with telephone triage services elsewhere.[1] Contrary to his view, evidence is now available suggesting that telephone triage can improve access to health care for disadvantaged groups, in particular rural and remote communities and indigenous populations.

Health Direct, funded by the health department of Western Australia, is Australia's first large telephone triage project.[2] It has received over 315 000 calls in its first 24 months of operation. Call handling targets have been met, and over 11% of the calls received in the past 12 months have been from rural and remote communities, many of which are among the most isolated in the world.

Health Direct has become a key strategy to help stretched rural health services after hours. It has achieved similar levels of customer satisfaction to those reported for NHS Direct, but at a cost per call of less than half that of NHS Direct[2 3] and an annual per capita cost of $A2.51 (£0.87). Additional Australasian data support the contention that telephone triage can address inequities in access to primary care services for low income and indigenous populations.[4 5] George laments the lack of serious disease among people who call NHS Direct, but there is more to medicine than serious illness.

Ongoing analysis of comparative data from similar services elsewhere in relation to performance, cost of delivery, and impacts on health and social systems may provide a valuable benchmark in furthering this debate.

Andrew J Wilson *managing director*
McKesson Asia-Pacific, PO Box 4069, Lane Cove, NSW 2066, Australia
andrew.wilson@mckesson.com.au

Valendar Turner *medical director*
Health Direct, PO Box 450, Leederville, WA 6903, Australia

1 George S. NHS Direct audited. *BMJ* 2002;324:558-9. (9 March.)
2 Turner VF, Bentley PJ, Hodgson SA, Collard PJ, Drimatis R, Rabune C, et al. Telephone triage in Western Australia. *Med J Aust* 2002;176:100-3.
3 Munro J, Nicholl J, O'Cathain A, Knowles E, Morgan A. *Evaluation of NHS Direct first wave sites. Final report of phase 1 research.* Sheffield: Medical Care Research Unit, University of Sheffield, 2000.
4 St George IM, Cullen MJ. The Healthline pilot: call centre triage in New Zealand. *N Z Med J* 2001;114:429-30.
5 Wilson A, Cullen M. The Greater Murray AccessLine. *Aust Psychol* 2001;9:351-5.

Nurses as NHS gatekeepers

Nurses tend to use social rather than medical model of care

EDITOR—The BMA's suggestion that nurses might act as gatekeepers[1] is not supported by the evidence from a three year study of NHS Direct West London linked to a cooperative population of 0.9m.[2]

Although the computer triage based on the medical model is intended to manage and control demand, evidence suggests that when nurses stopped being employed by the cooperative and became employees of NHS Direct, general practice referral patterns changed considerably. In particular, requests for visits by a general practitioner have now been transposed into the less costly requests for telephone advice from mobile visiting doctors. The nurses moved from being gatekeepers to being patient advocates, using a social rather than the more limited medical model of care. This greater sensitivity may also be why patients like talking to nurses.

The work by MORI, which performed the survey leading to the suggestion, is pointless if patients' attitudes were not surveyed at different points in the need continuum, as attitudes to health access change from the well to the ill.

Lastly, the BMA's proposal should surely have been in a joint report with the Royal College of Nursing if role boundaries are to be renegotiated in this way, especially if it has any chance of maintaining credibility with the public.

Annabelle L Mark *professor of healthcare organisation*
Middlesex University Business School, London HA5 5AG
a.mark@mdx.ac.uk

1 Dyer O. BMA says nurses could become NHS gatekeepers. *BMJ* 2002;324:565. (9 March.)
2 Mark AL, Shepherd IDH. *Don't shoot the messenger—an evaluation of the transition from HARMONI to NHS Direct in west London.* London: Middlesex University Business School, 2001.

Using nurses might influence developing countries

EDITOR—I found the BMA's proposal that nurses could become NHS gatekeepers interesting.[1] I come from a developing country where nurses and other staff (in some cases minimally trained on the job) are the mainstay of a health system that can afford few doctors.

The trend in several countries, such as Sudan, Somalia, and Barbados, has been to reduce the clinical duties of nurses and auxiliary workers. From my observations this reduction seems to be proportional to the countries' fortunes. This implies that people ask for the best available within the means. In Sudan I was the only doctor to a community of 85 000 spread out in villages over about 80 km^2.

A technological knowledge base demands highly trained staff. In a society with high demands gatekeeping requires basic skills in clinical diagnosis. It should also be supported by the communication and technology that can give prompt answers and aid decisions on whether to refer or request investigations or treatment, etc.

The proposal will be a blessing if it works. Besides being cost effective it might halt the trend to reduce the clinical role of staff and strengthen primary health care in

developing countries, which tend to copy developed countries.

Anthony Lwegaba *lecturer in social and preventive medicine*
University of West Indies, School of Clinical Medicine and Research, Queen Elizabeth Hospital, Bridgetown, Barbados, West Indies
lwegaba@lycos.com

1 Dyer O. BMA says nurses could become NHS gatekeepers. *BMJ* 2002;324:565. (9 March.)

Can nurse practitioners provide equivalent care to GPs?

Nurses and doctors working together can complement each other

Editor—Horrocks et al consider whether nurse practitioners working in primary care can provide equivalent care to doctors.[1] It has reinforced much of my experience as a nurse running a general practice under the auspices of a personal medical services pilot.

My practice has been open for three and a half years and operates the philosophy that patients should be seen by the most appropriate person to help them. On average, 65% of patients choose to see the nurse and a small proportion of those require referral back to a general practitioner.

Concerns for patient satisfaction arose from my need to ensure the acceptability of a nurse led service to patients. By establishing a mechanism to listen to patients' concerns about their health and including patient representatives in the management of the practice, we have developed a service with a high standard of patient satisfaction.

Some may argue that patient satisfaction is a fickle concept that is impossible to measure and that such commercial concepts have no place in healthcare provision. Using the general practice assessment survey's patient satisfaction questionnaire, our patients showed that they felt valued and commented that they thought that nurses made better listeners and had more time than doctors. In reality, this perceived time factor issue is not the case in my practice, as the nurse often sees more patients than the doctor and the doctor tends to spend more time with the patients than the nurse. This is not surprising as he tends to see important medical problems for most of his surgery. Patients also expressed confidence in the nurses' clinical abilities and are surprised that such a system has not been thought of before.

Whether this has anything to do with nurses versus doctors, I am not sure. Patients feel that they can relate better to a nurse as an equal. For this reason they may be able to communicate their needs more readily. Underneath the titles is the concept that the professional with the right skill should be the person the patient sees. Nurses and doctors working together can complement each other, generating an environment where patients receive prompt, competent care and staff enjoy job satisfaction.

Catherine Baraniak *project lead*
Meadowfields Nurse-Led General Practice, Chellaston, Derby DE73 1TQ
Catherine.Baraniak@gp-c81665.nhs.uk

1 Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002;324:819-23. (6 April.)

More studies of these nurses' technique are needed

Editor—Horrocks et al state in the title of their systematic review that its focus is the role of nurse practitioners in primary care.[1] They concede that ambiguity exists over the definition of a nurse practitioner yet go on to include studies involving nurses working in hospital departments.[2–4]

Before large sums of money are thrown at such projects, it would be wise to compare like with like. In addition, policy implementers ought to consider several other points not covered by Horrocks et al's review.

Firstly, general practitioner registrars in their final three months of training have a minimum of four years of postgraduate work experience, yet are deemed unfit to practise without first passing various elements of summative assessment. Nurse practitioners are not required to have their video consultation technique checked.

Secondly, if nurse practitioners wish to be considered as independent practitioners then they need to have their own comprehensive indemnity, so that litigation stops with them, rather than their employing practice.

Thirdly, during employment of a nurse practitioner on a three month trial basis, doctors at my practice asked our trainer to review each of the nurse practitioner's surgeries, as he would with a general practitioner registrar. The trainer thought that little insight was shown into why specific questions, investigations, or drugs were used. The basic understanding of the pathology and pharmacology lagged far behind the automated efficiency of following guidelines.

There certainly seems to be an evolving role for nurse practitioners, but future studies must clearly document the context in which patients are seen. There must also be some form of assessment of nurse practitioners' consultation technique, similar to that in general practice.

Ian O'Connor *general practitioner*
Oldcastle Surgery, Bridgend, South Wales CF31 3ED
elaineianoconnor@hotmail.com

1 Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002;324:819-23.
2 Hoekelman RA. What constitutes adequate well-baby care? *Pediatrics* 1975;55:313-26.
3 Saker M, Angus J, Perrin J, Nixon C, Nicholl J, Wardrope J. Care of minor injuries by emergency nurse practitioners or junior doctors: a randomised controlled trial. *Lancet* 1999;354:1321-6.
4 Cooper M. An evaluation of the safety and effectiveness of the emergency nurse practitioner in the treatment of patients with minor injuries: a pilot study. Glasgow: Accident and Emergency, Glasgow Royal Infirmary, 2001. (Typescript.)

Nurse practitioners increase access to quality health care for many patients

Editor—A few points in Horrocks et al's systematic review are worth discussing further.[1]

Firstly, they mention the lack of difference in health outcomes between nurse practitioners and doctors. Most of the trials included in their review are short term trials not designed to determine the health outcomes of chronic diseases such as diabetes, asthma, and hypertension, which may take decades to advance.

Secondly, they highlighted higher patient satisfaction among patients seen by nurse practitioners. Patient satisfaction correlates strongly with patient adherence[2] but, though important, is not a reliable measure of the standard of care. It certainly has no part to play in the ability to diagnose conditions and provide medical care. The fact that nurse practitioners spend more time with patients and order more tests proves their inability to diagnose conditions and provide subsequent care in a time limited and efficient manner.

The increasing shortages in the delivery of primary care facing developed nations have led to the question of how to provide cheaper care. This is central to the evolution of the nurse practitioner concept in the United States. We struggle to deal with issues concerning the cost of training health professionals, resource use, and the safety of health care.

It is naive to consider that the expert services given by doctors after their several years of training can be matched by the 24 months of training of a registered nurse. If this were true we would have stopped using doctors in primary care years ago; doctors would only be in specialty fields.

Nurse practitioners are not meant to be compared with or to replace doctors. Rather, they will increase access to quality health care for many patients whose needs are within the limited scope of their training. Unfortunately, a few of us go far beyond this simple goal and attempt to make comparisons when none exists.

Rahul Gupta *internist and primary care physician*
Seema Gupta *research associate*
Division of Internal Medicine, Florala Memorial Clinic, Florala, AL 36442-0219, USA
rsgu@hotmail.com

1 Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002;324:819-23. (6 April.)
2 Safran DG, Taira DA, Rogers WH, Kosinski M, Ware JE, Tarlov AR. Linking primary care performance to outcomes of care. *J Fam Pract* 1998;47:213-20.

More methodologically sound investigations are needed

Editor—Horrocks et al report higher levels of patient satisfaction among patients of nurse practitioners.[1] Though this is important, it is unclear whether patient satisfaction is a valid measure of the quality of practice. Moreover, of the five trials presented that looked at patient satisfaction, three asked prospective patients whether they found it

acceptable to see a nurse practitioner rather than a doctor.[2–4] Thus only those patients predisposed to accept nurse practitioners were included in the studies.

The authors also report that nurse practitioners undertook more investigations and had longer consultations than did doctors. The absolute difference of less than four minutes in consultation time is of questionable clinical importance and may reflect the practice setting used by the nurse practitioners. Whether the finding that nurse practitioners ordered more tests per patient is a marker of better or worse practice is unclear; it may inflate costs.

We believe that the authors' assertions about the quality of care are not supported by the data presented. The authors state that nurse practitioners identified physical abnormalities more often than doctors did, without qualifying that the cited data are from 1975 and refer only to well-baby examinations.[5] The observations that nurse practitioners made more complete records, communicated better, and were as proficient as doctors in ordering and interpreting *x* ray films are based only on patients with minor injuries seen in an accident and emergency department. It may not be reasonable to make the leap that such conclusions would hold true in adults more recently, outside the accident and emergency department, or in more seriously ill patients.

With these limitations in mind, we question the authors' ability to conclude that "nurse practitioners can provide care that leads to increased satisfaction and similar health outcomes." Even more dubious is the conclusion that they provided care of equivalent or superior quality to that provided by doctors.

We agree with the authors that the ability of nurse practitioners to identify rare but important health problems needs further research. Although the authors mention some of the weakness of their data, they propose that their review supports the increased involvement of nurse practitioners in primary care. We think that this conclusion is out of proportion to the data presented and that patients and policymakers would be better served by using this review as an indication of the need for more methodologically sound investigations.

William Rifkin *doctor*
Maimonides Medical Center, Brooklyn, NY 11219, USA
w drifk@hotmail.com

Arthur Rifkin *doctor*
Long Island Jewish Medical Center-Hillside Hospital, Glen Oaks, NY 11004, USA

John Horiszny *doctor*
Red Hook, NY 12507, USA

1 Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002;324:819-23. (6 April.)
2 Venning P, Durie A, Roland M, Roberts C, Leese B. Randomised controlled trial comparing cost effectiveness of general practitioners and nurse practitioners in primary care. *BMJ* 2000;320:1048-53.
3 Shum C, Humphreys A, Wheeler D, Cochrane MA, Skoda S, Clement S. Nurse management of patients with minor illnesses in general practice: multicentre, randomised controlled trial. *BMJ* 2000;320:1038-43.
4 Kinnersley P, Anderson E, Parry K, Clement J, Archard L, Turton P. Randomised controlled trial of nurse practitioner versus general practitioner care for patients requesting "same day" consultations in primary care. *BMJ* 2000;320:1043-8.
5 Hoekelman RA. What constitutes adequate well-baby care? *Pediatrics* 1975;55:313-26.

Australian GPs are beginning battle faced by other countries

Editor—It is fascinating that Australian doctors, though keen to manage public money to ensure that their personal incomes are maximised, are reluctant to engage in managing public money to achieve better outcomes for their patients.[1] Thompson is right in saying in her rapid response to Zinn's piece that general practitioners will receive no financial benefits and that the emphasis is on increasing prescribing quality and at the same reducing costs (bmj.com/cgi/eletters/324/7343/937#21623). There is much evidence that these are complementary goals, not conflicting ones.[2]

The reaction of the Australian Medical Association is the same as that of organised medicine to fundholding in Britain and to budget holding in New Zealand almost a decade ago. In New Zealand newly formed independent practitioner associations, similar in many respects to Australian divisions of general practice and the English primary care trusts, embarked on budget holding for pharmaceutical and pathology services.[3 4] Outrage followed from the New Zealand Medical Association, Royal New Zealand College of General Practitioners, pathologists, and specialists. But visionary general practice leaders, with a strong focus on quality, saw the opportunities being offered and pressed on; now nearly 90% of all general practitioners are in some form of pharmaceutical budget management. The process is strongly driven by quality.[5]

The main threat to the Australian divisions, though, is not the medical profession. As we have found in New Zealand, it is the bureaucrats who see the process as primarily a cost cutting strategy, not a quality strategy. Australian bureaucrats seem to support general practice, but the New Zealand experience will be relevant to Australia.[2–5] Success in such a controversial strategy depends on clearly articulated, common goals; full collaboration between all parties; appreciable financial support; and reasonable expectations about what can be achieved.

Laurence A Malcolm *professor emeritus*
Aotearoa Health, RD1 Lyttelton, New Zealand
laurence.malcolm@cyberxpress.co.nz

1 Zinn C. News roundup: Australia's plan to reduce drug spending attacked by doctors. *BMJ* 2002;342:937. (20 April.)
2 Malcolm L, Medawar C. *Towards a consumer-led strategy for the quality use of medicines in Australia.* Canberra: Department of Health and Aged Care, 2001. (www.health.gov.au/hbd/qum/index.htm)
3 Malcolm L, Mays N. New Zealand's independent practitioner associations: a working model of clinical governance? *BMJ* 1999;310:1340-2.
4 Malcolm L, Wright L, Barnett P. *The development of primary care organisations in New Zealand.* Wellington: Ministry of Health, 1999. (www.moh.govt.nz/)
5 Malcolm L, Wright L, Barnett P, Hendry C. *Clinical leadership and quality improvements in primary care organisations in New Zealand.* Auckland: Clinical Leaders Association of New Zealand, 2001. (www.clanz.org.nz)

Oncologist who stood up to US insurance companies lost work

Editor—From early reports of the success of high dose chemotherapy and autologous bone marrow infusion some oncologists, including me, recognised that the procedure was far from proved, dangerous, and expensive.[1] I was a reviewer for several insurance companies at the time and rejected all requests for payment for both primary and metastatic disease. The companies welcomed the first few opinions, but when lawsuits started to be filed consultation requests decreased and then stopped. Yet other physicians were still being sent consultations.

I got the message. It was cheaper to grant payment than to fight in court and lose. But I could not change my opinion, because the evidence for effectiveness did not exist. I saw the lack of randomisation, the selected patients, the data from studies for other purposes—it was all there, and obvious.

I was near retirement and wanted to start a new career reviewing insurance cases for appropriateness of care, but as companies dropped me from their lists I could not obtain consultations. Fortunately, I was asked to head the oncology division at a teaching hospital. For the next six years until retirement I never referred a patient with breast cancer for high dose chemotherapy and autologous bone marrow infusion; the patients and I were all the better for it.

I recall the article on variations and attitudes of insurance companies in the *New England Journal of Medicine*, and was surprised that it was published.[2] Things were getting out of hand, and as a labouring doctor I had little voice. I suppose I was one of the few people not surprised at the final outcome in 1999.

I now investigate anomalous claims of the complementary medicine system. Here again, pressure groups, deluded elected officials, and officials lacking wisdom and principle are mandating payment for even more implausible methods.

Wallace Sampson *editor*
Scientific Review of Alternative Medicine, Los Altos, CA 94022, USA
wisampson@cs.com

1 Welch HG, Mogielnicki J. Presumed benefit: lessons from the American experience with marrow transplantation for breast cancer. *BMJ* 2002;324:1088-92. (4 May.)
2 Peters WP, Rogers MC. Variation in approval by insurance companies of coverage for autologous bone marrow transplantation for breast cancer. *N Engl J Med* 1994;330:473-7.