

Differences between clinical trials and postmarketing use

Karin Martin,¹ Bernard Bégaud,¹ Philippe Latry,² Ghada Miremont-Salamé,¹ Annie Fourier¹ & Nicholas Moore¹

¹Département de Pharmacologie, Université Victor Segalen, Bordeaux 2. Bordeaux and ²Service médical d'Aquitaine. CNAM-TS. Bordeaux, France

Correspondence and reprint requests

Karin Martin, Département de Pharmacologie, Université Victor Segalen Bordeaux 2, 33 076 Bordeaux cedex, France.
Tel: + 33 5575 71561
Fax: + 33 5575 74660
E-mail: karin.martin@pharmaco.u-bordeaux2.fr

Keywords

clinical trials, drug safety, drug utilization, effectiveness, postapproval

Received

24 December 2002

Accepted

15 July 2003

Aims

Clinical trials constitute the gold standard to assess the efficacy and safety of new medicines. However, because they are conducted in standardized conditions far from the real world of prescription and use, discrepancies in patient selection or treatment conditions may alter both the effectiveness and risks. On the basis of three examples, our objectives were to study the differences between the characteristics of treated populations and treatment patterns in clinical trials and in postmarketing settings and to discuss the potential consequences on actual efficacy and safety.

Methods

Treated populations were compared with patients included in premarketing clinical trials. Comparisons were made on the basis of demographic characteristics and treatment patterns.

Results

Whatever the indicator and the drug studied, differences were observed: from 0.04% to 63% for tacrine, from 0% to 37% for celecoxib and from 6% to 52% for simvastatin, with possible consequences on the effectiveness and safety of the drug concerned. Our results confirm the under-representation of women and elderly patients in premarketing clinical trials, e.g. an M : F ratio of 4.6 in clinical trails of simvastatin vs 1.0 in the joint population. Moreover, the concomitant use of medicines was made extremely restrictive by the protocols of these trials while this was not the case in the postmarketing phase. This has possible consequences on the effectiveness and safety of the drug concerned.

Conclusions

These results plead for systematic *ad hoc* observational postmarketing studies for any novel and/or expensive medicine to assess the relevance of premarketing data.

Introduction

Clinical trials are the indisputable gold standard to demonstrate the efficacy of new medicines before marketing. However, they should be conducted in standardized conditions and exclude certain types of patients and/or certain situations, in order to make the statistical evaluation of efficacy and safety more efficient. Moreover, many clinical trials include a run-in period which reduces even further the interindividual variability [1]. In the real world of prescription and use, several parameters are

expected to vary which may alter both effectiveness and safety [2–4]. These changes are rarely explored though they may be relevant to real-world safety. For example, in the 1980s, hundreds of serious cases of cardiac arrhythmia, some fatal, occurred in patients over 70 years treated with bepridil. This age class had been excluded from premarketing studies [5]. Wieringa *et al.* [6, 7] have published two studies comparing pre- and postapproval settings and discussed possible consequences on effectiveness. In the first, they compared the

demographic characteristics between patients participating in phase III trials of cardiovascular drugs and those of actual users after approval. In the second, they compared comorbidity. In the present paper, we also focused on safety by considering a third population, the so-called *injured population*, i.e. patients having presented with an adverse drug reaction (ADR) attributed to the drug concerned and reported to the national pharmacovigilance system. We also considered more parameters which may alter the validity of phase III data both in terms of efficacy and safety.

For the purpose of this comparison, we chose three relevant drugs:

Tacrine: this drug was chosen as the 'historical' gold standard because it is intended to be used in a very specific population, i.e. patients with Alzheimer's disease. For this reason, the first 5000 patients treated in France were included in a systematic follow-up cohort with precise control of conditions of prescription and use.

Celecoxib and *simvastatin*: these are high-profile drugs with very many users, so they offer scope for investigating the large-scale consequences of any differences observed.

Methods

The three medicines studied (tacrine, simvastatin and celecoxib) belong to different pharmacotherapeutic classes and were launched in France in 1988, 1994 and 2001, respectively.

For each drug we defined three populations: (i) a *target population*, i.e. the population considered for inclusion in the clinical trials and to whom the conclusions concerning efficacy and safety may be applied; (ii) a *joint population* actually treated in 'the real world'; and (iii) an *injured population* defined as patients having presented an ADR attributed to the drug concerned and reported to the French pharmacovigilance system.

The data for the target populations were obtained from the data on clinical trials considered for approval in France, some of which have been published [8–23]. For tacrine, data on the joint population were taken from the PACO cohort (Pharmacosurveillance Active du COgnex) which included the first 5000 patients treated with this drug in France, from 1994 to 1999. For celecoxib and simvastatin, prescription data were obtained from a random sample of 500 prescription forms submitted for re-imbursement in January 2001 and recorded in the database of the Caisse Nationale d'Assurance Maladie des Travailleurs Sociaux (CNAM-TS) of the Aquitaine region (South-Western France).

For tacrine, the 'injured' population was defined by using data from the first year of the PACO follow-up. For the other two drugs, we assessed the spontaneous report cases recorded in the French Pharmacovigilance database between January 1st and May 15th 2001 for 100 and 200 mg celecoxib, and between January 1st 1999 and May 15th 2001 for 20 mg simvastatin. These durations of surveillance were chosen in order to have at least 100 reports for each drug.

If available, the following indicators were collected and compared for each population: age, gender, dosage, duration of treatment, comedications, indication and contra-indications.

The term of 'contra-indication' is used in a broad sense, i.e. a situation for which the efficacy and safety profile of the drug was not established. For example, the percentage of 'contra-indications' was the percentage of prescription forms including at least one drug excluded during pre-approval clinical trials.

In order to avoid too restrictive a definition of the target population, the extreme values of each indicator in the clinical trials were used to define the range of the reference, even if corresponding to very few patients. For example, if clinical trials had included patients between the ages of 18 and 65 years, a deviate was considered for patients of the joint population who were <18 or >65 years.

Differences were expressed as percentages of patients for which one or several parameters was/were out of the range of the values found in the target population, i.e. patients included in the clinical trials considered for approval. These trials constituted the gold standard for comparisons. For simvastatin and celecoxib, since post-approval data were obtained from a sample of the whole joint population, 95% two-sided confidence intervals (CI) were computed using the normal, binomial or Poisson methods as appropriate. This was not the case for tacrine, since the whole exposed population was studied. Qualitative and categorical data were compared by using the Chi²-Pearson test and the significance level was set at 0.05.

The potential consequences of the observed differences were discussed on the basis of both effectiveness and risk. For effectiveness, we considered that from a statistical point of view, the results of an experience apply only to a population having the same characteristics and to the same conditions of use. Consequently, drug efficacy was considered to be questionable in situations not represented during pre-approval clinical trials. The same applied for safety (potential risk). The purpose of comparing the injured population with the joint and target populations was to identify possible risk

factors if a given characteristic was over-represented in the injured population.

Results

For tacrine, the two original clinical trials included 863 patients. The 11 clinical trials for celecoxib included 5557 patients, and the trial for simvastatin included 2221 patients (Table 1). As previously mentioned, the PACO cohort included 4996 patients. For tacrine, the number of ADRs (whatever the level of seriousness) identified during the first year of the follow-up was 1173. The numbers of spontaneous reports of ADRs for celecoxib and simvastatin in the study period were 233 and 112, respectively.

Differences between the target and joint populations, for each selected indicator and for each drug, are shown in Tables 1 and 2. The magnitude of the difference varied according to the parameter and the drug considered. As expected, differences were lower for tacrine, except for dose schedule (63%) (defined by a 6-week period before reaching the maximal dosage) and for the dura-

tion of treatment (61%), e.g. patients treated for a longer period than in clinical studies.

Table 3 shows the aggregated percentages of patients who were outside the target population for at least one indicator. The percentages remained roughly the same, although were lower for tacrine.

Gender

Comparison with clinical trials showed a marked over-representation of women in the population actually treated with tacrine and simvastatin (Table 2).

Age

Similarly, the comparison with clinical trials showed a dramatic over-representation of patients over 70 years in the population treated with celecoxib (joint and injured populations) (Figure 1). Patients under 35 years and over 70 years were both over-represented in the population treated with simvastatin (Figure 2): 76% of patients were under 65 years in the target population while 57% were over this age in the joint population. We used mean

Table 1

Differences between target, joint and injured populations

Target population	Target limits	Joint population (% difference + CI)	'Injured' population (% difference + CI)
<i>Tacrine</i>			
<i>n</i> = 863		4996	1173
Age range	49–95 years	0.3	0.3
Duration of treatment	≤7.5 months	61	4.3
Dose	>120 mg after 6 months ¹	4	2.9
Dose	≤160 mg	0.04	0.2
Dose schedule	40–80, 80–120, 120–160 mg	63	NA
Contra-indications	NA	39	NA
Indication	10 ≤ MMSE ² ≤ 28	9	NA
<i>Celecoxib</i>			
<i>n</i> = 5557		500	233
Age range	20–93 years	1.0 (0.1, 1.8)*	0 (0, 1.5)*
Duration of treatment	≤24 weeks	–	0 (0, 1.5)
Contra-indications	NA	37 (33, 42)	31 (27, 35)
<i>Simvastatin</i>			
<i>n</i> = 2221		500	112
Age range (men)	35–70 years	39 (33, 44)*	32 (21, 44)*
Age range (women)	40–70 years	52 (46, 59)*	46 (30, 61)*
Duration of treatment	≤6.3 years	NA	6.0 (0.3, 12)
Contra-indications	NA	43 (39, 48)	27 (19, 36)

% difference: percentage of patients outside the range of values found in clinical trials; CI: 95% confidence interval; * $P < 0.01$ Chi² analysis on age distribution; ¹According to premarketing clinical trials, a dose under 120 mg after 6 months of treatment was ineffective; ²MMSE: Minimal Mental State Examination; NA: not applicable.

age instead of the central tendency for age because, except for celecoxib, the latter was not available for age in the pivotal phase III trials. Moreover, the populations could be more adequately compared by using age distribution.

Contra-indications

Table 1 shows that the percentage of patients in 'real life' (joint population) who used medicines excluded by the protocol during clinical trials varied from 37 to 43%.

Table 2

Comparison of male : female ratios between the three populations

	Target	Population Joint	Injured
Tacrine	0.9	0.6	0.5
Simvastatin	4.6	1	1.5
Celecoxib	0.4	0.5	0.5

This point is interesting because it questions the validity of extrapolating from the results of these trials, both in terms of effectiveness and safety. Conversely, comparison of the injured and joint populations did not make it possible to identify any significant over-representation of a given characteristic which might act as a risk factor for ADR (Table 4).

Since complete and/or detailed data were not available in the reports of clinical trials, comparison of comedICATIONS between the target and joint populations which could be a proxy for comorbidity was possible only for simvastatin. This clearly revealed different patterns between these two populations (Table 5).

Discussion

Unlike previously published studies, the main interest of the present paper is that the comparison involved three populations and was based on seven different parameters.

Our results are consistent with studies having confirmed the under-representation of women and elderly patients in premarketing clinical trials [6, 24, 25]. The consequences of such differences may be important both

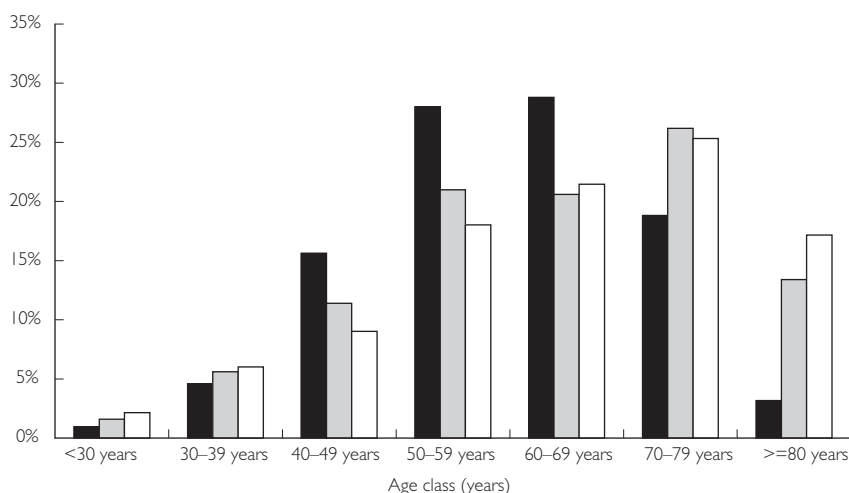
Table 3

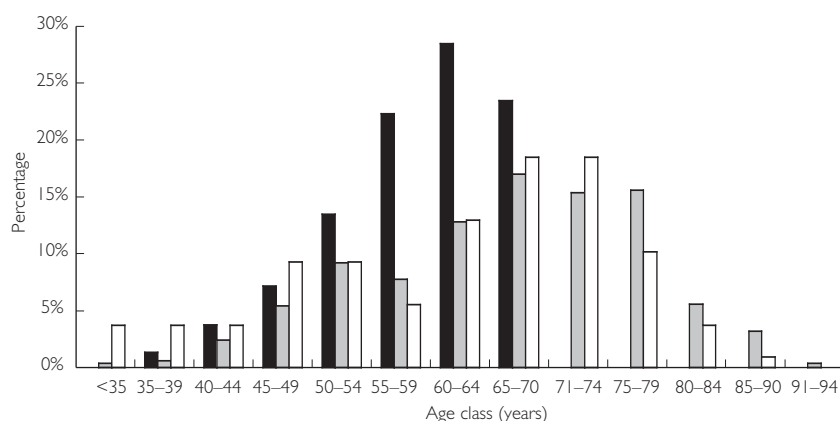
Percentages of patients outside target for at least one indicator, compared with the clinical trials population

Differences	Tacrine	Celecoxib	Simvastatin
Target population and observed joint population	5.8% (age, dose, indication)	33% (age, contra-indication)	63% (age, gender, contra-indication)
Target population and 'injured' population	4.6% (duration of treatment, age)	29% (age, contra-indication)	49% (age, gender, contra-indication)

Figure 1

Age distribution for the three populations for celecoxib. Target population (■), joint population (■) and injured population (□)



**Figure 2**

Age distribution for the three populations for simvastatin. Target population (■), joint population (■) and injured population (□)

Table 4

Assessment of possible risk factors (comparison of observed joint and injured populations)

Indicators	Joint population (1) (% difference and 95% CI)	Injured population (2) (% difference and 95% CI)	Ratio (1:2)
<i>Tacrine</i> (n)	4996	1173	
Age range (years)	0.3	0.3	1.0
Duration of treatment (months)	61	4.2	0.1
Daily dose	0.04	0	–
<i>Celecoxib</i> (n)	500	233	
Age range	1.0 (0.1; 1.8)	0 (0; 1.5)	–
Mean age (years)	63 (62, 65)	65 (63, 67)	1.0 (0.9; 1.1)
Contra-indications	37 (33, 42)	31 (27, 35)	0.8 (0.7; 1.1)
<i>Simvastatin</i> (n)	500	112	
Age range (men)	39 (33, 44)	32 (21, 44)	0.8 (0.5; 1.3)
Age range (women)	52 (46, 59)	46 (30, 61)	0.9 (0.5; 1.3)
Mean age (men)	65 (63, 66)	61 (58, 65)	0.9 (0.9; 1.0)
Mean age (women)	68 (67, 69)	62 (58, 68)	0.9 (0.8; 1.0)
Contra-indications	43 (39, 48)	27 (19, 36)	0.8 (0.7; 1.0)

CI: confidence interval.

for efficacy and safety. For example, efficacy could be different in women because of changes in pharmacokinetic parameters [26, 27]. This concern is particularly true for simvastatin since women accounted for 50% of the joint population while few had been included in clinical trials (18%). The same can be said for age since 42% of patients in the joint population were over 65 years, i.e. the limit of inclusion in clinical trials.

The use, in the joint population, of drugs excluded during clinical trials (from 27 to 43%) may also have altered the efficacy and safety of the drug considered. Obviously, exclusion of certain drugs during clinical trials is not only justified by safety concerns but may minimize possible interactions with the parameters

under study. Nevertheless, the differences in treatment patterns between clinical trials and the joint population question the validity of extrapolating from the results of preapproval clinical trials to the population actually treated. Indeed, the potential consequences of such differences should be assessed. Percentages of coprescription markedly differed between patients treated with simvastatin and those included in clinical trials. This could lead to differences in comorbidity associated with these populations, and again make the extrapolation of the results of preapproval clinical trials questionable. Moreover, these differences between the percentages are probably under-estimated since they were based on medicines marked on the same prescription form and

Table 5

Comparison of comedications between target and joint populations for simvastatin

Drugs	Target population (2) % patients	Joint population (1) % patients + CI	Ratio (1 : 2)	95% CI
Aspirin	37	12 (9, 15)	0.3	0.25, 0.41
β-adrenoceptor blockers	57	25 (21, 29)	0.4	0.38, 0.51
Calcium-channel blockers	32	20 (17, 24)	0.6	0.52, 0.75
Thiazide diuretics	7.0	0.2 (0.01; 1.1)	0.03	0.004, 0.2
Warfarin	1.0	3.8 (2.1, 5.5)	3.8	2.1, 7.0
Isosorbide mono/dinitrate	31	0 (0; 0.7)	–	–

CI: 95% confidence interval.

not on a face-to-face interview with the patient, e.g. self-medication.

For most of the indicators used, marked differences were observed between target and joint populations. The comparison was based on fewer indicators for celecoxib and simvastatin because the database of the reimbursement system did not provide any direct information on factors such as indication, actual daily dose, duration of treatment or drugs prescribed on a different prescription form. Observed differences between clinical trials and the 'real world' were less marked for tacrine than for the other two drugs. This was expected since the early marketing phase of tacrine was subject to a protocol (observational cohort study). Theoretically, in such strict conditions no differences would have been expected but this could be due to the fact that tacrine was the first and only drug made available for the treatment of a particularly serious condition, i.e. Alzheimer's disease, so it was tempting for practitioners to try it in patients outside of the approved indications.

As is always the case in observational studies, the validity of our results depends upon the representativeness of the samples used to describe the joint and injured populations.

For tacrine, the data were taken from a cohort study having enrolled all subjects treated with this drug, thus eliminating any potential selection bias. Concerning the other two drugs, the representativeness could appear questionable since the analysis was based on part of the joint population, i.e. a sample of 500 patients. However, because they were randomly extracted, these samples can be considered as representative of the whole population of the CNAM-TS database.

Moreover, since these two drugs are expensive, most patients likely sent their prescription form for re-

imbursement within a short period of time, whatever their socio-economical level, thus minimizing any potential selection bias. Considering the type of drugs studied, any biases due to seasonal changes in prescription and use can also be ruled out.

For celecoxib and simvastatin, the injured populations were defined by the reports found in the French pharmacovigilance database. Because of the unavoidable under-reporting that jeopardizes such surveillance systems, the injured population might not necessarily be representative of all patients having presented an ADR with these drugs. However, no data or previous findings support such a hypothesis.

Even so, it is extremely difficult to evaluate the possible consequences of the marked differences observed in our study. In terms of efficacy, implementation of postmarketing *ad hoc* studies such as pragmatic trials is no doubt required. Comparing the characteristics of the injured population, data readily available from pharmacovigilance databases, with those of target and joint populations appears to be a straightforward solution to identify possible risk factors, even though we did not find such factors with the drugs we studied. It would be interesting to replicate the approach proposed in other countries, mainly those in which medical databases are available, such as the G.P.R.D in the United Kingdom. This would allow a better description of indications and contraindications.

Conclusions

The marked differences observed in our study confirm that pre- and postapproval are two distinct worlds. At least for some types of patients, these differences could have consequences in terms of both effectiveness and safety. In view of the size of the exposed population in

the real world of prescription and use, it would be wise to conduct *ad hoc* evaluation studies when the actual population of users differs markedly from that of clinical trials in one or several indications. In this way, it would be possible to test the relevance of premarketing data for use in 'real life' settings.

This work was funded as a research project by a grant from the Fondation pour la Recherche Médicale (Paris, France).

The authors are indebted to the Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSaPS), to the Caisse Nationale d'Assurance Maladie d'Aquitaine and to Parke-Davis for their authorization to publish data from the different databases mentioned in this paper.

We also thank Ray Cooke who kindly supervised the English of this paper.

References

- Pablos-Méndez A, Barr G, Shea S. Run-in periods in randomized trials. Implications for the application of results in clinical practice. *JAMA* 1998; 279: 222–5.
- Black N. Why we need observational studies to evaluate the effectiveness of health care. *Br Med J* 1996; 312: 1215–18.
- Hemminki E. Problems of clinical trials as evidence of therapeutic effectiveness. *Soc Sci Med* 1982; 16: 711–12.
- Ray WA, Griffin MR. Evaluating drugs after their approval for clinical use. *N Engl J Med* 1993; 329: 2029–32.
- Jaillon P, Girard M, Ponsonaille J. Tolérance électrocardiographique du bédridil (cordium). *La Lettre Du Pharmacologue* 1991; 5: 202–6.
- Wieringa NF, De Graeff PA, Van der Werf GT, Vos R. Cardiovascular drugs. discrepancies in demographics between pre- and post-registration use. *Eur J Clin Pharmacol* 1999; 55: 537–44.
- Wieringa NF, Vos R, Van der Werf GTH, Van der Veen WJ, De Graeff PA. Co-morbidity of clinical trial versus 'real world' patients using cardiovascular drugs. *Pharmacoepidemiol Drug Saf* 2000; 9: 569–79.
- Bensen WG, Fiechtner JJ, McMillen JL, et al. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. *Mayo Clin Proc* 1999; 74: 1095–105.
- Bensen WG, Zhao SS, Burke TA, et al. Upper gastrointestinal tolerability of celecoxib, a cox-2 specific inhibitor, compared to naproxen and placebo. *J Rheumatol* 2000; 27: 1876–83.
- Emery P, Zeidler H, Kvien TK, et al. Celecoxib versus diclofenac in long-term management of rheumatoid arthritis: randomised double-blind comparison. *Lancet* 1999; 354: 2106–11.
- Farlow M, Gracon SI, Hershey LA, Lewis KW, Sadowsky CH, Dolan-Ureno J. A controlled trial of tacrine in Alzheimer's disease. *JAMA* 1992; 268: 2523–8.
- Geis GS, Hubbard RC, Woods EM, et al. Efficacy of celecoxib, a cox-2 specific inhibitor, in osteoarthritis of the hip. *Arthritis Rheum* 1999; 42(Suppl): S144.
- Goldenberg MM. Celecoxib, a selective cyclooxygenase-2 inhibitor for the treatment of rheumatoid arthritis and osteoarthritis. *Clin Ther* 1999; 21: 1497–513.
- Goldstein JL, Correa P, Zhao WW, et al. Reduced incidence of gastroduodenal ulcers with celecoxib, a novel cyclooxygenase-2 inhibitor, compared to naproxen in patients with arthritis. *Am J Gastroenterol* 2001; 96: 1019–27.
- Group Scandinavian Simvastatin Survival Study. Design and baseline results of the Scandinavian simvastatin survival study of patients with stable angina and/or previous myocardial infarction. *Am J Cardiol* 1993; 71: 393–400.
- Group Scandinavian Simvastatin Survival Study. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994; 344: 1383–9.
- Knapp MJ, Knopman DS, Solomon PR, et al. A 30-week randomized controlled trial of high-dose tacrine in patients with alzheimer's disease. *JAMA* 1994; 271: 985–91.
- Mckenna F, Borenstein D, Wendt H, Wallemark C, Lefkowitz JB, Geis GS. Celecoxib versus diclofenac in the management of osteoarthritis of the knee. *Scand J Rheumatol* 2001; 30: 11–18.
- Miettinen TA, Pyorala K, Olsson AG, et al. Cholesterol-lowering therapy in women and elderly patients with myocardial infarction or angina pectoris: findings from the Scandinavian Simvastatin Survival Study (4S). *Circulation* 1997; 96: 4211–18.
- Simon LS, Weaver AL, Graham DY, et al. Anti-inflammatory and upper gastrointestinal effects of celecoxib in rheumatoid arthritis. *JAMA* 1999; 282: 1921–8.
- Tive L. Celecoxib clinical profile. *Rheumatology* 2000; 39(Suppl 2): 21–8.
- Williams GW, Hubbard RY, Zhao W, Geis GS. Comparison of once-daily and twice-daily administration of celecoxib for the treatment of osteoarthritis of the knee. *Clin Ther* 2001; 23(2): 213–27.
- Zhao SZ, Dedhiya SD, Verburg K, Osterhaus JT. Celecoxib 200 mg administered once a day or in split doses has equal impact on health-related quality of life (HRQOL) of patients with osteoarthritis (OA). *Arthritis Rheum* 1999; 42(Suppl): S297.
- Schmucker DL, Vesell ES. Underrepresentation of women in clinical drug trials. *Clin Pharmacol Ther* 1993; 54: 11–15.
- Bugeja G, Kumar A, Banerjee AK. Exclusion of elderly people from clinical research: a descriptive study of published reports. *Br Med J* 1997; 315: 1059.
- Gleiter CH, Gundert-Remy U. Gender differences in pharmacokinetics. *Eur J Drug Metab Pharmacokin* 1996; 21: 123–8.
- Harris RZ, Benet LZ, Schwartz JB. Gender effects in pharmacokinetics and pharmacodynamics. *Drugs* 1995; 50: 222–39.