TOPICS IN CLINICAL TRIAL MANAGEMENT

Bridget-Anne Kirwan

TOPICS IN CLINICAL TRIAL MANAGEMENT

Bridget-Anne Kirwan

Topics in Clinical Trial Management

Acknowledgements

This thesis is based on studies conducted at SOCAR Research SA, Nyon, Switzerland.

The author gratefully acknowledges the patients, investigators, committee members and other participants in the PICO (PImobendan in COngestive heart failure) and ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine) trials; and the co-authors of the publications and documents incorporated in this thesis.

Financial support by SOCAR Research SA for the publication of this thesis is gratefully acknowledged.

ISBN 90-77595-30-9

Cover design and lay-out: To Κυμα
Printed by Optima Grafische Communicatie Rotterdam

Topics in Clinical Trial Management

Aspecten van het uitvoeren van experimenteel klinisch onderzoek

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 21 april 2004 om 13:45 uur
door

Bridget-Anne Kirwan

geboren te Waterford, Ierland

Promotiecommissie

Promotoren:	Prof.dr. J. Lubsen Prof.dr. A. Hofman
Overige leden:	Prof.dr. A.W. Hoes Prof.dr. P.J. Koudstaal Prof.dr. Th. Stijnen

For all my family

Publications and documents incorporated in this thesis

Chapter 2

The Pimobendan in Congestive Heart Failure (PICO) Investigators. Effect of pimobendan on exercise capacity in patients with heart failure: main results from the Pimobendan in Congestive Heart Failure (PICO) trial. *Heart* 1996;**76**:223-31

Chapter 3

Lubsen J, Kirwan BA. Combined endpoints: can we use them? *Stat Med* 2002; **21**:2959-70.

Chapter 4

Lubsen J, Poole-Wilson PA, Pocock SJ, van Dalen FJ, Baumann J, Kirwan BA, Parker AB. Design and current status of ACTION: A Coronary disease Trial Investigating Outcome with Nifedipine GITS. *Eur Heart J* 1998;**19** (Suppl I):I20-32.

Chapter 5

Kirwan BA, Lubsen J, Poole-Wilson PA on behalf of the ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) investigators. Treatment of angina pectoris: associations with symptom severity. *Int J Cardiol* 2004 (in press)

Chapter 6

Kirwan BA, Jonkers PR, van Dalen FJ, Lubsen J. SOCDAT®: a comprehensive clinical trial data and study management philosophy based on simultaneous display of scanned documents and corresponding database content. *Submitted for publication 2004.*

Chapter 7

ACTION Statistical Analysis Plan (SAP). Adapted from version 4, dated December 8, 2003 as approved by the Steering Committee on January 9, 2004. *Internal document ACTION study.*

Contents

Chapter 1

Introduction

"That principles of research do in fact exist or that there are persons qualified to expound them are not self-evident propositions".

Jerome Cornfield (1912-79) [1]

Today's clinicians are expected to practice evidence-based medicine, and to consult sources such as *Clinical Evidence*[2] when deciding on treatment. Results from randomised controlled clinical trials (RCTs) have become an essential part of the ever evolving state of knowledge upon which *Clinical Evidence* is based. Current treatment of prevalent conditions such as hypertension, hyperlipidaemia, acute myocardial infarction and heart failure would be unthinkable without the numerous and often very large trials that have been done. The physician who prescribes a beta-blocker, a statin, a thrombolytic agent or an angiotensin-converting enzyme inhibitor can be assured that these are beneficial because their "effectiveness has been demonstrated by clear evidence from RCTs".[2]

There is a large body of literature in textbooks and specialised journals on clinical trial methods, mostly statistical in nature. A useful overview may be found in Redmond and Colton.[3] What is perhaps insufficiently realised is that the statement '*we did a randomised, placebo controlled, double blind trial'* defines a procedural concept. Whether a "systematic tendency of any factors associated with the design, conduct, analysis and interpretation of the results … to make the estimate of treatment effect deviate from its true value"[4] was avoided can never be shown from the data. Even the most sophisticated statistical methods cannot compensate for flawed trial conduct.

In 1969, evidence from RCTs became mandatory for getting marketing approval from the US Food and Drug Administration (FDA).[5] This has meant that most clinical trials are executed today within a polygon of forces and interests.[6] While there is a tendency to accept results of trials at face value, in particular when published in a major peer-reviewed medical journal, problems have arisen that can often be attributed to inappropriate interference with trial conduct.[7] Trust in the results of a particular trial is therefore equivalent to trusting that its scientific integrity was ensured by managing trial conduct without undue interference.

The aim of this thesis is to show how clinical trial conduct can be managed while respecting the underlying scientific principles. Chapter 2 describes the main results of PICO (PImobendan in COngestive heart failure), a trial which investigated a positive inotropic agent in patients with heart failure using exercise toler-

ance as primary outcome. The results of this trial framed our thinking about how to analyse the balance between positive and untoward effects of treatment, thinking that is further elaborated in Chapter 3. This chapter also covers issues that relate to the use of so-called combined endpoints, a feature of many recent large trials focusing on clinical outcome. Several of the lessons learnt as described in Chapter 3 were implemented while designing the ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine) trial as described in Chapter 4. In Chapter 5 baseline data from the same trial are presented. At the time of writing, ACTION was still ongoing; results will be available in September 2004. In Chapter 6 we describe the database management system which was implemented to manage the ACTION study. In Chapter 7, the statistical analysis plan for this trial is reproduced. Finally, in a general discussion (Chapter 8) we focus on the trial management issues that arose during the conduct of the PICO and ACTION trials respectively.

Major medical journals will publish trial results only when the "data have been gathered and are presented in an objective and dispassionate manner".[8] It follows that scientific integrity cannot be an afterthought when the trial is finished. Scientific integrity must be ensured by appropriate trial conduct from the beginning. Otherwise, the results may never see the light of day.

REFERENCES

1. Cornfield J. Principles of research. *Am J Ment Defic* 1959;**64**:240-52
2. Godlee F *et. al*. (eds). Clinical Evidence. The international source of the best available evidence for effective healthcare. BMJ Publishing Group Ltd, Issue 10, London, 2003
3. Redmon CK, Colton T (eds). Biostatistics in clinical trials. John Wiley & Sons, Chisester, 2001
4. The European Agency for the Evaluation of Medicinal Products. Human Medicines Evaluation Unit. Note for guidance on statistical principles for clinical trials. ICH topic E9, CPMP/ICH/363/96, London, 1998 (www.emea.eu.int/pdfs/human/ich/036396en.pdf, accessed on February 4, 2004)
5. Pocock SJ. Clinical trials – a practical approach. John Wiley & Sons, Chisester, 1983
6. Horton R. The Dawn of McScience. The New York Review of Books. Volume 51, Number 4; March 11, 2004 (www.nybooks.com/articles/16954, accessed on February 25, 2004)
7. Horton R. The clinical trial: deceitful, disputable, unbelievable, unhelpful, and shameful – what next? Control Clin Trials 2001;22:593-604
8. Davidoff F, DeAngelis CD, Drazen JM et al. Sponsorship, authorship, and accountability. *N Engl J Med* 2001;**345**:825-6

Chapter 2

Effect of pimobendan on exercise capacity in patients with heart failure: main results from the PImobendan in COngestive heart failure (PICO) trial

ABSTRACT

Primary objective: To determine the effects of pimobendan 2·5 and 5 mg daily on exercise capacity in patients with chronic heart failure.
Design: A randomised, double blind, placebo controlled trial of the addition of pimobendan to conventional treatment with a minimum follow up of 24 weeks.
Setting: Outpatient cardiology clinics in six European countries.
Patients: 317 patients with stable symptomatic heart failure, objectively impaired exercise capacity and an ejection fraction of 45% or lower who were treated with at least an angiotensin converting enzyme inhibitor and a diuretic and who tolerated a test dose of pimobendan.
Results: Compared to placebo, both pimobendan 2·5 and 5 mg daily improved exercise duration (bicycle ergometry) by 6% (P = 0·03 and 0·05) after 24 weeks of treatment. At that time 63% of patients allocated to pimobendan and 59% of those allocated to placebo were alive and able to exercise to at least the same level as at entry (P = 0·5). No significant effects on oxygen consumption (assessed in a subgroup of patients) and on quality of life (assessed by questionnaire) were observed. Pimobendan was well tolerated. Proarrhythmic effects (24-hour electrocardiography) were not observed. In both pimobendan groups combined the hazard of death was 1·8 (95% confidence interval 0·9 to 3·5) times higher than in the placebo group.
Conclusions: Pimobendan improves exercise capacity in patients with chronic heart failure who are also on conventional treatment. The balance between benefit and risk of treatment with this compound remains to be established however.

Despite optimal treatment with diuretics and angiotensin converting enzyme inhibitors[1] many patients with chronic heart failure suffer from persistent symptoms, limited exercise tolerance and impaired quality of life. Also the prognosis continues to be poor. Hence, the treatment of this condition remains a therapeutic challenge.

Inotropic stimulation has met with little success. That digitalis glycosides improve exercise performance and quality of life is generally accepted but their effect upon mortality remains to be established. More recently phosphodiesterase inhibitors with both inotropic and vasodilating properties have been introduced.[2,3] Much has been expected of this class of substances[4] but long-term trials have shown that, although a moderate improvement in quality of life and exercise capacity can be achieved, mortality is increased.[5-7] A novel approach to inotropic stimulation is the direct sensitisation of cardiac myofilaments to cytosolic calcium. Pimobendan (UDCG 115 BS) combines this effect with myocardial cyclic adenosine monophosphate-dependent phosphodiesterase inhibition.[8] It is rapidly absorbed, the peak plasma level is reached after 1·5 hours and the effect lasts eight to 10 hours. Its main metabolite, UDCG 212, has similar pharmacodynamic properties as the parent compound. In patients with heart failure, oral or intravenous pimobendan produces lasting and dose dependent beneficial haemodymic effects. Myocardial energetics seem to be favourably influenced because the ratio between cardiac work and oxygen consumption remains unchanged or is even improved.[9] Systemic blood pressure is reduced slightly with higher doses of pimobendan. Although the phosphodiesterase inhibiting properties of pimobendan suggest arrhythmogenic potential, no important proarrhythmic effect has been observed in the clinical studies performed so far.

Several earlier trials in chronic heart failure patients showed beneficial effects when pimobendan was added to an optimal basic regimen with or without digitalis.[10-12] Also, a trial comparing pimobendan with the angiotensin converting enzyme inhibitor enalapril was performed.[13] To confirm these earlier findings, another placebo-controlled trial focusing on exercise capacity, oxygen consumption, and quality of life was undertaken with a longer (at least 24 weeks) treatment duration and in a larger group of patients than in previous trials.

PATIENTS AND METHODS

Patients

In 30 centres (see Appendix) patients were recruited if they had chronic moderate (New York Heart Association class II-III) heart failure and had given written informed consent. All patients were at least 18 years of age, had been clinically stable without important changes in background medication for at least 30 days, and had an ejection fraction of 45% or lower. At two maximal exercise tests at least two weeks apart, exercise times were within the age and sex-specific limits shown in table 1 and no more than one minute apart.

Table 1: Exercise protocol (maximal bicycle ergometry in sitting position) and limits of exercise time allowed at entry

	Age (years)						
Men	<65	65-69	70-74		≥75		
Women		<50	50-59	60-64	65-69	70-74	≥75
Workload (Watts):							
1st minute	0 W	0 W	0 W	0 W	0 W	0 W	0 W
2nd minute	30 W	20 W	20 W	10 W	10 W	10 W	10 W
Each next minute	+10 W	+10 W	+10 W	+10 W	+10 W	+10 W	+10 W
Limits (minutes:seconds):							
At least	3:00	3:00	3:00	3:00	3:00	3:00	3:00
Less than	10:00	10:00	10:00	10:00	9:00	8:00	7:00

Exclusion criteria were: stenotic, obstructive or infectious cardiac disease; exercise capacity limited by angina; on waiting list for transplantation, suspicion of digitalis toxicity; acute myocardial infarction, coronary revascularisation, or episode of syncope or cardiac arrest during the last three months; automatic cardiac defibrillator implanted; primary renal or hepatic disease; haemodynamically significant pulmonary embolism or severe pulmonary disease; any other still present life-threatening condition; previous participation in a trial with pimobendan; and anticipated problems with follow up or compliance. Women were excluded unless they had been sterilised or were at least two years post-menopausal. Baseline serum potassium had to be above 3·8 mmol/l, serum creatinine below 194 μmol/l and aspartate aminotransferase below 100 U/l.

The day before double blind medication was started, all patients were given a single test dose of pimobendan and were excluded if this dose produced significant symptoms or signs of intolerance.

Medication regimen

Background medication consisted of at least an angiotensin converting enzyme inhibitor and a diuretic. In addition, digitalis, nitrates and molsidomine were allowed. Other inotropic agents, phosphodiesterase inhibitors, ibopamine, β blockers, calcium antagonists and other vasodilators could not be given. The only antiarrhythmic agent allowed was amiodarone, which had to have been started at least three months before entry. In patients receiving amiodarone, the QT interval corrected for heart rate[14] (QTc) was monitored and the dose of this compound was reduced if QTc exceeded 480 ms.

Double blind medication consisted of either placebo, or pimobendan 2·5 or 5 mg daily, divided in two equal doses. Allocation was randomised and blocked by centre. Medication was also taken on the days of exercise testing. Investigators were instructed to halve the dose when serum creatinine rose to at least 2·5 mg/dl (220 μmol/l), or aspartate aminotransferase to at least 150 U/l; and to withdraw double blind medication if such elevation(s) persisted. Compliance was monitored by pill counts and was judged to have been good when at least 80% of the predicted number of double blind medication capsules had been used without interruption for more than 48 hours.

Trial design

The screening phase consisted of a minimum of two outpatient clinic visits at least two weeks apart and was followed by a 24-week efficacy phase on double blind medication consisting of clinic visits after 2, 4, 8, 12 and 24 weeks. Clinical status permitting, all planned follow up assessments (including exercise testing) were completed irrespective of deviations from the double blind medication regimen (intention-to-treat principle). After completion of the efficacy phase, patients who were still alive at that moment were followed for further clinical events until a common stopping date. Patients who consented continued to take double blind medication (extended follow up phase).

The primary outcome was exercise time measured at least twice during screening, and 4, 12 and 24 weeks after start of double blind medication. The same age and sex specific maximal bicycle ergometry protocol in the sitting position (table 1) was used in all centres. In a sub-study, gas exchange measurements were also performed.

Baseline ejection fraction was assessed by the investigator from a two dimensional echocardiogram, which was recorded on videotape and re-analysed later at the co-ordinating centre. If no analysable echocardiogram was available, a value obtained by another method was accepted. At each visit, the symptomatic state was assessed by the New York Heart Association classification.[15] Quality of life was evaluated by the Minnesota Living with Heart Failure questionnaire[16] administered twice during screening and after four, 12 and 24 weeks. Patient safety was assessed by monitoring vital signs (every visit), standard 12 lead electrocardiography preceding exercise testing, standard laboratory tests (screening, four, 12 and 24 weeks) and 24 hour electrocardiography (screening, four and 24 weeks). ECGs were analysed at the co-ordinating centre. Proarrhythmia was assessed by criteria described previously.[17]

While the trial was ongoing, patients were withdrawn when a clinically relevant violation of the selection criteria was detected at the co-ordinating centre. The decision to do so was taken by the steering committee before the medication code was broken. Cause of death was classified by the critical events committee, which had no access to the medication code. Sudden death was defined as witnessed death within one hour of onset of symptoms or unexpected, unwitnessed death.[18,19] This

committee also assessed whether missed exercise tests were due to the patient's cardiovascular condition.

Ethics

The trial was carried out in accordance with the Declaration of Helsinki,[20] had been approved by the institutional review board of each centre and was monitored according to European Union standards of Good Clinical Practice. For patient safety, episodes of ventricular tachycardia in excess of three consecutive beats observed in centrally analysed 24 hour electrocardiograms were reported to the investigator. Data on serious adverse events were regularly reviewed by an independent monitoring committee, which was in the possession of the treatment code from the start of the trial onwards.

Statistical methods

Sample size estimation was based on the results of an earlier trial with the same compound and the same primary outcome.[11] For exercise time and for the physical dimension of the Minnesota Living with Heart Failure questionnaire, 70 patients per group were required for a power of 90% at a two-sided significance level of 5% (based on charts of the power function for analysis of variance tests[21] and sums of squared treatment effects estimated from the earlier trial[11]). For peak oxygen consumption 45 patients per group were required. No interim analyses for efficacy were performed.

All patients who complied with the selection criteria and had started double blind medication were included in intention-to-treat analyses; no patient was excluded for protocol violations which had occurred during follow up. The primary pre-specified analysis of exercise time was limited to those patients who had at least the first follow up (four week) exercise test carried out and had shown good compliance (see medication regimen) up to the day of the test. If subsequent tests were not performed, whatever the reason, or were performed although compliance between tests had been poor, the last exercise time value obtained while compliance was good was carried forward. Tests of significance were performed by fixed-effects analysis of variance for repeated measurements.[22] All P values quoted are two-sided.

As a secondary analysis, changes in exercise time at 24 weeks were ranked and all patients assigned to placebo were compared by a standard non-parametric test with all patients assigned to pimobendan on an intention-to-treat basis. If the patient had died or if there had been a cardiovascular contraindication for exercise testing, the rank assigned was that below the patient with the lowest rank based on exercise time changes actually observed. If exercise time was missing for other reasons, the last available value was carried forward in calculating the rank.

Percentages were compared between groups by chi-square tests with one or two degrees of freedom for comparisons between two or three groups respectively. Hazards of all-cause mortality and of the combined event death or hospitalisation

for cardiovascular reasons were calculated by dividing the total number of events that had occurred by the number of person-years of follow up that patients had been at risk of the event considered. The time of first intake of double blind medication was regarded as the beginning of follow up, and the day the patient was last known to be alive, as the end. Medication groups were compared by hazard ratios and their 95% confidence intervals.[23] To assess whether the hazard ratio estimates for the combined endpoint death or hospitalisation for cardiovascular reasons were distorted by imbalance between treatment groups for clinical predictors of this event at baseline, the same hazard ratios were also estimated by Cox multivariate proportional hazard analysis.[24]

RESULTS

Patients, conduct and compliance

The trial was completed as planned. From March 1993 to June 1994, 333 patients completed the screening phase and had one test dose of pimobendan administered. None of these patients showed acute intolerance but two were subsequently excluded (one for hypokalaemia, one for worsening heart failure). Thus 331 patients were started on double blind medication. Before the medication code was broken 14 of these were excluded for violations of the selection criteria (four placebo, six pimobendan 2·5 mg/day, four pimobendan 5 mg/day; one excluded patient assigned to placebo died nine months after randomisation). Hence, 317 patients (108 allocated to placebo, 106 to pimobendan 2·5 mg/day, and 103 to 5 mg/day) are the basis of this report. All but one had been followed until November 1994 for clinical events. Selected entry characteristics are presented in table 2. Given the relatively small numbers in each treatment arm, randomisation resulted in well balanced groups.

Table 2: Entry characteristics*

		Pimobendan (daily dose)		
	Placebo	2·5 mg	5 mg	Total
Number of patients	108 (100%)	106 (100%)	103 (100%)	317 (100%)
Age (years)	65·1 (10·3)	64·8 (9·1)	66·6 (9·1)	65·5 (9·5)
Male	79%	78%	83%	80%
History:				
Ischaemic aetiology	72%	67%	67%	69%
Documented MI	65%	63%	57%	62%
Coronary bypass grafting	18%	26%	21%	22%

		Pimobendan (daily dose)		
	Placebo	2·5 mg	5 mg	Total
Treated for high blood pressure	23%	36%	20%	26%
Diabetes	14%	15%	20%	16%
Median (range) duration of HF (months)	34 (1–225)	35 (2–480)	35 (2–284)	35 (1–480)
Hospitalised >24 hours for HF	73%	66%	70%	70%
NYHA II / III	51% / 49%	56% / 44%	50% / 50%	52% / 48%
Basic regimen:				
ACE-inhibitor and diuretic	100%	100%	100%	100%
Digitalis glycosides	63%	55%	59%	59%
Nitrates	43%	44%	46%	44%
Molsidomine	7%	7%	5%	7%
Amiodarone	9%	11%	17%	12%
Anticoagulants	56%	46%	51%	51%
Observations:				
Regular sinus rhythm	78%	74%	73%	75%
Ejection fraction (%)†	28 (7)	26 (6)	28 (7)	27 (7)
End-diastolic volume (ml)‡	312 (104)	315 (95)	308 (113)	312 (104)
Exercise time (seconds)	435 (97)	417 (100)	416 (110)	423 (102)
Drop of SBP during exercise	3%	4%	7%	4%
Total MLHF score§	27·9 (18·4)	26·5 (16·8)	25·8 (19·1)	26·8 (18·1)
QTc interval >480 msec	14%	18%	18%	17%

* Data are given either as a percentage or as a mean (SD), unless indicated otherwise. ACE, angiotensin converting enzyme; HF, heart failure; MI, myocardial infarction; MLHF, Minnesota Living with Heart Failure questionnaire; NYHA, New York Heart Association classification; SBP, systolic blood pressure; QTc, QT-interval corrected for heart rate

† No entry value was available for one patient assigned to pimobendan 5 mg/day (local values were accepted when central analysis was not possible).

‡ No entry values were available for 11 patients assigned to placebo, 7 patients assigned to pimobendan 2·5 mg/day and 10 patients assigned to pimobendan 5 mg/day

§ No entry value was available for one patient assigned to pimobendan 2·5 mg/day.

Table 3: Exercise testing and medication compliance during efficacy phase

		Pimobendan (daily dose)		
	Placebo	2·5 mg	5 mg	Total
Number of patients	108	106	103	317
Exercise test done:				
4 weeks	103	99	98	300
12 weeks	97	92	88	277
24 weeks	94	86	88	268
Exercise test done and good compliance:				
4 weeks	101	97	94	292
4 and 12 weeks	93	82	82	257
4, 12 and 24 weeks	88	73	79	240
Exercise test not done due to death:				
4 weeks	2	3	3	8
12 weeks	5	5	10	20
24 weeks	6	13	11	30
Exercise testing contra-indicated:				
4 weeks	1	3	2	6
12 weeks	3	4	4	11
24 weeks	4	2	3	9
Exercise test not done for other reasons:				
4 weeks	2	1	0	3
12 weeks	3	5	1	9
24 weeks	4	5	1	10

Table 3 summarises the extent to which exercise testing was performed in relation to compliance with the double blind medication regimen. For the primary analysis of exercise time 92% of patients (292/317) were available. In 76% (240/317) all three exercise tests were performed and compliance was good throughout the efficacy phase. One or more exercise test results was carried forward in the primary analysis for 52 patients. In 10 patients there was a reason other than death or a contra-indication for not performing the planned 24-week exercise test. For these the last exercise test result was carried forward in the ranked secondary analysis of change in exercise capacity at 24 weeks (see statistical methods).

Based on the same criteria for inclusion as for the primary analysis of exercise time, 35 patients on placebo, 34 on pimobendan 2·5 mg/day and 34 on pimobendan 5 mg/day respectively were available for analysis of gas exchange measurements. Similarly, 103 patients on placebo, 100 on pimobendan 2·5 mg/day and 94 on pimobendan 5 mg/day were available for a per-protocol analysis of Minnesota Living with Heart Failure questionnaire scores.

Exercise capacity and peak oxygen consumption

The mean changes in exercise time (relative to the last baseline test) for the patients included in the primary analysis are shown in figure 1. After four, 12 and 24 weeks respectively, the mean treatment effects of pimobendan 2·5 mg/day relative to placebo were 13, 27 and 29 seconds (P = 0·03); and of pimobendan 5 mg/day 19, 17 and 28 seconds (P = 0·05).

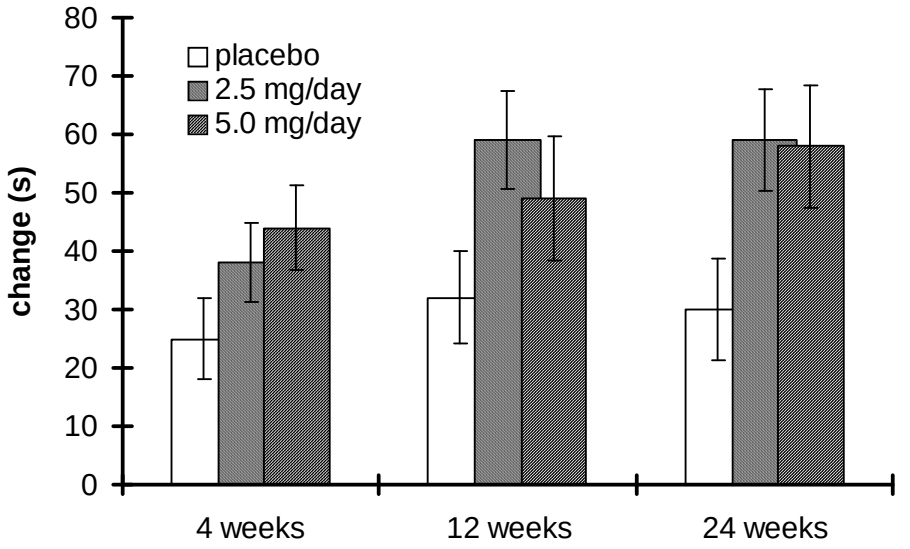


Figure 1: Mean (SEM) of change in exercise time after four, 12 and 24 weeks

For the 292 patients entered in the primary analysis (101 placebo, 97 pimobendan 2·5 mg/day and 94 5 mg/day; table 3), the bars show by treatment group the mean changes in exercise time relative to baseline) after four, 12 and 24 weeks. The last available value was carried forward when data for 12 or 24 weeks was missing. For the 101 patients on placebo in this analysis, the mean entry exercise time was 437 seconds and the mean changes were (standard errors between brackets): +25 (7) after four weeks, +32 (8) after 12 weeks and +30 (9) seconds after 24 weeks. In the pimobendan 2·5 mg/day group the mean changes were +38 (7), +59 (8) and +59 (9), and in the 5 mg/day group +44 (7), +49 (11) and +58 (10) seconds respectively (overall P = 0·06; 2·5 mg/day versus 5 mg/day P = 0·9; 2·5 mg/day versus placebo P = 0·03, 5 mg/day versus placebo P = 0·05; all pimobendan versus placebo P = 0·02).

Figure 2 shows the cumulative distributions of changes in exercise time at 24 weeks for the 108 patients assigned to placebo and the 209 patients assigned to pimobendan (both dose groups combined), based on intention-to-treat and taking into account exercise tests which were not possible because of death or a contraindication. Of the 209 patients assigned to pimobendan, 63% were able to exercise at 24-

weeks to at least the same level as at entry. Among the 108 patients assigned to placebo, this was the case in 59% (P = 0·5, non-parametric test).

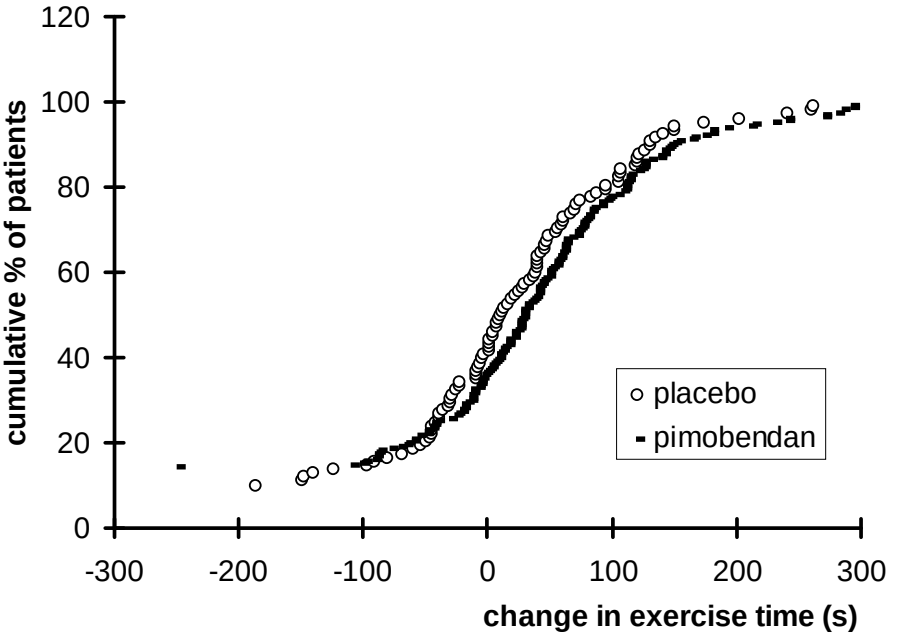


Figure 2: Cumulative distributions of changes in exercise time at 24 weeks based on intention-to-treat for placebo (108 patients), and for both pimobendan groups combined (209 patients)

In ten cases in the placebo group (6 + 4, table 3) and 29 in the combined pimobendan groups (13 + 11 + 2 + 3; table 3, P = 0·2), the 24-week exercise test was missing because of death or a cardiovascular contra-indication for exercise testing. Hence, for placebo the curve starts at the 10·2th (11/108) and for pimobendan at the 14·4th (30/209) percentile. Whenever the 24-week test was missing for other reasons (4 placebo, 5 + 1 pimobendan; table 3), the last available value was carried forward. The medians do not differ significantly (difference between medians = 18 seconds, P = 0·5) but in this trial more patients assigned to pimobendan showed improvement in exercise time than patients assigned to placebo, despite their higher incidence of death or contra-indications for exercise testing.

In the subgroup of 35 patients allocated to placebo who had gas exchange measurements the mean peak oxygen consumption was 13·1 ml/min.kg at entry; and 13·7, 13·3 and 13·9 ml/min.kg after four, 12 and 24 weeks respectively. At these time points, the mean effects of pimobendan 2·5 mg/day (relative to placebo, per-protocol analysis) on this variable were 0·2, 0·8 and 0·4 (P = 0·4) and of pimobendan 5 mg/day 0·6, 0·8 and 0·3 ml/min.kg respectively (P = 0·4).

New York Heart Association class and quality of life

Of those assigned to placebo, 4% (4/108) were in a better New York Heart Association class at least once during follow up than at entry and never worsened or died before the 24 week visit. The corresponding figure for those assigned to pimobendan was 10% (20/209; P = 0·06).

For the 103 patients on placebo analysed for Minnesota Living with Heart Failure questionnaire scores, the mean total score at entry was 27·7 units; and 25·0, 25·6 and 26·1 units after four, 12 and 24 weeks respectively. At these time points, the mean effects of pimobendan 2·5 mg/day (relative to placebo, per-protocol analysis) were –0·4, –1·6 and –1·2 (P = 0·5) units and of pimobendan 5 mg/day – 0·0, +0·7 and –1·0 units respectively (P = 0·9).

Clinical outcome

During the efficacy phase proarrhythmia based on 24-hour electrocardiography was observed in 20 patients assigned to placebo, 15 to pimobendan 2·5 mg/day and 14 to pimobendan 5 mg/day. When sudden cardiac death was considered as proarrhythmia also, these numbers became 24, 22 and 19 respectively (P = 0·5).

Data on the clinical outcome at the end of the efficacy phase (i.e. at 24 weeks) is given in table 4. Double blind medication was stopped or reduced significantly more often in patients allocated to pimobendan than to placebo (11 + 8 = 19 placebo, 21 + 13 = 34 pimobendan 2·5 mg/day, 16 + 13 = 29 pimobendan 5 mg/day; P = 0·04).

Data on mortality and hospitalisation for cardiovascular reasons during the entire trial (that is, efficacy and extended follow up phase combined) based on intention-to-treat is given in table 5. The mean follow up was 11 months and 44% of patients were followed for at least one year. In total, 47 patients died. While taking double blind medication, eight died on placebo, 14 on pimobendan 2.5 mg/day and 11 on pimobendan 5 mg/day. The causes of death are also given in table 5. More than half of all deaths were classified as sudden. Cardiovascular deaths caused by myocardial infarction (one in each treatment group), worsening heart failure (one placebo, two pimobendan 2·5 mg/day and five pimobendan 5 mg/day), multiple organ failure (one placebo), cerebrovascular accident (one pimobendan 2·5 mg/day) and peripheral vascular disease (one pimobendan 2·5 mg/day). The two non-cardiac deaths were caused by bladder cancer (placebo) and pneumonia (pimobendan 5 mg/day).

Although there was a trend towards a higher mortality hazard among patients allocated to pimobendan, neither comparison with placebo was statistically significant at the 5% level (each confidence interval includes one, table 5). In both pimobendan groups combined, the mortality was 1·8 times higher (0·9 to 3·5) than in the placebo group. Similarly, there was a non-significant trend towards higher hazards of the combined event death or first hospitalisation for cardiovascular reasons.

Table 4: Clinical events (excluding extended follow up)

	Number of patients with event*			Ranked clinical outcome at 24 weeks†		
Event:	Placebo (*N* = 108)	Pimobendan 2·5 mg/day (*N* = 106)	Pimobendan 5 mg/day (*N* = 103)	Placebo (*N* = 108)	Pimobendan 2·5 mg/day (*N* = 106)	Pimobendan 5 mg/day (*N* = 103)
Death (all causes)	6	13	11	6	13	11
Acute myocardial infarction	2	1	1	1	0	1
Hospitalisation for worsening heart failure‡	0	1	4	0	1	2
Hospitalisation for arrhythmia§	4	3	5	3	1	5
Hospitalisation for other cardio-vascular reasons	14	17	16	8	10	5
Medication for heart failure added or increased‖	29	25	22	17	12	10
Anti-arrhythmic agents added	6	7	6	1	3	0
Other cardiac medication added or increased¶	3	5	0	1	0	0

Double blind medication stopped by physician	11	21	16	5	5	2
Double blind medication reduced by physician	8	13	13	4	5	7
Hospitalisation for non-cardiovascular reasons**	12	13	18	3	1	4
None of the above	–	–	–	59	55	56

* Each patient with at least one of the events listed counted.

† Each patient counted once and assigned to the uppermost applicable event listed.

‡ In-hospital addition of intravenous inotropics (including digitalis).

§ In-hospital addition of an anti-arrhythmic agent (including digitalis).

|| Diuretics, ACE-inhibitors, intravenous inotropics, digitalis, nitrates or molsidomine added or increased at any time.

¶ Calcium-antagonists, beta-blockers or other vasodilators added or increased at any time.

** Except hospitalisation for procedures already planned before entry

Table 5: Mortality and mortality or first hospitalisation for cardiovascular reasons (intention-to-treat; including extended follow up)

	Placebo (*N* = 108)	Pimobendan 2·5 mg/day (*N* = 106)	HR 2·5/pl* (95% CI)	Pimobendan 5 mg/day (*N* = 103)	HR 5/pl* (95% CI)
Deaths (all causes)	11	20		16	
Sudden cardiac death†	7	15		7	
Other cardiovascular death†	3	5		6	
Non-cardiovascular death†	1	-		1	
Unknown cause	-	-		2	
Hazard‡ (per 100 person-years)	10·8	21·3	2·0 (0·9 - 4·1)	17·4	1·6 (0·7 - 3·4)
No. of deaths or 1st hospitalisation	27	40		33	
Hazard‡ (per 100 person-years)	29·5	48·5	1·6 (1·0 - 2·7)	41·2	1·4 (0·8 - 2·3)

* Hazard ratios (HR) comparing patients assigned to pimobendan 2·5 mg/day and to pimobendan 5 mg/day respectively with patients assigned to placebo; with 95% confidence intervals (CI).

† As assessed by the Critical Events Committee

‡ Number of events divided by person-time of follow up until event or until end of extended follow up

The adjusted estimates from the Cox regression model of the hazard ratios for death or hospitalisation for cardiovascular reasons were 1·5 comparing pimobendan 2·5 mg/day with placebo (95% confidence interval 0·9 to 2·5), and 1·2 (0·7 to 2·1) comparing pimobendan 5 mg/day with placebo. Conditionally independent and significant ($P < 0·05$) covariates associated with a reduced risk in the Cox model were (hazard ratios between brackets): change of systolic blood pressure during exercise (0·983/mm Hg rise), resting systolic blood pressure (0·980/mm Hg), exercise time (0·998/second), and previous coronary bypass grafting (0·5). Covariates associated with an increased risk were previous hospitalisation for heart failure (2·6), history of symptomatic arrhythmia (1·6) and history of high blood pressure treated with drugs (1·7). The adjusted estimates are similar to those given in table 5. Conditional on the covariates mentioned, age, sex, NYHA class, ejection fraction and use of digitalis were not associated with the risk.

DISCUSSION

The main finding of this trial confirms earlier findings that pimobendan improves exercise capacity.[10-12] Relative to placebo, the magnitude of this effect seemed to increase until 12 weeks of treatment and then to persist at the same level at 24 weeks (figure 1). This is not only the largest placebo-controlled trial to date but also the first one with a treatment period of this duration. This sustained effect has therefore not been reported before.

The largest earlier placebo-controlled trial studied 198 patients in NYHA class III who were treated for 12 weeks with either placebo, or 2·5, or 5, or 10 mg/day of pimobendan.[11] Mean exercise duration increases (rounded to whole seconds) of 30, 68, 122 and 81 respectively were observed. Adding these mean increases to the mean entry exercise time in the placebo group (536 seconds) yields mean follow up exercise durations of 566, 604, 658, and 617 seconds respectively for placebo or 2·5, or 5 or 10 mg/day of pimobendan. Hence, in that trial the effect of treatment expressed as the percentage difference compared placebo was [(604 - 566) x100]/566 or 7% for treatment with 2·5 mg/day of pimobendan. Calculated in the same way, the treatment effects for 5 and for 10 mg/day of pimobendan were 16% and 9% respectively. In another smaller earlier trial, 10 treatment effects calculated in similar manner of 16% and 13% for 5 and 10 mg/day of pimobendan respectively were observed.

Based on the data given in table 2 and figure 1 and by the same method of calculation, treatment effects at 24 weeks in this trial were 6% for 2·5 and also 6% for 5 mg/day of pimobendan. Thus the magnitude of effect of 2·5 mg/day of pimobendan in this trial was similar to that in the largest earlier trial[11] but the effect of 5 mg/day was considerably less than observed before.[10,11] Based on the dose-response relations observed in these earlier trials, the 10 mg/day dose was not studied in this trial. Contrary to the largest earlier trial,[11] no difference was observed between the 2·5 and 5 mg/day dosages with respect to their effect on exercise duration. One can only speculate why the effect on exercise duration of the 5 mg/day

dose in this trial was less marked than observed earlier. In the largest earlier trial,[11] all patients were in NYHA class III. In this trial that was the case in 48% (table 2). We found no relation between the effect of pimobendan and NYHA class at entry. Hence, it seems unlikely that this can explain the difference between this and the earlier result. Similarly, it is unlikely that the difference in angiotensin converting enzyme inhibitor use (80% in the largest earlier trial, all patients in the present one) or digitalis use (88% in the largest earlier trial, 59% in the present one) can explain the difference. Whether there is a relationship between the use of angiotensin converting enzyme inhibitors and the effect of pimobendan is not known and in our data there was no relation between the effect of pimobendan and use of digitalis. A factor may have been that patients in this trial were older than in the earlier trial (65·5 mean age in this one, 61 in the largest earlier trial) and that ischaemic heart disease was a more frequent aetiology (69% in this trial and 44% in the largest earlier trial). This, and the different exercise methodology used,[25] may have influenced the results. Alternatively, the larger efficacy of pimobendan 5 mg/day relative to 2·5 and 10 mg/day in the largest earlier trial[11] may to a certain extent have been a chance finding. An argument in favour of this that the 10 mg dose has produced the largest haemodynamic improvement.[10] In any case, this trial together with earlier trials[10,11] suggest that the therapeutic range of pimobendan is 2·5 to 5 mg/day.

This trial did not reproduce the statistically significant effects of pimobendan peak oxygen consumption[10,11] and on Minnesota Living with Heart Failure questionnaire[11] scores seen in earlier trials. Again, one can only speculate about the reasons. Peak oxygen consumption was measured only in a sub-group of patients and the power may have been insufficient. The Minnesota Living with Heart Failure questionnaire may be less sensitive to changes in patient well being and/or less reproducible when this instrument is used in different cultural settings than the one for which it was originally developed and validated.

This trial followed a classical pattern for exercise capacity trials in heart failure. Several additional design features, however, allowed us to clarify further the efficacy and safety of adding pimobendan to the prevailing basic heart failure regimen. As in the largest earlier trial,[11] we randomised only those patients who tolerated a test dose of pimobendan. No patient was excluded because of acute intolerance. In fact, the present data show that pimobendan in the dosages studied is well tolerated. The number of patients in whom treatment had to be discontinued was small (table 4). No clinically relevant signs of proarrhythmia were observed. The number of patients in whom proarrhythmia was observed by 24-hour electrocardiography based on accepted criteria[17] was highest in the placebo group even when sudden cardiac death in the absence of proarrhythmia based on 24-hour electrocardiography was also regarded as proarrhythmia.

When the clinical situation permitted, follow up exercise testing was performed in patients in whom pimobendan had been discontinued. This allowed for the ranking of patients based on exercise capacity at 24 weeks and on intervening clinical events as shown in figure 2. We believe this to be an important additional

evaluation of the efficacy of chronic heart failure treatment. Effects on clinical outcome may either dilute or enhance effects on exercise capacity.[26] In the present case, the trend towards a higher mortality in patients started on pimobendan raises the question whether the benefit related to the positive effect of pimobendan on exercise capacity in the survivors is negated by the higher mortality risk associated with this treatment. The ranking in figure 2 shows that patients started on pimobendan have a higher chance of surviving 24 weeks *and* improving their exercise capacity than patients started on placebo. The difference was, however, not statistically significant. The major limitation of this trial is therefore that it does not allow a definitive conclusion about the benefit/risk ratio of treatment with pimobendan and it is this benefit/risk ratio which would merit further study in a larger trial.

To collect further data on the safety of pimobendan, all patients were followed until a common stopping date after they completed the 24-week efficacy phase, and those who consented were kept on medication. This design feature optimises the number of patient-years of follow up which can be accumulated within a given time. Hence it was possible to evaluate mortality and the combined rate of death or first hospitalisation for cardiovascular reasons over a mean follow up of 11 months based on intention-to-treat. Even though the differences were not significant, the hazards of both events were higher in the pimobendan than in the placebo groups (table 5). This was largely due to an increased incidence of sudden cardiac death, in particular in the pimobendan 2·5 mg/day group. The results of Cox multivariate proportional hazard analysis showed that these differences are unlikely to be related to imbalance between treatment groups for important predictors of clinical outcome.

Because of its different mechanism of action,[9] the trend towards a higher mortality in patients treated with pimobendan in this trial is unexpected despite earlier findings with predominantly cyclic adenosine monophosphate dependent phosphodiesterase inhibitors such as milrinone and enoximone.[6,7] Our findings may be attributable to chance because the trend observed was not statistically significant. Another potential explanation is a hitherto unknown interaction between digitalis and a positive inotropic compound such as pimobendan. The excess mortality in the pimobendan groups in this trial occurred predominantly in patients who were also on digitalis at entry. Thirty five of the 47 deaths observed (table 5) occurred in patients on digitalis at entry, divided as follows over the treatment groups: eight placebo, 15 pimobendan 2·5 mg/day and 12 pimobendan 5 mg/day. The corresponding numbers of deaths among patients not on digitalis at entry were three, five and four respectively. The number of events in each subgroup was small, this sub-group analysis had not been specified in advance and the apparent interaction between the effect of pimobendan on mortality and the use of digitalis at entry was not statistically significant. Nevertheless, our data raise the question whether the benefit/risk ratio of treating heart failure patients with positive inotropic agents depends on co-treatment with digitalis. Theoretically, calcium sensitisation enhances the effects of digitalis, which acts by increasing the ionised calcium concentration

within the myocardial cell.[27] Possibly, this leads to overstimulation. Rather than being dependent on co-treatment with digitalis, the benefit/risk ratio may also depend on the prognosis in general. In this trial, patients on digitalis were sicker (as evidenced by lower ejection fractions) and had a worse prognosis than patients not on digitalis. In any case, these potential sources of effect modification of pimobendan, and possibly of other positive inotropic agents, warrant further investigation. The extent to which such questions can be answered in placebo controlled trials will depend to a large extent on the outcome of the currently ongoing large-scale digitalis trial.

In conclusion, as regards exercise capacity this trial confirmed the efficacy of the addition of pimobendan to the basic regimen of patients with chronic heart failure. The effect of pimobendan as monotherapy has not been studied in controlled trials thus far. Pimobendan was well-tolerated and clinically relevant proarrhythmia was not observed. There was a trend towards improved clinical condition in patients treated with pimobendan, but also towards a higher mortality. Hence, the balance between benefit and risk, which may depend also on concomitant use of digitalis, remains to be established.

APPENDIX

The Pimobendan in Congestive Heart Failure (PICO) Investigators consisted of:

CLINICAL INVESTIGATORS

Belgium: G Heyndrickx (O-L Vrouweziekenhuis, Aalst), M Renard (ULB Hôpital Erasme, Brussels). *Germany:* S Beil, W Doering (Städt. Krankenhaus München-Schwabing); E. Kauder (Kreiskrankenhaus Tuttlingen), H Klepzig, C Schacherer (Klinikum der Johann Wolfgang Goethe-Universität, Frankfurt), M Haass, C Krüger (Abt Innere Medizin III, Universitätsklinikum Heidelberg), P Limbourg (Stadtkrankenhaus Worms), E Liomin (Städt. Krankenhaus Singen), B. Maisch (Klinikum der Philipps-Universität, Marburg), M Schartl (Universitätsklinikum Rudolf Virchow, Berlin), H Simon (Städt. Krankenhaus Düren). *The Netherlands* (*member of Working Group on Cardiovascular Research): PJLM Bernink* (Martini Ziekenhuis, Groningen), PhW Fels* (Drechtsteden Ziekenhuis, Dordrecht), HJM Dohmen* (Groot Ziekengasthuis, Den Bosch), PHJM Dunselman*, MRP Baselier (Ignatius Ziekenhuis, Breda), NJ Holwerda* (St. Elisabeth Ziekenhuis, Tilburg), GJ Laarman* (O-L Vrouwe Gasthuis, Amsterdam), CM Leenders (Havenziekenhuis, Rotterdam), K van Leeuwen* (Ziekenhuis Overvecht, Utrecht), HR Michels* (Catharina Ziekenhuis, Eindhoven), M van der Ent, WJ Remme (Sticares Cardiovascular Research Foundation, Rotterdam); AJAM Withagen* (R de Graaf Gasthuis, Delft). *Norway:* T Gundersen (Aust-Agder Central Hospital, Arendal), JE Otterstad, G Frøland (Vestfold Sentralsykehus, Toensberg), TR Pedersen (Aker Sykehus, Oslo). *Portugal:* JRG Ferreira, E Prazeres de Sá (Hospital de Santa Maria, Lisbon), R Seabra-Gomes, V Gil (Hospital Santa Cruz, Carnaxide). *Sweden:* N-J Abdon (Uddevalla Hospital, Uddevalla), J Jonsson (Varberg Hospital, Varberg).

STEERING COMMITTEE

H Just, chairman (Freiburg, Germany). ÅC Hjalmarsson (Gothenburg, Sweden), WJ Remme (Rotterdam, The Netherlands).

DATA MONITORING AND ETHICAL REVIEW COMMITTEE

SJ Pocock (London School of Hygiene and Tropical Medicine, UK), T Meinertz (Universitätskrankenhaus Eppendorf, Hamburg, Germany).

CO-ORDINATING CENTRE, CORE LABORATORY FOR ECHOCARDIOGRAPHY AND 24-HOUR ELECTROCARDIOGRAPHY, AND ON-SITE MONITORING

Sociéte pour la Recherche Cardiologique (SOCAR) SA, Givrins, Switzerland (FJ van Dalen, J-M Dumont, P Jonkers, BA Kirwan, D La Framboise, L Le Roux, J Lubsen; M-J Roulin, M St John Sutton; D Sekarski, A Sublet, X-F Suckow).

SPONSOR

Boehringer Ingelheim Deutschland GmbH (M Haehl; J Heinrich-Nols)

REPORT PREPARED BY:

J Lubsen, H Just, ÅC Hjalmarsson, D La Framboise, WJ Remme, J Heinrich-Nols, J-M Dumont, P Seed (London School of Hygiene and Tropical Medicine, UK), X-F Suckow.

REFERENCES

1. Packer M. Treatment of chronic heart failure. Lancet 1992;340:92–5.
2. Kulka PJ, Tryba M. Inotropic support of the critically ill patient. A review of the agents. *Drugs* 1993;**45**:654–67.
3. Scholz H, Meyer W. Phosphodiesterase-inhibiting properties of newer inotropic agents. *Circulation* 1986;**73**:III99–108.
4. Colucci WS, Wright RF, Braunwald E. New positive inotropic agents in the treatment of congestive heart failure. *N Engl J Med* 1986;**314**:349–58.
5. The xamoterol in severe heart failure study group. Xamoterol in severe heart failure. *Lancet* 1990; **336**:1–6 [Erratum, *Lancet* 1990;**336**:698].
6. Uretsky BF, Jessup M, Konstam MA , Dec GW, Leier CV, Benotti J, *et al*. Multicenter trial of oral enoximone in patients with moderate to moderately severe congestive heart failure. Lack of benefit compared with placebo. *Circulation* 1990;**82**:774–80.
7. Packer M, Carver JR, Rodeheffer RJ, Ivanhoe RJ, DiBianco R, Zeldis SM, *et al* for the PROMISE Study Research Group. Effect of oral milrinone on mortality in severe chronic heart failure. *N Engl J Med* 1991;**325**:1468–75.
8. Hagemeijer F. Calcium sensitization with pimobendan: pharmacology, haemodynamic improvement, and sudden death in patients with chronic congestive heart failure. *Eur Heart J* 1993;**14**: 551–66.
9. Hasenfuss G, Holubarsch C, Heiss W, Just H. Influence of UDCG-115 on hemodynamics and myocardial energetics in patients with idiopathic dilated cardiomyopathy. *Am Heart J* 1989; **118**: 512–9.
10. Katz SD, Kubo SH, Jessup M, Brozena S, Troha JM, Wahl J, *et al*. A multicenter, randomized, double-blind, placebo-controlled trial of pimobendan, a new cardiotonic and vasodilator agent, in patients with severe congestive heart failure. *Am Heart J* 1992;**123**:95–103.
11. Kubo SH, Gollub S, Bourge R, Rahko P, Cobb F, Jessup M, *et al* for the Pimobendan Multicenter Research Group. Beneficial effects of pimobendan on exercise tolerance and quality of life in patients with heart failure. Results of a multicenter trial. *Circulation* 1992;**85**:942–9.

12. Sasayama S, Asanoi H, Kihara Y, Yokawa S, Terada Y, Yoshida S, *et al*. Clinical effects of long-term administration of pimobendan in patients with moderate congestive heart failure. *Heart Vessels* 1994;**9**:113–20.
13. Remme WJ, Krayenbühl HP, Baumann G, Frick MH, Haehl M, Nehmiz G, *et al* for the Pimobendan-Enalapril Study Group. Long-term efficacy and safety of pimobendan in moderate heart failure. A double-blind parallel 6-month comparison with enalapril. *Eur Heart J* 1994;**15**:947–56.
14. Bazett HC. An analysis of the time-relations of electrocardiograms. *Heart* 1920;**7**:353–8.
15. The Criteria Committee of the New York Heart Association. *Nomenclature and criteria for diagnosis of diseases of the heart and great vessels.* 9th ed. Boston: Little, Brown & Co; 1994:253–6.
16. Rector TS, Kubo SH, Cohn JN. Patient's self-assessment of their congestive heart failure: Part 2. Content, reliability and validity of a new measure, the Minnesota Living with Heart Failure Questionnaire. *Heart Failure* 1987;**3**:198–209.
17. Morganroth J, Borland M, Chao G. Application of a frequency definition of ventricular proarrhythmia. *Am J Cardiol* 1987;**59**: 97–9.
18. Myerburg RJ, Castellanos A. Cardiac arrest and sudden cardiac death. In: Braunwald E, ed. *Heart disease. A textbook of cardiovascular medicine.* 4th ed. Philadelphia: WB Saunders Company, 1992:756–89.
19. Cleland JGF, Erhardt L, Hall AS, Winter C, Ball SG on behalf of the AIRE Study Investigators. Validation of primary and secondary outcomes and classification of mode of death among patients with clinical evidence of heart failure after a myocardial infarction: a report form the Acute Infarction Ramipril Efficacy (AIRE) Study Investigators. *J Cardiovasc Pharmacol* 1993; **22** (Suppl 9):S22–7.
20. Declaration of Helsinki. The World Medical Assembly, Inc. Ferney-Voltaire, France, September 1989.
21. Pearson ES, Hartley HO. Charts of the power function for analysis of variance tests derived from the non-central F-distribution. *Biometrika* 1951;**38**:112–30.
22. Armitage P, Berry G. *Statistical methods in medical research*. 3rd ed. London: Blackwell Sci Publ, 1994:207–32.
23. Rothman KJ. *Modern epidemiology.* Boston: Little, Brown and Company, 1986:170–1.
24. Armitage P, Berry G. *Statistical methods in medical research*. 3rd ed. London: Blackwell Sci Publ, 1994:484–7.
25. Narang R, Swedberg K, Cleland JGF. What is the ideal study design for evaluation of treatment for heart failure? Insights from trials assessing the effect of ACE inhibitors on exercise capacity. *Eur Heart J* 1996;**17**:120–34.
26. Lubsen J. Exercise testing as outcome in congestive heart failure trials. *Drugs* 1994;**7**(Suppl 4): 25–30.

27. Feldman AM. Classification of positive inotropic agents. *J Am Coll Cardiol* 1993;**22**:1223–7.

Chapter 3

Combined Endpoints: Can we use them?

SUMMARY

Analysing specific non-fatal events in isolation may lead to spurious conclusions about efficacy unless the events considered are combined with all-cause mortality. The use of combined endpoints has therefore become widespread, at least in cardiovascular disease trials. Combining all-cause mortality with selected non-fatal events is useful because event-free survival, an important criterion in therapy evaluation, is addressed in this manner. In many clinical trials symptoms, signs or paraclinical measures (e.g., blood pressure, exercise duration, quality of life scores) are used as endpoints. If the patient died before the endpoint was measured, or it was otherwise not possible to perform follow up assessments as planned, the effect of treatment on these endpoints may be distorted if the patients concerned are ignored in the analysis. Examples are given of how distortion can be avoided by including all patients randomised in an analysis that uses a ranked combined endpoint based both on clinical events and on paraclinical measures. A distinction is made between a pseudo intention-to-treat analysis that disregards study medication status at the time of endpoint assessment, but is confined to patients with data; and a true intention-to-treat analysis that takes into account all patients randomised based on a ranked combined endpoint.

In clinical trials there are usually multiple outcomes that can be observed during follow up. To avoid multiplicity it is customary to focus on one specific outcome and to design the trial so that an efficacy question with respect to one pre-specified primary endpoint can be answered.

Focusing on one particular outcome has several limitations. A disease such as atherosclerosis may result in different disease manifestations that all have the same underlying cause. A treatment that affects atherosclerosis can therefore affect several different disease manifestations at the same time even when only patients presenting with one type of atherosclerotic disease are considered. Recently cardiovascular disease trials have combined several different causally related clinical events into one combined primary endpoint. Multiplicity is avoided in this manner. The power of the trial to detect a treatment effect may be enhanced and the overall clinical effect of certain treatments may be better understood when the total burden of a disease is evaluated by a combined endpoint.

Another limitation of focusing on one particular outcome occurs when the trial has an endpoint that pre-supposes that follow up assessments are complete after study medication has been taken for the specified period of time. Blood pressure in patients with hypertension, or exercise duration in patients with heart failure, are relevant examples. In trials focusing on such endpoints it is unavoidable that some planned follow up assessments will not be performed for reasons that relate to the clinical course of the patient following start of study medication. Confining the analysis to patients for whom follow up assessments were performed ignores the clinical outcome in those patients who were started on study medication but for whom follow up assessments were incomplete. It is not always appreciated that this may distort the comparison for the endpoint considered. Similarly, it is not always appreciated that distortion may occur when a specific clinical event, such as hospitalisation, is analysed while ignoring circumstances (such as death) that preclude the occurrence of the event considered.

Outcome information used as endpoints in clinical trials may be ranked hierarchically (Figure 1).[1] Distortion can occur when the analysis for an endpoint other than all-cause mortality (level 1 in Figure 1) ignores information from higher levels. Whether such distortion has occurred can be assessed by using endpoints that combine information from several levels, as will be shown in the examples that follow. Their purpose is also to show that combined endpoints often allow for a better understanding of treatment effects.

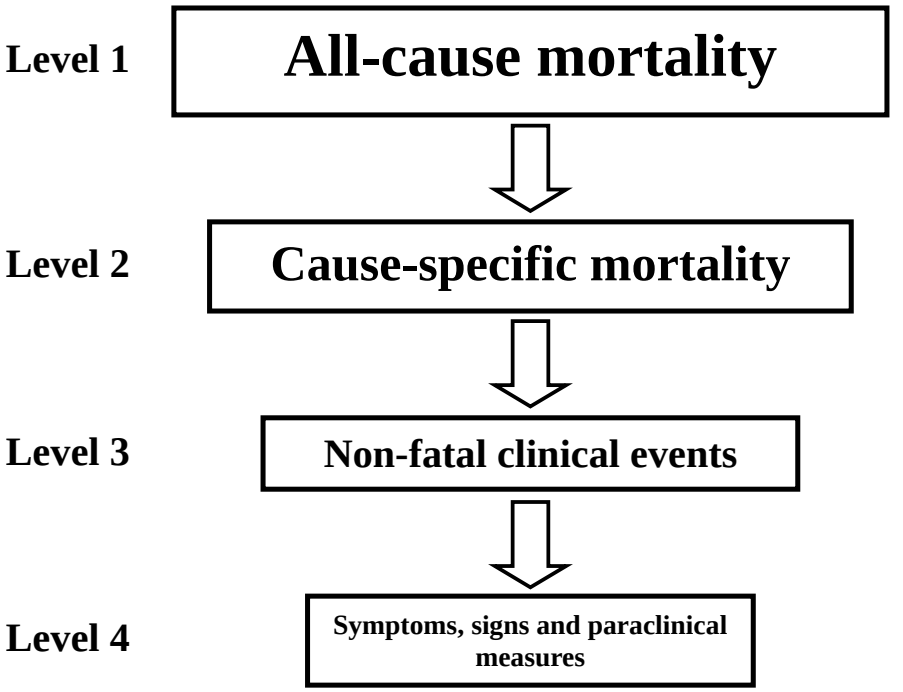


Figure 1: Hierarchical levels of outcome information to be considered in the analysis of clinical trial data

Level 4 information (e.g. blood pressure, symptoms) can only be obtained in patients who are still alive and must be considered in the context of non-fatal clinical events that may have affected the information. Non-fatal clinical events can only occur in patients who have survived up to the time of the event. Specific causes of death can only occur in patients who didn't die from another cause.

COMBINING MULTIPLE CLINICAL EVENTS

We first consider how combining clinical events from levels 1 – 3 (Figure 1) can help to understand what is achieved clinically by a certain treatment. Suppose that hospitalisation is the event of interest. Hospitalisation is a clinical event that belongs to level 3. Death – level 1 – reduces the risk of subsequent hospitalisation to zero. Hence, death and hospitalisation cannot be considered as independent entities. It would therefore be appropriate to report trial data on death and hospitalisation as shown in Table 1. Note that in this table all possible outcomes are considered and that they are displayed in mutually exclusive categories. It would be appropriate to consider only hospitalisation for certain reasons, such as aggravation of a specific underlying condition common to all patients at the time of randomisation. All-cause mortality, however, must be included to ensure that all patients en-

tered in the trial can be assigned to one of the four outcome categories considered in Table 1.

Table 1: Structure of data on death and hospitalisation (hypothetical data)

	Treatment	
	Active	Control
Died but never hospitalised during follow up (%)	15	5
Hospitalised and died during follow up (%)	5	15
Hospitalised, alive at the end of follow up (%)	20	20
None of the above (%)	60	60

Suppose that the percentages in Table 1 represent the results of a trial comparing active treatment with control. Total mortality is the same for both groups: 20%. The percentage of patients without either of the two events considered – the outcome category 'none of the above' in Table 1 – is also identical: 60%. The only effect of active treatment is a shift from hospitalisation and death (5% for active, 15% for control) to death without hospitalisation (15% for active, 5% for control). Because of this, the number of patients that were hospitalised is affected too (25% for active, 35% for control).

For the imaginary data in Table 1 it would be inappropriate to report only the total mortality and the number of patients that were hospitalised at any time. This would suggest that, without affecting mortality, active treatment reduced the risk of hospitalisation. The information that the percentage of patients without any of the two events considered was the same is suppressed. That this is not necessarily appreciated in published reports is shown in the following example.

Example 1: the dofetilide trial

This trial compared dofetilide – a drug that stabilises heart rhythm – with placebo in patients with heart failure.[2] In total, 762 patients were assigned to dofetilide and 756 to placebo. The primary endpoint was death from any cause. Hospitalisation for worsening heart failure was one of several secondary endpoints. The report contains a figure (reproduced here as Figure 2) showing that the probability of survival over time is essentially the same for both treatment arms, but that the probability of freedom from hospitalisation for worsening heart failure is favourably affected by dofetilide. The survival curves address the first and second mutually exclusive outcomes listed in Table 1 combined ('died but never hospitalised', and 'died and hospitalised'). The freedom from hospitalisation curves on the other hand address the second and third mutually exclusive outcomes listed in Table 1 combined ('died and hospitalised', and 'hospitalised, alive at the end of follow up'). Because of the resulting overlap, it is not possible to draw any conclusions based on the curves shown in Figure 2 about each of the four mutually exclusive out-

comes. In the report no data are given on death without hospitalisation (the 'died but never hospitalised'), or on survival without hospitalisation (the 'none of the above' outcome category).

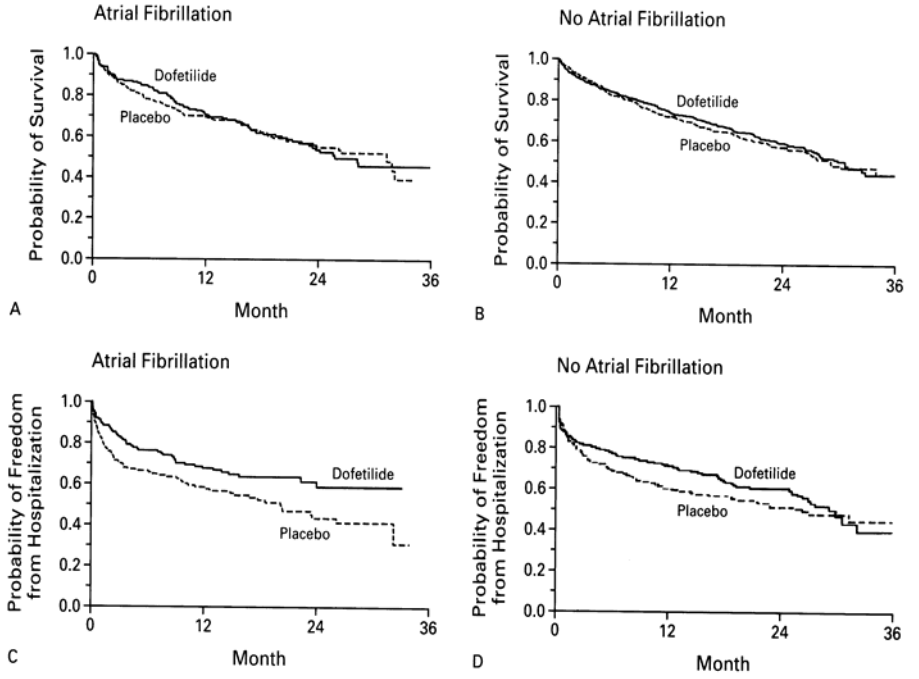


Figure 2: Kaplan-Meier estimates of the probability of survival and of freedom from hospitalisation for worsening congestive heart failure, according to the presence or absence of atrial fibrillation at base line

(A) Probability of survival among patients with atrial fibrillation at baseline (hazard ratio for the dofetilide group, 1·01; 95 percent confidence interval, 0·75 to 1·36). (B) probability of survival among patients without atrial fibrillation at baseline (hazard ratio, 0·94; 95 percent confidence interval, 0·78 to 1·13). (C) Probability of freedom from hospitalisation for worsening congestive heart failure among patients with atrial fibrillation at baseline (hazard ratio, 0·64; 95 percent confidence interval, 0·46 to 0·91). (D) Probability of freedom from hospitalisation for worsening congestive heart failure among patients without atrial fibrillation at baseline (hazard ratio, 0·80; 95 percent confidence interval, 0·65 to 0·98). Reproduced by permission of Massachusetts Medical Society from reference [2].

The abstract states that dofetilide was effective in reducing the risk of hospitalisation for worsening heart failure, but had no effect on mortality.[2] That there was no effect on mortality, and that there were fewer patients in the dofetilide than in the placebo arm who were hospitalised for worsening heart failure, is clear from Figure 2. Whether this is relevant is another matter. Many patients and doctors would consider it clinically useful to prescribe a drug that, while not prolonging life per se,

keeps the patient out of hospital; a drug in other words that prolongs hospitalisation-free survival. Few would however take a drug if its only effect were to be that death occurs outside hospital rather than after admission. The dofetilide report[2] does not address hospitalisation-free survival. Hence, the clinical relevance of the conclusion in the abstract is unclear.

Similarly, it makes no sense to consider only cause-specific mortality. A patient who died from a cause other than the one considered and who was not hospitalised would not be counted. As is shown in the next example, this causes problems of interpretation that affect many trial reports when several different clinical events are combined into one primary endpoint.

Example 2: the HOPE trial

In the HOPE trial, 4645 patients with risk factors for coronary disease were randomised to the ACE-inhibitor ramipril and 4652 to placebo.[3] The mean follow up was five years and the combined primary endpoint was cardiovascular death, or myocardial infarction (MI), or stroke. The numbers of patients with events are given in Table 2 in the same manner as listed in Table 3 in the original report[3].

Table 2: Incidence of the primary outcome and of deaths from any cause in the HOPE trial

	Assigned treatment	
	Ramipril *N* = 4645	Placebo *N* = 4652
MI, stroke or CV death (primary endpoint)	651	826
Cardiovascular (CV) death	282	377
Myocardial Infarction (MI)	459	570
Stroke	156	226
Non-CV death	200	192
All deaths	482	569

Reproduced by permission of Massachusetts Medical Society from Table 3 in the original report.[3]

There were 651 and 826 patients in the ramipril and placebo arms respectively who either died from a cardiovascular cause, or had MI, or stroke. The number of patients who died from a cardiovascular cause, had MI or stroke are also given in the report and in Table 2. To fully understand the manner in which the results are presented, it is useful to add up the number of patients listed as cardiovascular death, MI and stroke respectively. For the ramipril group this yields 282 + 459 + 156 = 897. The reason that this exceeds the 651 patients who reached the combined primary endpoint in this arm is that a patient who had an MI, then a stroke and then died from a cardiovascular cause is counted only once among the 651 patients who

reached the combined primary endpoint. The same patient is also counted once in each of the component outcome categories of the primary endpoint. The same patient is counted three times, however, when the counts of outcome events are summed.

Just as the freedom from hospitalisation curves in Figure 2 do not represent hospitalisation-free survival, the Kaplan-Meier curves for the primary endpoint given in the HOPE report do not represent MI- and stroke-free survival. In the report and in Table 2, the numbers of non-cardiovascular deaths are also given. Nonetheless, it is not possible to derive the outcome category 'none of the above' from the data given. Adding the 200 non-cardiovascular deaths to the 651 primary endpoints in the ramipril group does not yield the number of patients with any of the events considered because a patient who had an MI followed by non-cardiovascular death is counted twice in the total of 200 + 651 = 851 events. Thus, it is not possible to determine from the HOPE report how many patients were alive at the end of follow up and also free of MI and stroke.

COMBINING A CONTINUOUS OUTCOME WITH CLINICAL EVENTS

In many trials an endpoint belonging to level 4 (Figure 1) is used as criterion for efficacy. Clinical events belonging to levels 1 – 3 will by nature interfere with the assessment of an endpoint belonging to level 4. In analysing the results for such an endpoint this is often ignored. Important information may be suppressed in this manner, as is shown in the following example.

Example 3: The PICO trial

In the PICO trial, pimobendan (a drug that improves the heart's pump function) was compared with placebo in patients with heart failure.[4] Change from baseline in exercise duration was the primary criterion for efficacy. This is an endpoint that belongs to level 4. As regards its design, PICO is a standard 'surrogate endpoint' trial for the indication concerned. During a stabilisation phase, patients were selected based on repeated exercise tests to assess baseline exercise duration and clinical stability. Eligible patients were randomised to one of two different dosages of pimobendan, or placebo. Follow up exercise tests were done after 4, 12 and 24 weeks of treatment. Unless there was a clinical contraindication for exercise testing, follow up tests were also performed in patients for whom study medication had been stopped before the end of the trial, a design improvement that is not necessarily standard practice in this type of study. All patients were followed for clinical events until the planned end of the trial. Hence, data were also available on the reasons why follow up exercise tests were not performed.

The routine analysis of exercise capacity data for a trial such as PICO is limited to patients who had exercise testing during follow up. One so-called 'per protocol' approach is to consider only patients who completed at least one follow up exercise test on study medication. Another is to consider all patients who performed at least one follow up exercise test but to disregard whether the patient was

still taking study medication at the time of the test. This second approach is often described as 'intention-to-treat'.

The routine analyses of the PICO trial have been published.[4] As expected based on previous similar trials, there was a statistically significant positive effect on exercise capacity. Relative to placebo, there was also a trend towards an increased mortality in patients assigned to pimobendan. This raises the question to what extent the positive effect on exercise capacity is negated by the negative effect on mortality. In the publication, this question is addressed by an analysis that combines information on exercise duration for patients who underwent the 24-week exercise test (level 4) with information on clinical outcome (levels 1 – 3). The results are summarised in Table 3.

Table 3: PICO trial: ranked clinical outcome at 24 weeks

	Assigned treatment	
	Pimobendan *N* = 209 (100%)	Placebo *N* = 108 (100%)
Exercise test performed: same or higher exercise duration than at baseline	132 (63%)	64 (59%)
Exercise test performed: lower exercise duration than at baseline	48 (23%)	34 (31%)
Too sick to undergo exercise testing	5 (2%)	4 (4%)
Died before 24 weeks	24 (12%)	6 (6%)

Adapted from The Pimobendan in Congestive Heart Failure (PICO) Investigators.[4]

In this table all patients randomised (209 for both pimobendan dose groups combined, 108 for placebo) are represented based on a ranking of clinical outcome at the planned end of the trial into four mutually exclusive and exhaustive categories. The ranking has an unarguable order and starts with the best possible outcome: patients who were physically able to undergo the 24-week exercise test and who performed at least as well at this test as at baseline. The worst ranked outcome considered in Table 3 is death. Two intermediary outcomes are shown also. Note that the analysis in Table 3 is truly 'intention-to-treat' in the sense that all patients randomised are counted. Note also that because of this the percentages in the table can now directly be extrapolated to clinically meaningful probabilities. Patients who start taking pimobendan have a better chance to be alive after 24 weeks and to be able to exercise at least as well as at baseline as patients on placebo do (63% versus 59%). This is so despite the fact that patients who start taking pimobendan have twice the chance of dying within 24 weeks as patients on placebo (12% versus 6%). When only patients with exercise testing data at 24 weeks are considered, the chances of at least maintaining exercise capacity are 132/(132 + 48) = 132/180 = 73% for pimobendan and 64/(64 + 34) = 64/98 = 67% for placebo. However, while the effect

of pimobendan appears superior, the comparison of 73% versus 67% is not clinically meaningful because it ignores that more patients assigned to pimobendan than to placebo (29 versus 10 respectively) were unable to exercise at 24 weeks either because they were bed-ridden, or because of death. These patients should be taken into account when assessing the effect of a drug on a surrogate endpoint such as exercise duration. The example shows how this can be done based on a ranking that takes both the continuous outcome and clinical events into account.

In the example based on the PICO trial, ranking at one specific point in time after a relatively short duration of follow up was considered. The same principle can be applied to the analysis and presentation of results from trials with long-term follow up when the data allow ranking of the clinical condition at several time points during follow up.

Example 4: the Stockholm metoprolol trial

In the Stockholm metoprolol trial, metoprolol (a beta-blocker) was compared to placebo in patients who had MI.[5] All patients (154 metoprolol, 147 placebo) were followed for three years by regular out-patient clinic visits. At each visit the clinical status of those still alive was ranked on a five-point scale[6] as shown in Table 4 based on information belonging to levels 3 and 4.

Table 4: Stockholm metoprolol trial: total mean survival and mean number of days spent in different categories of clinical status

	Assigned treatment	
	Metoprolol *N* = 154	Placebo *N* = 147
Total mean survival (days)	992	964
1. After atherosclerotic complication	95	151
2. In NYHA class III/IV	78	91
3. In NYHA class II	541	546
4. In NYHA class I with side effects of treatment	57	20
5. In NYHA class I without side effects of treatment	221	156

NYHA, New York Heart Association. Adapted from Olsson *et al.*[6]

Category 1 – 'atherosclerotic complication' – was assigned to patients who were alive but who had sustained a reinfarction or a stroke, or who had undergone coronary bypass surgery, or a leg amputation for peripheral vascular disease (level 3). When a patient was assigned to this category at any one follow up visit, he/she was assigned to the same category at all subsequent follow up visits as long as he/she was alive. Categories 2 – 5 on the five-point scale describe the patient's sympto-

matic status as based on the New York Heart Association functional class and on the presence of suspected side effects of treatment (level 4). As long as the patient did not belong to category 1 at the time of a follow up visit, he/she could be in any of the categories 2 – 5. For each patient, the survival time in the study was taken as the time that had elapsed from randomisation to either death or end of follow up. This time was then subdivided for the time spent in each of the 5 categories. If the patient was in a different symptomatic status category at a given follow up visit than at the previous one, the assumption was made that the transition had occurred half-way between the two visits. Individual survival times and their subdivisions were then averaged for all patients in each of the two treatment arms (Table 4). The analysis shows that patients assigned to metoprolol on the average live longer (992 days compared to 964 days for placebo), and are also on the average free of cardiac symptoms or side effects for a longer period of time, 221 days compared to 156 days for placebo (Table 4). An analysis of this type combines information from all levels and the presentation of the results is better for patients who are being informed about what can be expected from the long-term use of treatment.

DISCUSSION

Claims based on trials published in major journals affect how drugs are marketed and how patients are treated. The example taken from the dofetilide trial raises questions about the basis for the claim made in the report that the treatment concerned reduced the risk of hospitalisation for worsening heart failure.[2] This particular claim is clinically relevant only if dofetilide increased hospitalisation-free survival. In this regard, no data was given in the report however. When asked in writing, the first author of the dofetilide trial report[2] was unable to provide the relevant data, citing that the project team had been dissolved.

The conclusions from the HOPE study were: “Ramipril significantly reduces the rates of death, MI, and stroke in a broad range of high-risk patients who are known to have a low ejection fraction or heart failure”.[3] This wording suggests that ramipril improves MI and stroke-free survival. While this conclusion is undoubtedly correct, the extent to which this is so cannot be assessed from the data published. It would have been helpful if the relevant table in the report had contained a line ‘none of the above’. The effect of ramipril on MI and stroke-free survival could then have been calculated directly from the data given.

Apart from the number of patients who reached the combined primary endpoint, the HOPE investigators also reported the number of patients who died from a cardiovascular cause, had MI or had a stroke at any time. That this was done was made clear by a footnote to the relevant table in the HOPE report.[3] Not all trials that rely on a combined primary endpoint report component endpoints in the same manner. In the NORDIL study[7] – which compared two different regimens for treating hypertension – the number of patients who reached each component endpoint was reported in the same manner as for HOPE. NORDIL was published back-to-back with a similar study called INSIGHT.[8] For the latter study, only the numbers

of patients who reached each component endpoint as the first manifestation of the primary endpoint were reported. The reader of the NORDIL report can find how many patients had a stroke. For INSIGHT on the other hand one can only find how many patients reached the primary endpoint because of a stroke. The total number of patients with a stroke is not given because a patient who had an MI followed by stroke and who is alive at the end of the study is counted in INSIGHT as MI only.

It is possible to report clinical events in a manner that also allows the reader to define combined endpoints. An example of how to do this can be found in the PICO report.[4] The principle is to define a priori a rank order from 'worst' to 'best' for the clinical events to be considered. For HOPE[3] this rank order could have been: 'death' – 'stroke' – 'MI'. That stroke precedes MI is based on a (perhaps debatable) judgement that morbidity from stroke is usually more serious than from MI. For two treatments the relevant results table must have four columns. In the first two, the total number of events is given by treatment arm. In the second two, the final outcome of each patient is counted once by treatment arm based on the worst event that occurred. For instance, a patient who sustains two MIs followed by a stroke but who is alive at the end of the study, contributes two MIs and one stroke to the event columns. The same patient is counted only once as a stroke in the outcome columns as this is the worst event that occurred. Reporting clinical events in this manner allows the reader to create directly from the table the combined event categories 'death or stroke', and 'death or stroke or MI'. Those interested in cost effectiveness are better served also by this presentation of events as the costs of care are determined to a large extent by the number and kind of morbid events that occur.

As explained earlier, MI and stroke-free survival is not addressed in the HOPE report.[3] This leaves relevant clinical questions concerning what can be achieved by treatment with ramipril unanswered. Because of the lower cardiovascular mortality in the ramipril arm, the number of patient-years of follow up in the ramipril arm must exceed that in the placebo arm. Even if ramipril, relative to placebo, does not increase the rate of non-cardiovascular death per unit time of patient follow up, this will result in a higher absolute number of non-cardiovascular deaths in the ramipril than in the placebo arm. In the case of HOPE, the difference is small – only 8 more cases of non-cardiovascular death in the ramipril arm (Table 2) – but goes in the expected direction. Obviously, this must be taken into account when assessing the impact of ramipril on mortality and morbidity. Incorporating all-cause mortality in a combined endpoint such as that used in HOPE has the additional advantage of removing bias that may have affected the classification of the cause of death. Bias in cause of death classification can occur even in double blind trials.[9,10]

The examples from the PICO[4] and the Stockholm metoprolol[5,6] trials expand on the relevance of event-free survival in therapy evaluation. Some treatments are associated with a risk of death that must be accepted in order to achieve symptomatic improvement, or avoid morbidity. Hip replacement is an obvious example. Were one to compare hip replacement with medical management in a large random-

ised trial, there is no doubt that the hip replacement arm would show a higher mortality than the medical management arm because of the risk of surgery. Mortality is not necessarily an appropriate surrogate for overall clinical benefit, and nobody seems to bother much about the surgical risk of hip replacement. The symptomatic improvement in survivors of surgery is apparently so important that the risk is considered worth taking.

A similar situation exists with respect to drugs in the same class as pimobendan. These drugs improve exercise capacity in patients with heart failure, but also increase mortality. The conventional presentation of the results of a trial such as PICO is a per-protocol analysis for the continuous outcome of interest. While information on clinical outcome is usually given also, it is generally not possible to summarise from published results the data in the same manner as in Table 3. Because of this it is impossible to judge whether the higher mortality risk negates symptomatic improvement. Therapy selection requires a value judgement. Those who are severely disabled by reduced exercise capacity may want to take the drug to increase the chance of enjoying symptomatic improvement. Those who are more concerned about dying will avoid the drug. It should be emphasised that only a presentation of the data as in Table 3 can show whether there is room for personal preference in treatment selection. Had the mortality in the pimobendan group been even higher, it is conceivable that the difference between the groups concerning the chances of being alive after 24 weeks and being able to exercise at least as well as at baseline would have disappeared completely. This could occur even if an analysis confined to patients who were able to undergo exercise testing at 24 weeks showed a positive effect.

We consider it inappropriate to use the term ‘intention-to-treat’ when analysis is confined to patients who were able to undergo exercise testing at 24 weeks. Such an analysis should perhaps be called ‘pseudo intention-to-treat’ to distinguish it from a real intention-to-treat that incorporates all patients randomised, as is the case in a mortality trial. Carrying forward the last available exercise duration value obtained for patients who died or who were unable to exercise does not solve the problem of ‘pseudo-intention-to-treat’. All patients randomised may be included in the analysis, but the reasons that follow up assessments were missed are ignored. In the PICO trial, the positive effect on exercise capacity was partly negated by the trend towards a higher mortality in patients treated with pimobendan. Had this trend been in the other direction, incorporation of mortality in the analysis for exercise capacity would have enhanced the apparent treatment effect. It is generally accepted that drugs from the same class as ramipril reduce morbidity and mortality in patients with heart failure, but do not have a major effect on exercise capacity. A reanalysis of data from exercise capacity trials with ramipril in heart failure has shown that the latter conclusion may have been the result of failing to take morbidity and mortality into account when analysing the effect of these drugs on exercise capacity.[11]

The example taken from the Stockholm metoprolol trial[6] shows how an analysis that combines several different types of follow up information available at one time point (as for the PICO trial[4]) can be extended to several time points. This method is particularly suited to describe effects on morbidity and mortality observed in trials with a long duration of follow up. In the Stockholm metoprolol trial, the treatment both prolonged mean survival and mean survival without symptoms.[6] In the case a certain treatment reduced mean survival because of an increased mortality risk, but prolonged mean survival without symptoms because of a positive effect on symptoms, an analysis of this nature would show this. Hence, such methods should be particularly appropriate to assess therapies that combine a positive effect on symptomatic status with no or an adverse effect on mortality. The Stockholm metoprolol trial had a fixed planned follow up of three years. This is not required for this analysis method; individual survival times and their subdivisions can readily be calculated also for trials with a variable duration of follow up.

An essential feature of Kaplan-Meier curves showing the probability of freedom from hospitalisation for worsening heart failure as in Figure 1 is that death preceding hospitalisation is considered as censoring. Kaplan-Meier analysis is based on the assumption that "there is no information in the times of censored observation".[12] In our view censoring for death is a violation of this principle. The same objection can be made to the Kaplan-Meier curves showing the probability of the primary endpoint over time that may be found in the reports on the HOPE[3] and other trials.[7,8] Censoring because of non-cardiovascular death cannot be considered as non-informative.

Although exceptions can be quoted both from past[13] and from ongoing trials,[14] ignoring non-cardiovascular death from a cardiovascular combined primary endpoint has become widespread at least in cardiovascular disease trials. One reason for this may relate to the temptation of choosing the primary endpoint such that one can expect the p-value to be as low as possible. Including non-cardiovascular deaths in a cardiovascular combined primary endpoint will generally lead to higher p-values unless the treatment reduces non-cardiovascular death also. This effect on the p-values-to-be is entirely predictable and the statement that 'it was prespecified to consider only cardiovascular death in the primary analysis' is neither reassuring nor relevant.

We have dealt neither with the question how to assess the overall statistical significance of differences between groups for any of the analyses discussed, nor how to assess the statistical significance of the individual components of a combined endpoint. The analysis for the PICO trial[4] summarised in Table 3 is an example of a worst-rank score analysis with informatively missing observations as discussed recently by Lachin.[15] Power calculation methods for composite ranking of survival and non-fatal outcome have been published by McMahon *et al*.[16] These authors deal with the power of a rank test, assuming that the treatment has a beneficial effect on the non-fatal outcome but no effect on mortality. This assumption

may not always be tenable and further theoretical development in this area would be valuable.

The examples given focused on single trials. The same methods could be used in meta-analyses. For instance, one could estimate survival curves based on several large mortality trials. From other studies one could obtain estimates of ranked symptomatic status for survivors at various time points. Results could then be combined into sets of curves showing survival in each category of symptomatic status considered and could then be presented in the same form as in Table 4 based on a surface area calculations.

Finally, it must be noted that the use of a combined endpoint should be considered in the design stage of a trial. As an example, consider again an exercise capacity trial in patients with heart failure such as PICO. In the latter, follow up exercise tests were also performed in patients who had been withdrawn from study medication.[4] It is customary to define both the primary endpoint and the primary analysis for the endpoint chosen before the trial starts. If the primary analysis specified is a per-protocol analysis confined to patients who had their follow up assessments of the chosen endpoint while still on study medication, it is tempting to omit the assessments that remain after premature withdrawal of study medication because these will not be used in the primary analysis. The protocol of the PICO trial specified that exercise tests had to be performed as planned also in patients who were withdrawn from study medication before the end of the study. Had this not been done, the ranked analysis at the end of follow up presented earlier would not have been possible. It should be standard practice always to perform all assessments as planned also for patients who are withdrawn from study medication ahead of schedule.

REFERENCES

1. Charache S, Terrin ML, Moore RD, *et al*. Design of the multicentre study of hydroxyurea in sickle cell anemia. *Control Clin Trials* 1995;**16**:432-46
2. Torp-Pedersen C, Møller M, Bloch-Thomsen PE, *et al*. Dofetilide in patients with congestive heart failure and left ventricular dysfunction. *N Engl J Med* 1999;**341**:857-65
3. The Heart Outcomes Prevention Evaluation Study Investigators. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *N Engl J Med* 2000;**342**:145-53
4. The Pimobendan in Congestive Heart Failure (PICO) Investigators. Effect of pimobendan on exercise capacity in patients with heart failure: main results from the Pimobendan in Congestive Heart Failure (PICO) trial. *Heart* 1996; **76**:223-31
5. Olsson G, Rehngvist N, Sjögren A, Erhardt L, Lundman T. Long-term treatment with metoprolol after myocardial infarction. Effect on 3-year mortality and morbidity. *J Am Coll Cardiol* 1985;**5**:1428-37

6. Olsson G, Lubsen J, Van Es GA, Rehnqvist N. Quality of life after myocardial infarction: effect of long term metoprolol on mortality and morbidity. *Br Med J* 1986;**292**:1491-93
7. Hansson L, Hedner T, Lund-Johansen P, Kjeldsen SE, Lindholm LH, Syvertsen JO, Lanke J, de Faire U, Dahlöf B, Karlberg BE for the NORDIL study group. Randomised trial of effects of calcium antagonists compared with diuretics and beta-blockers on cardiovascular morbidity and mortality in hypertension: the Nordic Diltiazem (NORDIL) study. *Lancet* 2000;**356**:359-65
8. Brown MJ, Palmer CR, Castaigne A, de Leeuw PW, Mancia G, Rosenthal T, Ruilope LM. Morbidity and mortality in patients randomised to double-blind treatment with a long-acting calcium-channel blocker or diuretic in the International Nifedipine GITS study: Intervention as Goal in Hypertension Treatment (INSIGHT). *Lancet* 2000;**356**:366-72
9. The Anturane Reinfarction Trial Research Group. Sulfinpyrazone in the prevention of sudden death after myocardial infarction. *N Engl J Med* 1980;**302**: 250-6.
10. Temple R, Pledger GW. The FDA's critique of the anturane reinfarction trial. *N Engl J Med* 1980;**303**:1488-9.
11. Lubsen J, Chadha DR, Yotof YT, Swedberg K. Meta-analysis of five exercise capacity trials in chronic congestive heart failure: effect of ramipril on mortality and morbidity. *Am J Cardiol* 1996;**77**:1191-6
12. Altman DG. Practical statistics for medical research. Chapman & Hall, London, 1991, p 385.
13. Scandinavian Simvastatin Survival Study Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994,**344**:1383-9.
14. Lubsen J, Poole-Wilson PA, Pocock SJ, van Dalen FJ, Baumann J, Kirwan BA, Parker AB. Design and current status of ACTION: A Coronary disease Trial Investigating Outcome with Nifedipine GITS. *Eur Heart J* 1998;**19** (Suppl I):I20-I32.
15. Lachin JM. Worst-rank score analysis with informatively missing observations in clinical trials. *Control Clin Trials* 1999;**20**:408-22
16. McMahon RP, Harrell Jr FE. Power calculation for clinical trials when the outcome is a composite ranking of survival and a non-fatal outcome. *Control Clin Trials* 2000;**21**:305-12

Chapter 4

Design and current status of ACTION: A Coronary disease Trial Investigating Outcome with Nifedipine GITS

ABSTRACT

Aims: To present the design of ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS), an ongoing multi-centre clinical outcome trial with nifedipine GITS (Gastro-Intestinal Therapeutic System) in patients with stable angina pectoris.

Methods: At least 6,000 patients with optimally treated stable angina without depressed left ventricular function are randomised in equal proportions to either nifedipine GITS or matching placebo (starting dose 30 mg, maintenance dose 60 mg once daily). Patients are followed for at least four years. The primary endpoint, to be analysed by assigned treatment, includes all-cause mortality, acute myocardial infarction, emergency coronary angiography for refractory angina, overt heart failure, debilitating stroke and peripheral revascularisation. For this endpoint, the trial has a power of 95% to detect a relative risk reduction of 18% at the 5% level of significance, and is large enough to exclude an excess mortality caused by nifedipine GITS of 3·1 deaths per 1,000 years of treatment or greater. The prespecified early termination rule is more conservative in the case of a beneficial effect than in the case of an adverse effect of nifedipine GITS. The first patient was randomised on November 29, 1996. By the end of April 1998, about 5,200 patients had been started on study medication.

Conclusions: Results will be available in the autumn of 2003.

The most common symptoms in patients with stable atherosclerotic coronary artery disease are regular attacks of anginal chest discomfort due to transient myocardial ischaemia.[1] The clinical management of this 'stable angina' syndrome has two objectives: to prevent or alleviate anginal attacks, and to prevent complications of the disease (such as sudden cardiac death, myocardial infarction and overt heart failure). Despite important recent advances (coronary revascularisation, lipid lowering), current treatment provides no cure and many patients remain symptomatic. Hence, maintenance treatment with anti-anginal drugs is often needed.

For many years, nitrates, beta-blockers and calcium channel antagonists (CCAs) have been the cornerstones of treatment for angina.[2,3] Nifedipine is the most widely used CCA for this indication. Historically, anti-anginal drugs have been introduced based primarily on proof of efficacy in reducing the symptoms. Their long-term safety has been of less concern although the safety of some anti-anginal drugs (notably the beta-blockers) is supported by the positive results of trials in patients with a history of acute myocardial infarction.[4] All this has changed due to the recent debate on the safety of CCAs, which resulted from a meta-analysis of trials in coronary disease[5,6] and observational data in hypertension.[7] The single most important conclusion on which all sides in this debate agreed is that there is a lack of data from well-designed long term trials.[8-16]

ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) is a multi-centre, randomised, double blind, placebo controlled trial that has started recruitment on November 29, 1996. Its primary objective is to assess, relative to placebo, the effect of nifedipine on the cardiovascular event-free survival of ambulatory patients who otherwise receive optimal treatment for stable angina and who do not have severely depressed left-ventricular function. The planned sample size is at least 6,000 patients with mean treatment duration of five years.

ACTION is the first large-scale trial with an anti-anginal drug in patients with stable angina that focuses on the long-term clinical outcome rather than on the anginal symptoms. The GITS (Gastro-Intestinal Therapeutic System) extended-release formulation of nifedipine is used. During its passage through the intestine, the GITS tablet releases active substance at a constant rate. Chronic once daily administration results in an even plasma level above the therapeutic minimum concentration for 24 hours.[17] The debate on the safety of nifedipine was based on data from studies using the original capsule formulation of this compound, which causes large plasma variations.[18] It is therefore opportune that the ACTION trial is undertaken, not only because of the general concern about the safety of CCAs, but also because a GITS formulation of nifedipine has become available that avoids large plasma concentration variations. This potentially more optimal formulation[17,19,20] has thus far not been tested in large scale long term trials in patients with stable angina. Below, we describe the main design features of this ongoing trial.

ACTION is not the only trial that will provide data on the long term safety of nifedipine GITS. To assess the efficacy and safety in hypertension of the same CCA, the INSIGHT (International Nifedipine GITS Study – Intervention as a Goal

in Hypertension Treatment) trial was mounted.[21] Recruitment of INSIGHT was completed in 1995 and 6,570 patients have been randomised. Results are expected in 1999.

TRIAL ORGANISATION

ACTION has been designed and is carried out by an independent research group. The role of the sponsor (Bayer) is limited to drug supply and on-site monitoring. The sponsor is not involved in data base management and has no direct access to the study medication code (see also below). A brief overview of the organisational structure of ACTION is as follows (a list of participants is given in Appendix I):

Coordination

The trial is coordinated by an independent research institute, which is responsible for overall trial management, data base management, quality assurance and reporting. These tasks are handled from two regional coordinating centres: one for Canada and one for the rest of the world. The same research institute is also responsible for the scientific programme of all trial related meetings.

Investigators

More than 300 investigators in 19 countries in Europe, Canada, Israel, Australia and New Zealand are participating. Most of these are hospital-based cardiologists. A complete list of participating centres can be obtained upon request. Investigators are required to inform the Steering Committee about participation in other studies that compete for the same type of patient.

Steering Committee

This committee approved the final version of the protocol and maintains scientific integrity while the trial is ongoing up to reporting of the main results. It has no access to the study medication code and has an independent chairman. The protocol specifies that members, who also represent the various participating countries, must discuss potential conflicts of interest with the chairman. Meetings are usually attended by designated non-voting representatives of the other parties involved but the committee meets also in the absence of these.

Executive Committee

This committee consists of the chairman, the three co-chairmen of the Steering Committee and designated representatives of the other parties involved. It is responsible for day-to-day management of the trial.

Data Monitoring and Ethical Review Committee

This committee consists of four clinical scientists with a statistician-chairman who are otherwise not involved in any way with the trial or its sponsor. Its main task is

to perform the pre-specified interim analyses. The random treatment allocation plan for the trial was generated by the chairman and access to the study medication code is controlled by him.

Critical Events Committee

This committee, which consists of six investigators and an external neurologist, decides on the final diagnostic classification of critical clinical events.

Core laboratories for echocardiography

Locally recorded echocardiograms are analysed centrally by a standardised computer-assisted method at the core laboratories of the regional coordinating centres. Investigators cannot participate in the study before the core laboratory has approved a test recording.

Study medication supply

Study medication for clinics in Canada and Israel is packed and supplied by the pharmacy of Bayer Inc, Etobicoke, Ontario, Canada. Medication for the rest of the world originates from the pharmacy of Bayer AG, Leverkusen, Germany.

On-site monitoring

During recruitment each centre is monitored on site according to good clinical practice (GCP) every six to eight weeks. After recruitment has been completed, the frequency will be reduced to every three months. Local Bayer subsidiaries are responsible for on-site monitoring. In some countries, sub-contractors have been engaged by the sponsor.

Endpoint verification

For all serious adverse events, the documentation and relevant patient data are verified on-site by coordinating centre personnel before the data are submitted to the Critical Events Committee for diagnostic classification.

Auditing

The trial will be audited by the Clinical Quality Assurance department of the sponsor. In addition, independent audits may be performed upon instruction from the Steering Committee.

ETHICS

ACTION is carried out in accordance with the provisions of the Declaration of Helsinki (last revised version, 1987) and ICH guidelines for good clinical practice (GCP). [22] Local ethics committees approved the protocol and the informed consent procedure for each centre. The trial is monitored on-site according to GCP.

To protect patient safety while the trial is in progress, an independent monitoring committee *(see trial organisation)* will perform pre-specified interim analyses.

PATIENT SELECTION

ACTION focuses on the following three subgroups of patients with angina pectoris:

1. those with a history of acute myocardial infarction who have had angina ever since, or have developed angina again after an angina-free period;
2. those who had a coronary revascularisation procedure but continued to have angina, or developed recurrent angina;
3. those who have angina presumably caused by coronary artery disease, but have no history of infarction or revascularisation.

The specific inclusion and exclusion criteria are listed in Table 1.

Table 1: Selection criteria

Inclusion criteria

1. Age 35 years or older.
2. In a stable clinical condition for at least one month while requiring oral and/or transdermal treatment for symptomatic angina either to prevent, or to treat, anginal attacks.
3. On unchanged oral treatment for symptomatic angina without a calcium channel blocker during the last two weeks.
4. Presence of at least one of the following criteria for coronary artery disease:
   - 4.1 Unequivocal myocardial infarction (presence of two of the following three: typical chest pain lasting at least 30 minutes, cardiac enzymes at least twice the upper limit of normal, ECG changes typical for acute myocardial infarction) at least three months prior to start of study medication.
   - 4.2 PTCA or CABG at least three months prior to start of study medication.
   - 4.3 Coronary artery disease documented by angiography.
   - 4.4 If coronary angiography has never been performed: a positive exercise test ($\geq$ 1 mm ST-depression at maximal exercise), or a perfusion defect (thallium scan).
5. Screening left-ventricular ejection fraction at least 40%.
6. On lipid-lowering treatment, if such treatment is indicated based on current internationally accepted guidelines.
7. Ambulatory, able to come to the out-patient clinic on his/her own and willing to participate based on a signed Declaration of Consent in accordance with national laws and regulations.

Exclusions related to medical history

1. Coronary or peripheral revascularisation; acute myocardial infarction, unstable angina, syncope, stroke or major surgery within three months before start of study medication.
2. Planned coronary angiography or revascularisation (PTCA or CABG).
3. Known intolerance to dihydropyridine calcium channel blockers.
4. Clinically significant valvular disease.

5. Severe obstructive airway disease.
6. Unstable insulin-dependent diabetes mellitus requiring frequent changes in insulin dosing.
7. Chronic or intermittent diarrhoea, ulcerative colitis, regional enteritis or any other gastro-intestinal condition that could result in incomplete absorption of nifedipine GITS.
8. Severe gastro-intestinal stenosis which could hinder the passage of GITS tablets.
9. Presence of any condition other than coronary artery disease which limits life expectancy.

Exclusions related to current symptoms or findings

10. Clinically significant heart failure, based on the presence of at least two of the following: in NYHA class II or higher, peripheral oedema, presence of rales over two-thirds of the chest, pulmonary congestion on a chest X-ray, cardio-thoracic ratio greater than 0·5.
11. Symptoms of orthostatic hypotension or supine systolic blood pressure 90 mm Hg or lower.
12. Systolic blood pressure 200 mm Hg or higher and/or diastolic blood pressure 105 mm Hg or higher despite blood-pressure lowering treatment.
13. Plasma or serum creatinine above twice the local upper limit of normal.
14. Alanine aminotransferase (ALAT; old name GPT = glutamic-pyruvic transaminase) or aspartate aminotransferase (ASAT; old name GOT = glutamic-oxaloacetic transaminase) above three times the local upper limit of normal.

Exclusions related to current treatment

15. On a daily dose of diuretics exceeding 20 mg furosemide, 5 mg bendroflumethiazide or equivalent.
16. On combination therapy with an ACE-inhibitor and a diuretic for heart failure.
17. Treated with any of the following which could not be stopped: calcium channel blockers, cardiac glycosides (unless given for supra-ventricular arrhythmias), other positive inotropic agents, class I or III anti-arrhythmics other than amiodarone or sotalol, cimetidine, anti-psychotic and anti-epileptic drugs, rifampicin or rifampine.

Miscellaneous exclusions

18. Problems with compliance or follow up anticipated.
19. Pregnancy, breast feeding or risk of pregnancy (fertile women using an acceptable method of contraception according to local regulations can participate).
20. Participation in another trial or study.

Only patients in stable clinical condition who require oral and/or transdermal treatment either to prevent, or to treat, anginal attacks are eligible. It is not required that patients are actually symptomatic at the moment study medication is started; it suffices that in the past anti-anginal treatment was started specifically to treat anginal attacks.

As nifedipine is contra-indicated in patients with depressed left ventricular function, inclusion criterion 5 (Table 1) specifies that the left-ventricular ejection

fraction must be at least 40%. To establish that this is the case, a two-dimensional echocardiogram must be made during screening. This is recorded on videotape and re-analysed centrally. Investigators are not required to postpone start of study medication until the Core Laboratory has also assessed the ejection fraction. If an echocardiogram could not be obtained (some patients have an unsuitable anatomy for this purpose), participation based on an ejection fraction value measured by another method is allowed. Other criteria are used to exclude patients who may develop heart failure because of another cardiac condition (exclusion criterion 4, Table 1), or already have clinical heart failure (exclusion criterion 10 and 15-17). The remaining exclusion criteria in Table 1 either refer to incompatible medication which cannot be stopped, or to co-morbidity that could interfere with the assessment of the effect on nifedipine GITS on cardiovascular event-free survival.

ENDPOINTS

The primary endpoint is the combined rate of death from any cause, unequivocal acute myocardial infarction, emergency coronary angiography for refractory angina, hospitalisation for overt heart failure, debilitating stroke, and peripheral revascularisation. This endpoint can also be viewed as 'cardiovascular event-free survival', i.e. survival free of the events mentioned.

All patients will be followed for all events that may occur until death or until the planned end of the trial. Hence, event rates will also be reported separately for each specific event incorporated in the primary endpoint. Unless done on an emergency basis, coronary angiography and coronary revascularisation procedures without complications are not included in the primary endpoint.

The occurrence of the events incorporated in the primary endpoint is monitored by standard reporting procedures for serious adverse events. As soon as a serious adverse event has been reported, investigators are required to supply additional pre-specified clinical information to allow for diagnostic classification by the Critical Events Committee according to pre-set criteria. This committee has no access to the study medication code, even if the treating physician has broken the code to decide on further treatment in an emergency.

STUDY MEDICATION DOSE, SUPPLY AND BLINDING

The study medication (which is added to treatments already in use, see below) consists of either nifedipine GITS or matching placebo. The starting dose is 30 mg once daily. If this is well tolerated, the dose must be increased to 60 mg once daily within six weeks. Study medication is then continued at this dose until the planned end of the trial (i.e. at least four years for the last patient started on study medication). If signs of intolerance to the treatment regimen occur, investigators are encouraged to first attempt resolving the problem by adjusting the concomitant medication regimen. If this is not successful, the dose of study medication may be reduced again to 30 mg; or study medication may be interrupted. Clinical situation permitting, investigators are always allowed to either restart study medication; or

increase the dose to 60 mg. The aim is to keep patients on study medication as much as possible. The only restriction is that ejection fraction must be reassessed by echocardiography, and must be above 40%, before study medication can be re-started after a clinical event that could have reduced ventricular function (such as acute myocardial infarction, bypass surgery).

In the case study medication had been started in violation of the selection criteria, investigators are instructed by the coordinating centre to stop study medication.

Initially study medication was allocated centrally by telephone every time the patient visited the clinic. Because of logistic problems, this system was stopped in March 1997 and replaced by a conventional system of drug supply and preservation of blinding. This included the generation of a new random treatment allocation plan, stratified by centre. Centres are now supplied directly by the pharmacies of the sponsor with numbered study medication packs that contain both the 30 and 60 mg dose for at least six months of treatment. No telephone call is required to identify the study medication to be handed out.

Blinding is maintained throughout in several ways. The chairman of the Data Monitoring and Ethical Review Committee generated the currently used random treatment allocation plan and sent one copy to each of the two pharmacists responsible for packing purposes only. Coding envelopes for emergency unblinding by an investigator were also prepared. The on-site monitors check that these normally remain unopened. Unblinding for the purpose of reporting to regulatory authorities can only be obtained on a case by case basis through protected access to a telephone system operated by the chairman of the monitoring committee. A treating physician other than the investigator who has no access to the coding envelope can also use this system in a medical emergency. All instances of code breaking are reported to the coordinating centre (without mentioning the actual code to maintain blinding). The coordinating centres have no access to the study medication code until the end of the trial. For the purpose of confidential interim analyses, the chairman of the monitoring committee will combine the code with data supplied by the coordinating centres.

CONCOMITANT TREATMENT

After start of study medication, all patients continue to receive the concomitant treatment regimen on which they have been stabilised before. Drugs that cannot be combined with study medication are specified in the selection criteria (see Table 1).

During screening, any patient not already on lipid lowering treatment must be evaluated for this treatment based on at least a screening assessment of total cholesterol. Investigators are free how to assess and treat elevated lipids, but must comply with current nationally accepted guidelines for lipid lowering.[23] In patients who appear to require lipid lowering during screening, treatment must be started at least one week before start of study medication.

Participation in ACTION does not restrict access to other generally accepted anti-anginal treatments other than CCAs. There are no restrictions as regards changes in concomitant treatment after start of study medication provided that the adjustments made are compatible with the use of nifedipine. Coronary revascularisation may be performed to treat angina that cannot be managed satisfactorily by adjusting the concomitant treatment regimen. Whether study medication is continued or interrupted during such a procedure, or during any other clinical event, is left to the discretion of the investigator. Study medication must be interrupted however when an indication arises for compounds mentioned in exclusion criterion 17 (Table 1).

FOLLOW UP

Two and six weeks after start of study medication, patients are seen at the outpatient clinic to assess whether the dose of study medication can be increased to 60 mg once daily. From then onwards the patient must be seen at the outpatient clinic at least every six months, and must be contacted in-between by telephone.

During outpatient clinic visits, the evolution of symptoms and signs, vital signs, changes in concomitant treatment and compliance with study medication must be documented. Also, a standard 12-lead electrocardiogram must be made. Routine laboratory tests are required 6 months after start of study medication; and after 2, 4, and 6 years. Telephone contacts require only that the patient is questioned about the evolution of symptoms and treatment compliance.

Unscheduled clinic visits for moderate or severe adverse events must be documented in similar manner as planned visits. Laboratory tests, electrocardiography and echocardiography (for assessment of ejection fraction when a decision must be taken about restarting previously interrupted study medication) done on clinical indication must be documented also. Otherwise, all clinical events and those diagnostic procedures which either carry a risk to the patient (such as angiography), or may affect treatment (such as exercise testing), must be documented as adverse events in a manner which is standard for phase III clinical trials. Serious adverse events must be reported to the coordinating centre within 24 hours of the investigator's awareness.

All patients who have taken at least one tablet of study medication (and are therefore considered as randomised) will be followed in this manner until a 'common stopping date' is announced to investigators. At present, this common stopping date is planned for four years after the last patient has been randomised. Planned visits and contacts take place irrespective of withdrawal of study medication ahead of schedule as long as the patient is willing and able to co-operate.

While participating in ACTION, patients are not allowed to participate in any other trial, ancillary or side-arm study unless this other trial or study has been approved by the Steering Committee.

DATA MANAGEMENT AND QUALITY ASSURANCE

Investigators enter data directly on conventional Case Report Forms printed on triplicate NCR paper. These are arranged visit by visit and contain extensive instructions and relevant definitions. All treatments are recorded in treatment logs. Standard adverse event forms are used to document adverse events and relevant clinical procedures that have been carried out. After verification of the data entered according to GCP, one copy of each completed Case Report Form page is removed by the on-site monitor and sent immediately by mail to the coordinating centre. There, a concurrent database is maintained. All incoming forms, original standard 12-lead electrocardiograms and copies of laboratory reports are scanned for electronic storage. After scanning, documents are archived. Data is entered in a specially developed relational data base management system. Data entry screens show the scanned document on the left half of the screen and the data entry fields on the right half. Scanned documents remain accessible on-line after data entry. The data entry module contains on-line range and logical checks. For data that are found missing, illegible or inconsistent, data clarification forms are generated which are sent to the on-site monitor for resolution.

Certain events must be reported immediately by the investigator by telefax on pre-printed forms directly to the coordinating centre. Examples are: informed consent and start of screening, start of study medication, serious adverse events and premature withdrawal of study medication. The reporting of serious adverse events complies with national regulatory requirements.

STATISTICAL CONSIDERATIONS

Sample size and power

Of all trials reported before ACTION began, the Scandinavian Simvastatin Survival Study[24] (4S) is closest as regards patient selection, size and duration of follow up. In the simvastatin arm of 4S, the combined hazard of death, acute myocardial infarction, emergency coronary angiography for refractory angina, overt heart failure, debilitating stroke and peripheral revascularisation (the primary endpoint of ACTION) was about 5·6 events per 100 patient-years (see Appendix II). This hazard was used as a basis for sample size calculations.

The planned size is 6,000 patients divided equally between nifedipine GITS and placebo who are followed for a minimum of four, and a mean of five years. For this size and mean follow up, the magnitude of the effect of nifedipine relative to placebo (expressed as a hazard ratio) that can be detected with a given power was calculated for varying assumptions about the hazard of the primary endpoint in the placebo arm (see Appendix II). Results are shown in Figure 1 *(next page)*.

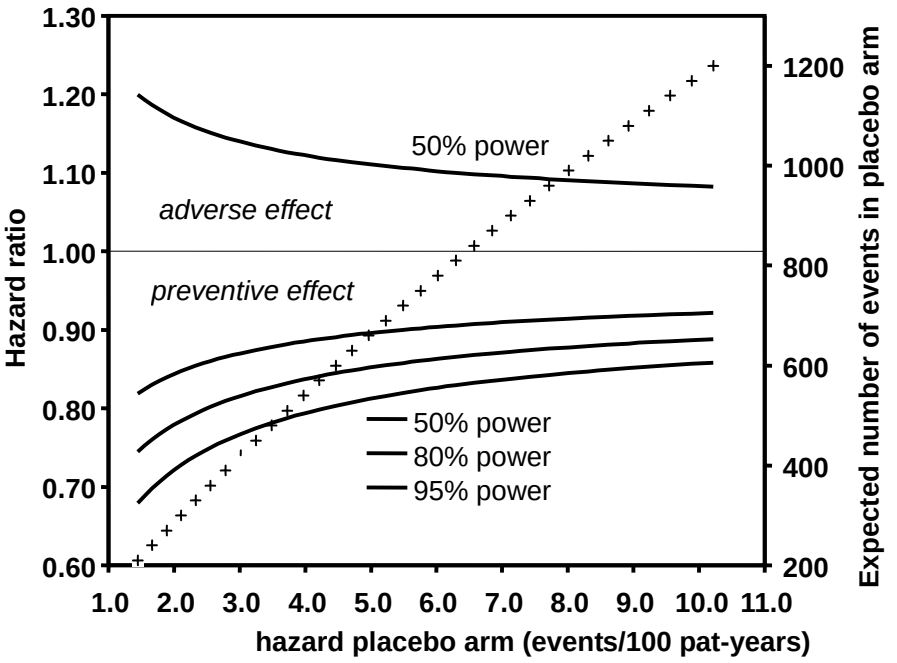


Figure 1: The hazard ratios that can be detected with a given power for a given hazard in the placebo arm

Assuming N = 3,000/arm, mean observation time = 5 years and two-sided $\alpha = 0{\cdot}05$, the hazard ratio (vertical axis to the left) which can be detected with a given probability (power) is plotted for the power choices shown (uninterrupted curves) against the hazard in the placebo arm, expressed as number of events per 100 patient-years (horizontal axis). The figure also shows (+++ curve) the expected number of events in the placebo arm (vertical axis to the right) and a 50% power curve for detecting an adverse effect of nifedipine GITS.

When the hazard of the primary endpoint in the placebo group is assumed to be 5·6 (as in the simvastatin arm of 4S, see above), Figure 1 shows that ACTION has 95% power to detect a hazard ratio of 0·82 (i.e. an 18% reduction of the primary endpoint by nifedipine, relative to placebo) with a 5% two-sided level of significance. The power to detect hazard ratios of 0·86 (14% reduction) and 0·90 (10% reduction) is 80% and 50% respectively (Figure 1). These hazard ratios represent relatively modest effects of nifedipine on the primary endpoint.

ACTION is also undertaken to confirm that nifedipine GITS is safe for chronic clinical use. To assess the excess event rate attributable to nifedipine GITS that can be excluded when the event rates are equal in both treatment arms, two-sided confidence intervals around a hazard ratio of one were calculated for varying assumptions about the hazard in the placebo group (see Appendix II). Results are shown in Figure 2 (*next page*).

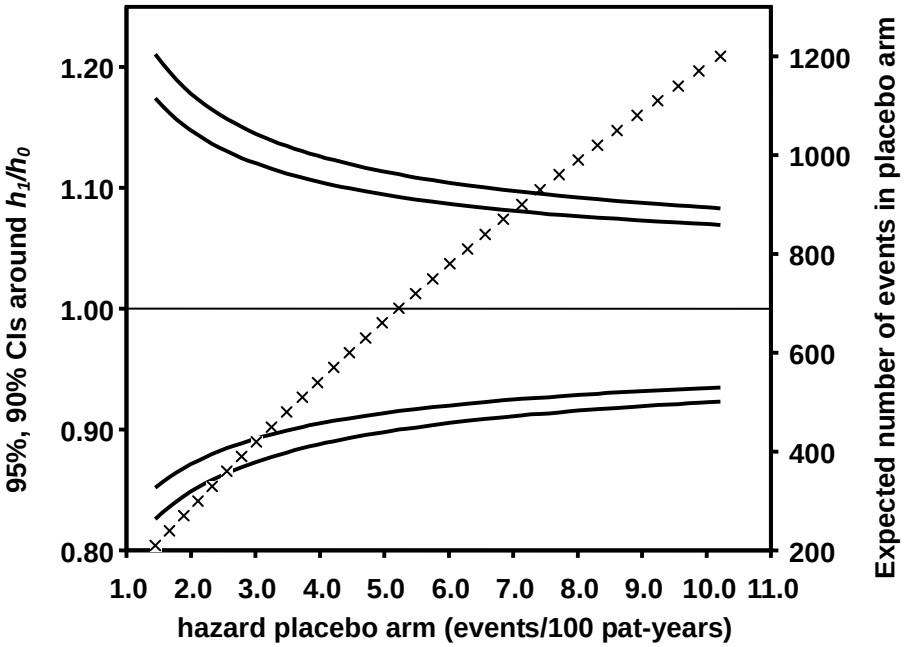


Figure 2: Two-sided 90% and 95% confidence intervals around a hazard ratio of one for a given hazard in the placebo arm

Assuming N = 3,000/arm, mean observation time = 5 years and two-sided α = 0·05, two-sided 90% (inner curves) and 95% (outer curves) confidence intervals (CIs) around a hazard ratio of one (vertical axis to the left) are plotted for a given hazard in the placebo arm, expressed as number of events per 100 patient-years (horizontal axis). The figure also shows (+++ curve) the expected number of events in the placebo arm (vertical axis to the right).

When the hazard of the combined endpoint monitored for safety (all-cause death, myocardial infarction and stroke) in both ACTION treatment arms is assumed to be 3·2 per 100 patient-years (as observed in the simvastatin arm of the 4S study, see Appendix II), the 95% confidence interval ranges from 0·88 to 1·14 (around a hazard ratio of one). The 90% confidence interval is shown also in Figure 2. Hence, when the combined rates of death, non-fatal acute myocardial infarction and stroke are as assumed, an increase in the rate of these events caused by nifedipine GITS in excess of 14% can be excluded with confidence. Assuming a mortality of 1·5 per 100 patient-years (as in the simvastatin arm of 4S) in both treatment arms, the corresponding upper limit of the 95% confidence interval is 1·21. In absolute terms this represents the exclusion of an excess mortality of 3·1 deaths per 1,000 years of treatment or greater.

For the hazards mentioned above, the hazard ratio's which can be detected for a given power and the number of events to be expected (calculated with the formu-

lae given in Appendix II, assuming n = 3,000 and t = 5 years) are tabulated in Table 2.

Table 2: Relative and absolute reductions which can be detected with a given power

Hazard placebo arm (events/100 patient-years)	Hazard ratio	Relative reduction (%)	Expected number of events placebo	Expected number of events nifedipine	Absolute reduction
5·58 (primary criterion for efficacy):					
95% power	0·82	18	731	614	117
80% power	0·86	14	731	639	92
50% power	0·90	10	731	667	64
3·23 (primary criterion for safety):					
95% power	0·77	23	447	353	94
80% power	0·82	18	447	373	74
50% power	0·87	13	447	395	52
1·45 (mortality):					
95% power	0·68	32	210	144	66
80% power	0·74	26	210	158	52
50% power	0·82	18	210	173	37

N = 3,000/arm. Mean observation time = 5 years. Two-sided α = 0·05

Guideline for early termination

The guideline for early termination as defined in the protocol is reproduced in Appendix III. The guideline is based on 30,000 planned patient-years of follow up (6,000 patients with a mean follow up of 5 years) and distinguishes between a preventive and an adverse effect of nifedipine observed at a pre-specified interim analysis. For a preventive effect (efficacy), the criterion for early termination is the primary endpoint of the study evaluated after 10,000 and after 20,000 patient-years of follow up. For an adverse effect (safety), the criterion monitored is the combined rate of death of any cause, acute myocardial infarction and stroke. This criterion is evaluated after every 5,000 patient-years of follow up. Although the guidelines for efficacy and for safety are both based on a two-sided p-value of 0·05, the boundary shape is different. As a consequence, the boundary for efficacy is initially more conservative than for safety.

Statistical analysis

The essential features of the planned statistical analysis are as follows. Two analysis populations are distinguished: all-randomised and valid-for-efficacy. The former consists of all patients who have taken at least one tablet of study medication and the latter of the subset of patients who are documented to comply with the selection criteria. Using Kaplan-Meier plots and log-rank tests for comparing censored time-to-event data, the primary endpoint will be analysed by assigned study medication. Effects of nifedipine GITS relative to placebo will be expressed as hazard ratios with 95% confidence intervals. Analyses will be presented both for the all-randomised and for the valid-for-efficacy population but analyses for the all-randomised population will be considered as 'primary'. It is expected that the two analyses will give essentially the same results.

To assess the relationship between patient characteristics at entry and event hazards, multivariate proportional hazard regression analysis according to Cox[25] will be used; with age, sex, NYHA class, ejection fraction, history of myocardial infarction, use of a beta-blocker at entry, use of lipid-lowering therapy at entry and use of a CCA before entry as pre-specified co-variates. Modification of the effect of nifedipine GITS by these patient characteristics (subgroup analysis) will be assessed using interaction tests.

In the protocol several secondary analyses are specified, some of which address pharmaco-economic issues, generalisability to specific patient sub-groups based on a multivariate risk score and effect of nifedipine GITS on NYHA class over time.[26]

CURRENT STATUS AND PLANNING

The first patient was recruited on November 29, 1996. After 205 patients had been started on study medication, recruitment was interrupted on March 16, 1997 due to the problems encountered with central telephone allocation of study medication. After the drug supply system had been redesigned, recruitment was resumed on April 22, 1997. By the end of April 1998; about 5,200 patients had been randomised and recruitment was ahead of schedule.

Recruitment is planned to be completed in July 1999 at the latest. To make up for patients who are entered in violation of selection criteria, it is expected that about 7,000 patients must be started on study medication. The last patient started will have a minimum of four years of follow up. Hence, the main results of ACTION are expected to be available in the autumn of 2003.

DISCUSSION

ACTION is designed to resolve whether nifedipine GITS favourably affects the long-term clinical outcome of patients with stable angina (secondary prevention). Nifedipine GITS is registered in many countries for the treatment of anginal symptoms, but nowhere for secondary prevention. In the case the results of ACTION show that nifedipine GITS indeed has a secondary preventive effect, an expansion of its registered indications may be sought. For this reason, ACTION has not been designed as a phase IV trial of an already marketed drug, but as a phase III trial intended as a basis for registration. Its design follows the current standard pattern for phase III trials, including detailed case documentation, immediate reporting of serious adverse events and on-site monitoring according to GCP.

Several design features of ACTION were a matter of debate before the trial was started. The selection criteria in earlier versions of the protocol required the presence of coronary artery disease but did not require that the patient be treated for anginal symptoms. In the light of the recent debate on the safety of nifedipine, it was considered inappropriate to treat asymptomatic patients with this compound even in the context of a clinical trial. Hence, inclusion was restricted to patients who are on anti-anginal treatment started to treat anginal symptoms (rather than for another indication, such as secondary prevention as in the case of a beta-blocker). Anginal symptoms must have occurred after the last infarct or revascularisation procedure. Hence, anti-anginal treatment other than a CCA (i.e. predominantly nitrates and/or beta-blockers) is part of the concomitant treatment regimen of all patients. It is therefore a limitation of ACTION that the study will not address the role of nifedipine GITS monotherapy in stable angina.

As eventual conclusions based on trial results are customarily driven by p-values, it is desirable to pre-specify one primary endpoint. All-cause mortality as a primary criterion was rejected for a variety of reasons. Death is not the only outcome that matters in symptomatic patients. Also, there were doubts that mortality could significantly be reduced even further in a patient group with already a low mortality due to recent advances in treatment. This being so, a composite endpoint which takes into account several aspects of the effect of a treatment is the best alternative. For ACTION, major cardiovascular event-free survival was adopted as the primary endpoint. All-cause mortality is by definition included in this endpoint. Non-fatal acute myocardial infarction, emergency coronary angiography for refractory angina, hospitalisation for overt heart failure, debilitating stroke and peripheral revascularisation were included as all of these represent cardiovascular morbidity which physicians aim to prevent when treating patients with stable angina. An additional argument to include stroke and peripheral revascularisation was that these two might possibly also be prevented by nifedipine due to its blood pressure lowering[27,28] and anti-atherosclerotic effects.[29,30] On the other hand, coronary revascularisation procedures as such were not included as these nowadays are so frequent that they can hardly be considered as morbidity to be avoided unless peri-procedural complications such as infarction or stroke occur. The primary endpoint chosen cov-

ers an important aspect of clinical safety since it includes death from non-cardiovascular causes. It is conceivable that a drug induces an excess mortality but at the same time increases morbidity-free survival (as is the case for many common surgical procedures). If this is the case for nifedipine GITS, ACTION will show this.

As regards the study medication, the option was considered to use 30 mg GITS tablets throughout and allow the investigator freedom to adjust the dose individually while maintaining the blinding. This option was rejected because 60 mg nifedipine GITS once daily is the standard recommended dose. Also considered was the option to precede start of double blind study medication by an open-label run-in phase on nifedipine GITS to assess tolerance and initial effects of the drug on symptoms and vital signs. While this may provide relevant additional information and allows exclusion of patients who do not tolerate the drug, this option was rejected because of the added complexity and because nifedipine GITS in the starting dose of 30 mg once daily is generally well tolerated.

The concomitant treatments that are allowed follow directly from the objectives of ACTION and the type of patient involved. Patients with symptomatic angina cannot be left untreated. Hence, all anti-anginal drugs and all other medication that can be combined with nifedipine are allowed as concomitant treatment. On the usual indications, a coronary revascularisation procedure may be performed at any time after start of study medication. The only restrictions are drugs that may interact with nifedipine, or are used for a cardiac or another condition that renders the patient unsuitable for the trial.

Lipid lowering as a component of the concomitant treatment regimen is of special interest as lipid lowering has been shown to have important secondary preventive effects in patients with coronary artery disease.[24] If investigators were left completely free in prescribing lipid lowering, and if the occurrence of symptomatic angina induces clinicians to assess lipids and start lipid lowering treatment, the fraction of patients using this treatment at the end of the trial in the placebo arm would be higher than in the nifedipine arm despite randomisation and double blinding. When this occurs, nifedipine would not have a fair chance to show a secondary preventive effect. To prevent unbalance as regards lipid lowering as much as possible, investigators are required to evaluate all patients for this type of treatment during screening, and start treatment before start of study medication if indicated.

After an initial estimate, detailed sample size calculations were performed to assess the power of the trial for the sample size and duration of follow up chosen. The guideline for early termination after an interim analysis was constructed in such a way that the overall significance level assumed in the power calculations is maintained. The methods used are standard and computationally simple (see Appendix II) but the presentation of the results in Figure 1 is perhaps unusual. The figure shows that the hazard ratio that can be detected with a given power depends strongly on the event hazard in the placebo arm for low event rates. For higher event rates, the relationship is less steep. The hazard of the primary endpoint in the

placebo arm of ACTION is likely to be higher than the hazard of the same endpoint in the simvastatin arm of the 4S study[24] on which sample size calculations were based (see Appendix II) as all patients in this arm of 4S were treated with a potent lipid lowering agent. In ACTION this is unlikely to be the case.

The power calculations showed that ACTION has ample statistical power to detect relatively small effects on cardiovascular event-free survival. The trial is on the other hand likely to be criticised by some as having insufficient power for mortality only. This potential criticism has been specifically considered by a panel of international experts which was consulted while the design was finalised. Based on the calculations performed to assess the excess mortality which can be excluded with confidence in the case the mortality turns out to be the same in both treatment arms (see also Figure 2 and Appendix II), there was unanimous agreement that the chosen sample size of ACTION is sufficient to assess the clinical safety of a drug of which the efficacy in treating anginal symptoms is not in dispute.

The guideline for early termination (see Appendix III) is initially more conservative in the case of a beneficial than in the case of an adverse effect of nifedipine GITS. Various arguments can be made that a guideline for early termination should be asymmetrical. The reason in this case was that nifedipine GITS has no future when interim data show an adverse effect on the endpoints monitored for safety (death, myocardial infarction and stroke) as alternative anti-anginal agents are available. On the other hand, when no safety problem in this sense is observed 'ad interim' there is interest to complete the trial as planned in order to estimate the magnitude of any positive effects precisely. The guideline is constructed in such a way that there is probability of 0·025 that nifedipine GITS is declared unsafe while in fact no adverse effect on the endpoints monitored for safety exists.

An important question that must be considered already in the design phase of a trial is how to handle patients who are started on study medication in violation of the selection criteria. Under pressure to recruit, investigators tend to develop their own opinions on the interpretation and/or the clinical relevance of certain selection criteria. Hence, the inclusion of patients who are in fact not eligible is unavoidable unless the data are verified for each individual patient still in screening before investigators are allowed to start study medication. This is difficult to arrange in a large multi-centre trial and causes an additional burden to patients who must wait, or even come back another time, to hear whether clearance has been obtained to start study medication.

To avoid ambiguity when the results are presented, the protocol requires that investigators must follow ineligible patients started on study medication until the planned end of the trial. Such patients are considered as belonging to the all-randomised population in the analysis plan, which will be analysed based on the conventional concept of 'intention-to-treat' analysis. The protocol also specifies an analysis by treatment assignment that is confined to eligible patients only (the valid-for-efficacy population). This analysis will show whether the primary analysis is unduly distorted by the presence of ineligible patients.

The strategy for the handling of ineligible patients has unwanted consequences for trial management. Investigators who are required to follow ineligible patients as planned require to be supported for this. This is in fact an extra incentive for investigators to start study medication in ineligible patients. By diligent monitoring and exclusion of investigators who repeatedly start study medication in ineligible patients an effort is made to limit the number of ineligible patients entered into ACTION. At the same time a scientific discussion on how to handle ineligible patients entered in clinical trials is overdue.

The protocol prohibits that patients participate also in another trial or study unless approved by the Steering Committee. This was inserted to discourage investigators to arrange ancillary studies of their own that may interfere with the objectives of ACTION. While the trial was designed, it was proposed to assess both clinical and coronary angiographic outcomes in one large trial by also performing angiography during follow up in all patients. Complications that result from angiography for the purpose of a clinical study must be counted as clinical events but are unlikely to be affected by treatment. This would tend to dilute the effect of treatment on clinical outcome. For ethical reasons, angiograms for purpose of a clinical study cannot be left unanalysed until the study is completed. Hence, the management of patients in a trial with angiographic outcomes will not be representative for 'usual medical care', with angiography done only on indication. Based on these considerations, the study procedures of ACTION were limited to what is 'usual medical care' for patients with stable angina. Several ancillary studies in ACTION patients that do not require procedures that influence patient management, or carry a risk, have been approved by the Steering Committee. Ancillary studies focusing on quality of life, coronary calcification assessed by fast computer tomography and non-invasive assessment of carotid intima-media thickness are underway in subsets of patients. The prognostic value of QT-dispersion is also assessed in a subset, as is the evolution of left-ventricular mass by an additional echocardiogram at the end of the trial. In order to stay as close as possible to the intended nature of the trial, patients may participate only in one ancillary study that requires repeated additional procedures during follow up.

Large randomised trials have been criticised because the results are difficult to generalise to individual patients seen in clinical practice. This may be so because the clinical entity entered in a trial is not well defined, the treatment studied is not what would normally be done in any case, and/or because the outcome assessed is not clinically relevant. In designing ACTION, an effort has been made to avoid all three. When the results become available in 2003, this study will redefine the usefulness of nifedipine GITS in patients with stable angina.

APPENDIX I

ACTION participants are:

PRINCIPAL INVESTIGATOR AND STUDY CHAIRMAN
PA Poole-Wilson, Cardiac Medicine, National Heart & Lung Institute, Dove House Street, London SW3 6LY, UK, phone +44-171-351 81 79, fax +44-171-823 33 92.

INVESTIGATORS
313 centres in Australia (5), Austria (4), Belgium (6), Canada (49), Denmark (24), Finland (4), France(17), Germany(33), Greece(5), Israel(18), Italy(38), The Netherlands(28), New Zealand(3), Norway(18), Portugal(8), Spain(20), Sweden(17), Switzerland(3), UK(13).
Regional Coordinating Centres: SOCAR Research SA, PO Box 2564, 1260 Nyon 2, Switzerland, phone, fax +41-22-994 43 09 (for Europe, Israel, Australia and New Zealand); SOCAR Cardiac Research Ltd., 4120 Yonge Street, Suite 306, North York, Ontario M4N 2G3 Canada, phone +1-416-226 16 00, fax +1-416-226 44 40 (for Canada).

SPONSOR
Bayer AG, Leverkusen, Germany.

STEERING COMMITTEE
PA Poole-Wilson, London, UK (chairman); H Just, Freiburg, Germany (co-chairman); M. Motro, Tel-Hashomer, Israel (co-chairman); JD Parker, Toronto, Canada (co-chairman); MG Bourassa, Montreal, Canada; T Dart, Melbourne, Australia; J-M Detry, Brussels, Belgium; K Fox, Edinburgh, Scotland; P Hildebrandt, Frederiksberg, Denmark; A Hjalmarsson, Goteborg, Sweden; GP Molhoek, Enschede, the Netherlands, J-E Otterstad, Toensberg, Norway; P Rizzon, Bari, Italy; R Seabra-Gomes, Carnaxide, Portugal; J Soler-Soler, Barcelona, Spain; S Weber, Paris, France.

EXECUTIVE COMMITTEE
PA Poole-Wilson (chairman); H Just, M Motro, JD Parker, BA Kirwan, J Lubsen, FJ van Dalen, AB Parker, M Llewellyn.

DATA MONITORING AND ETHICAL REVIEW COMMITTEE
SJ Pocock, London, UK (chairman); J-P Boissel, Lyon, France; WW Parmley, St-Francisco, USA; W Rutishauser, Geneva, Switzerland; L Wilhelmsen, Gothenburg, Sweden.

CRITICAL EVENTS COMMITTEE
N Danchin, Vandoeuvre-les-Nancy, France (chairman), A Battler, Beer Sheva, Israel; A Bayes de Luna, Barcelona, Spain; P Dunselman, Breda, The Netherlands; S Glasser, Tampa, USA; P Koudstaal, Rotterdam, the Netherlands; G Sutton, Uxbridge, UK.

APPENDIX II

Calculation of power and confidence intervals

In the 2,221 patients allocated to simvastatin in the 4S study[24] the following events were observed: 182 deaths, 164 definite non-fatal acute myocardial infarctions, 12 intervention associated non-fatal acute myocardial infarctions and 47 non-fatal strokes. Overt heart failure requiring a change in heart failure treatment and refractory anginal chest pain were not specifically mentioned in the 4S report but are likely to have been incorporated in the 295 patients who had an ‘acute non-myocardial infarction coronary heart disease’ event in the simvastatin arm. Taken

together this represents 700 events that are included in the primary endpoint of ACTION.

The Kaplan-Meier estimate of the 5-year survival in the simvastatin arm was about 93% (cf. figure 1 of the original 4S report). Assuming that survival was exponential with a constant hazard (cf. equation 2 below), this corresponds with a hazard of death of *-[ln(0·93)]/0·05* or 1·45 per 100 patient-years. Based on this data, the hazard of the primary endpoint to be expected in the placebo arm of this study is therefore of the order of *700* x *1·45/182* or 5·58 per 100 patient-years. Similarly, the expected hazard of the endpoints monitored for safety (based on 182 deaths, 164 definite non-fatal acute myocardial infarctions, 12 intervention associated non-fatal acute myocardial infarctions and 47 non-fatal strokes or 405 events in 4S) is *405* x *145/182* or 3·23 per 100 patient-years.

Power calculations were based on:

1. the null-hypothesis *(H_o)* that there is no difference between allocation to nifedipine GITS and to placebo,
2. a two-sided significance test for the rejection of H_o with a level α of 0·05,
3. a planned sample size *n* of 3,000 per arm with a mean follow up by assigned treatment of five years.

To assess the probability that H_o will be rejected (i.e. the power of this study) as a function of the risk reduction by nifedipine GITS and of the event hazard in the placebo arm, the following standard formula for the calculation of sample sizes when proportions in two equal groups are compared was used:[31]

$$n = \frac{p_1 \cdot (1 - p_1) + p_0 \cdot (1 - p_0)}{(p_1 - p_0)^2} \times f(\alpha, \beta) \qquad (1)$$

In equation 1, p_1 and p_0 represent the event probabilities in the nifedipine GITS and placebo arms respectively. The power (1-β) and the significance level α are set by the standard Normal distribution function $f(\alpha,\beta)$, which has values of 13·0, 7·9 and 3·8 for a power of 95%, 80% and 50% respectively at a two-sided α of 0·05.

Event probabilities were converted to hazards by exponential approximation, assuming a constant hazard:[32]

$$p_i = 1 - e^{-h_i \cdot t} \text{ for } i = 1 \text{ and } i = 0 \qquad (2)$$

In this equation p_i is as defined above and h_i stands for the event hazard in a given treatment arm. As a mean observation time of 5 years is planned, *t* was set to 5 years.

For various choices of the event hazard h_o in the placebo arm, p_o was calculated (by equation 2 with t = 5 years). Next, equation 1 was solved for p_1 (with n = 3,000) for each of the three values of $f(\alpha,\beta)$ given above. Using again equation 2, the p_1 thus obtained was then converted to the corresponding hazard h_1 in the nifedipine GITS arm. Finally, the hazard ratio was calculated as h_1/h_0. Results of these calculations are shown in Figure 1.

To calculate the expected lower and upper bounds of two-sided 90% and 95% confidence intervals around an observed hazard ratio *(HR)* equal to one, the following formulas were used.[33] The standard error of *ln(HR)* was taken as:

$$se\left[\ln(HR)\right] = \sqrt{\frac{1}{e_1} + \frac{1}{e_0}}$$

where e_1 and e_0 stand for the number of events in the nifedipine GITS and placebo arms respectively.

The lower and upper bounds *LB* and *UB* of a confidence interval *(CI)* of *HR* follow from:

$$LB, UB = e^{\ln(HR) \pm z \cdot se\left[\ln(HR)\right]}$$

For various choices of equal event hazards in the nifedipine GITS and placebo arms (h_1 and h_o), p_1 and p_o were calculated (by equation (2) above, with t = 5 years). From these, the expected numbers of events e_1 and e_0 were obtained, based on n = 3,000 per arm. Lower and upper confidence bounds around a hazard ratio of one were then obtained by the formula above by taking z = 1·645 for two-sided 90% and z = 1·96 for 95% confidence intervals. Results of these calculations are shown in Figure 2.

APPENDIX III

Guideline for early termination

The total number of patient-years of observation planned is 6,000 x 5 or 30,000 patient-years. Interim analyses will be done each time an additional 5,000 patient-years of observation has been collected (i.e. five interim and one final analysis). The first one will be performed shortly before the end of patient recruitment.

The option to stop the trial because the combined hazard of death, acute myocardial infarction and stroke appears to be *increased* by nifedipine GITS will be considered at each of these interim analyses and the trial will be stopped if the negative effect of nifedipine GITS (relative to placebo) on this combined rate is significant at the level of p = 0·01. This implies that there is a probability of 0·025 ('Pocock rule'[34]) that the trial will be stopped for an adverse effect on the endpoints monitored for safety while in fact no such effect exists.

The option to stop the trial because the combined hazard of death, acute myocardial infarction and stroke appears to be *reduced* by nifedipine GITS will be considered only at the 2nd and the 4th interim analysis mentioned above, i.e. only after 10,000 and after 20,000 patient-years of observation have been collected. Based on a hazard in the placebo group of 3·23 per 100 patient-years, approximately 160 and 320 cases of death, non-fatal acute myocardial infarction or stroke respectively are expected to have accrued at these time points in the placebo group.

The guideline for stopping is:

first look: $z \geq 3{\cdot}47$ ($p < 0{\cdot}0005$),

second look: $z \geq 2{\cdot}45$ ($p < 0{\cdot}014$).

The calculation was performed with the EaSt™ package[35] to correct for 'multiple looks' in sample size calculation for survival data and was based on the following assumptions:

1. two-sided $\alpha = 0{\cdot}05$
2. a number of looks of three,
3. boundary shape as described by O'Brien and Fleming.[36]

It is noted that the *z*-values mentioned depend only on specific assumptions about α, the number of looks and the type of boundary chosen.

REFERENCES

1. Gersh BJ, Braunwald E, Rutherford JD. Chronic coronary artery disease. In: Braunwald E, ed. Heart Disease. A textbook of cardiovascular medicine. Fifth Edition. WB Saunders Company, Philadelphia, 1992:1289-365.
2. Fox KM, Davies GJ. Pathophysiology, investigation and treatment of chronic stable angina. In: Julian DG *et al*. eds. Diseases of the Heart. WB Saunders and Co, London, 1996:1000-26.
3. Robertson RM, Robertson D. Drugs used for the treatment of myocardial ischemia. Chapter 32 in: Goodman and Gilman's The Pharmacological Basis of Therapeutics, 9th edition, McGraw-Hill, New York, 1996:759-79.
4. Smith SC Jr, Blair SN, Criqui MH, *et al*. (the Secondary Prevention Panel). Preventing heart attack and death in patients with coronary disease. *Circulation* 1995;**92**:2-4.
5. Furberg CD, Psaty BM, Meyer JV. Nifedipine. Dose-related increase in mortality in patients with coronary heart disease. *Circulation* 1995;**92**:1326-31.
6. Furberg CD, Psaty BM. Corrections to the nifedipine meta-analysis. Letter to the Editor. *Circulation* 1996;**93**:1475-6.
7. Psaty BM, Heckbert SR, Koepsell TD, *et al*. The risk of myocardial infarction associated with antihypertensive drug therapies. *JAMA* 1995;**274**:620-5.
8. Buring JE, Glynn RJ, Hennekens CH. Calcium channel blockers and myocardial infarction. A hypothesis formulated but not yet tested. Editorial comment. *JAMA* 1995;**274**:654-5.
9. Frohlich ED, Zanchetti A. Antihypertensive therapy:safety and efficacy of drugs and publications. Editorial. *Hypertens* 1996;**27**:317-8.
10. Kloner RA. Nifedipine in ischemic heart disease. Editorial. *Circulation* 1995; **92**:1074-78.
11. Lenfant C. The calcium channel blocker scare. Lessons for the future. *Circulation* 1995;**91**:2855-6.
12. Mancia G, Van Zwieten PA. How safe are calcium antagonists in hypertension and coronary heart disease? *J Hypertens* 1996;**14**:13-7.

13. Marwick C. FDA gives calcium channel blockers clean bill of health but warns of short-acting nifedipine hazards. *JAMA* 1996;**275**:423.
14. Opie LH and Messerli FH. Nifedipine and mortality. Grave defects in the dossier. Editorial. *Circulation* 1995;**92**:1069-73.
15. Rydén L, Malmberg K. Calcium channel blockers or beta receptor antagonists for patients with ischaemic heart disease. What is the best choice? Editorial. *Eur Heart J* 1996;**17**:1-2.
16. Yusuf S. Calcium antagonists in coronary artery disease and hypertension. Time for reevaluation? Editorial. *Circulation* 1995;**92**:1079-82.
17. Brogden RN, McTavish D. Nifedipine gastrointestinal therapeutic system (GITS). A review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in hypertension and angina pectoris. *Drugs* 1995;**50**:495-512.
18. Sorkin EM, Clissold SP, Brogden RN. Nifedipine. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic efficacy, in ischaemic heart disease, hypertension and related cardiovascular disorders. *Drugs* 1985;**30**:182-274.
19. Krakoff LR, Bravo EL, Tuck ML, Friedman CP and the Modern Approach to the Treatment of Hypertension (MATH) Study Group. Nifedipine gastrointestinal therapeutic system in the treatment of hypertension. Results of a multicenter trial. *Am J Hypertens* 1990;**3**:318S-325S.
20. Parmley WW, Nesto RW, Singh BN, Deanfield J, Gottlieb SO and the N-CAP Study Group. Attenuation of the circadian patterns of myocardial ischemia with nifedipine GITS in patients with chronic stable angina. *J Am Coll Cardiol* 1992;**19**:1380-9.
21. Brown MJ, Castaigne A, Ruilope LM *et al*. INSIGHT: International nifedipine GITS study intervention as goal in hypertension treatment. *Journal of Human Hypertension* 1996;**10**(Suppl. 3):S157-60.
22. ICH Harmonised Tripartite Guideline for Good Clinical Practice. Brookwood Medical Publications, Brookwood, 1996.
23. Pyörälä K, De Backer G, Graham G, Poole-Wilson PA, Wood D on behalf of the Task Force. Prevention of coronary heart disease in clinical practice. Recommendations of the Task Force of the European Society of Cardiology, European Atherosclerosis Society and European Society of Hypertension. *Eur Heart J* 1994;**15**:1300-31.
24. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;**344**:1383-9.
25. Cox DR. Regression models and life tables. *J R Statist Soc B* 1972;**34**:187-200.
26. Olsson G, Lubsen J, Van Es GA, Rehnqvist N. Quality of life after myocardial infarction: effect of long term metoprolol on mortality and morbidity. *Br Med J* 1986;**292**:1491-3.

27. Bravo EL, Krakoff LR, Tuck ML, Friedman CP and the Modern Approach to the Treatment of Hypertension (MATH) Study Group. Antihypertensive effectiveness of nifedipine gastrointestinal therapeutic system in the elderly. *Am J Hypertens* 1990;**3**:326S-332S.
28. Staessen JA, Fagard R, Thijs L, *et al*. for the Systolic Hypertension-Europe (Syst-Eur) Trial Investigators. Morbidity and mortality in the placebo controlled European trial on isolated hypertension in the elderly. *Lancet* 1997; **350**:757-64.
29. Loaldi A, Polese A, Montorsi P, *et al*. Comparison of nifedipine, propranolol and isosorbide dinitrate on angiographic progression and regression of coronary arterial narrowings in angina pectoris. *Am J Cardiol* 1989;**64**:433-9.
30. Lichtlen PR, Hugenholtz PG, Rafflenbeul W, Hecker H, Jost S, Deckers JW on behalf of the INTACT group investigators. Retardation of angiographic progression of coronary artery disease by nifedipine. Results of the international nifedipine trial on antiatherosclerotic therapy (INTACT). *Lancet* 1990;**335**:1109-13.
31. Pocock SJ. Clinical trials: a practical approach. J Wiley & Sons, Chichester, 1983, page 125.
32. Rothman KJ. Modern epidemiology. Little Brown & Co, Boston, 1986a, page 31.
33. Rothman KJ. Modern epidemiology. Little Brown & Co, Boston, 1986b, page 171.
34. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977;**64**:191-9.
35. CYTEL software corporation. EaSt: a software package for the design and interim monitoring of group sequential clinical trials. Cambridge, 1993.
36. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;**35**:549-56.

Chapter 5

Treatment of angina pectoris: associations with symptom severity

ABSTRACT

Objective: To evaluate whether the frequency of anginal attacks in medically treated patients with stable angina is related to the intensity of anti-anginal treatment, the clinical history and coronary anatomy.

Methods: Analysis of baseline data from the ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) study, an ongoing placebo-controlled trial in 7,669 patients with stable angina pectoris who require anti-anginal treatment.

Results: Prior to randomisation, 8% of 7,669 patients had no anginal attacks, 63% had occasional, 22% had regular, 4% had frequent and 3% had daily attacks. Men (79% of all patients) and patients with a history of MI (51%) had less frequent anginal attacks ($P<0{\cdot}0001$). The number of coronary angiograms ever performed (70% had at least one angiogram), the extent of angiographic coronary disease (32% of those who had angiography had more than two-vessel disease), a history of peripheral artery disease (12%), the number of anti-anginal drugs used (64% were prescribed two or more such medications) and a history of revascularisation (a history of coronary bypass surgery was present in 23% and of balloon dilatation in 26%) were each positively associated with anginal attack frequency.

Conclusions: For the majority of patients with chronic stable angina not on a calcium-antagonist, medical treatment with other anti-anginal drugs is sufficient to control symptoms and only a minority of patients are refractory to medical treatment. Invasive treatments for chronic stable angina are only needed in a small proportion where symptoms persist.

Stable angina pectoris caused by ischaemic heart disease (IHD) is an easily recognised clinical entity. The frequency of anginal attacks is usually stabilised by the use of drugs tailored to the needs of the individual patient and the spectrum of treatment ranges from monotherapy to the use of multiple drugs and interventional procedures. Abnormalities found on coronary angiography vary greatly, even when only patients with typical angina pectoris without a history of myocardial infarction are considered.[1]

The effect on symptoms of anti-anginal drugs used in clinical practice has been extensively assessed by placebo-controlled clinical trials but there is little evidence for a positive effect on mortality. The benefit from coronary artery bypass grafting (CABG) compared to medical treatment in selected patients with regard to symptoms and mortality is generally accepted.[2] Percutaneous transluminal coronary angioplasty (PTCA) has been shown to reduce angina but not mortality.[3] Only limited information is available concerning to what extent treating physicians both attempt and succeed in reducing the frequency of anginal attacks, using therapies which have been proven in clinical trials.

One way to address this question would be to perform an open randomised trial in patients who present with anginal symptoms, comparing a stepped protocol of medical and invasive treatment with no treatment at all. For ethical reasons such a trial will never be done. An alternative way of answering this question would be to determine whether there is a relationship between the frequency of anginal attacks and the intensity of treatment in a cross-sectional study of patients with stable angina. If the degree of response to treatment varies between patients, one would expect such a study to show variability concerning the intensity of treatment amongst patients who report a similar frequency of anginal attacks. If some patients were more or less refractory to certain treatments, one would expect that patients who report frequent anginal attacks are treated more intensively than patients who have less frequent attacks.

ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) is an on-going multicentre, prospective, randomised, double blind, placebo controlled trial. Its primary objective is to assess the effect of nifedipine on cardiovascular event-free survival of ambulatory patients who otherwise receive optimal treatment for stable angina. Its design and methods have been described elsewhere.[4] Prior to randomisation detailed data on the clinical history and current treatment were collected. These data allowed us to assess the associations between the frequency of anginal attacks and variables such as the intensity of medical treatment, the coronary anatomy and clinical characteristics.

METHODS

Patients

ACTION focuses on the following three subgroups of patients with angina pectoris:

1. Those with a history of acute myocardial infarction who subsequently had angina or who developed angina after an angina-free period.
2. Those with a history of coronary revascularisation who either continued to have angina or who developed angina after an angina-free period.
3. Those with anginal complaints and a positive exercise test but without a history of myocardial infarction, coronary angiography or revascularisation.

The ACTION inclusion and exclusion criteria have been described in detail elsewhere.[4] Angina pectoris was defined as typical chest discomfort localised in the central part of the chest with or without radiation and elicited by physical or psychological stimuli. The symptoms were required to have been relieved gradually by rest or quickly by nitroglycerin. Included in the definition were breathlessness, fatigue, or dyspnoea that behaved in a similar manner.[5] Only patients in a stable clinical condition requiring oral and/or transdermal treatment either to treat, or to prevent, recurrent anginal attacks were eligible. It was not a requirement for patients to be symptomatic at the time of randomisation. It sufficed that anti-anginal treatment had been started in the past specifically to treat anginal attacks. Patients with clinical heart failure or an ejection fraction below 40% were excluded, as were patients with other important clinical conditions or contra-indications for nifedipine GITS.

ACTION patients are followed up with 3-monthly contacts for a minimum of four years. The mean follow up will be five years and results are expected in 2004.

Data collection

The pre-randomisation data used in the present analysis was collected using a paper case report form (CRF). Using a pre-coded question, anginal attack frequency was documented as no attacks, occasional (less than one attack/week), regular (1 – 3 attacks/week), frequent (4 – 6 attacks/week) or daily attacks. Functional class was rated using the New York Heart Association (NYHA) classification scale.[6] The total number of coronary angiograms (CAGs), PTCAs and CABGs ever performed, the clinically significant lesions ever identified, and whether lesions were dilated or bypassed, were documented by items in the CRF as shown in Figure 1. Information concerning the history of myocardial infarction, claudication, transient ischaemic attacks and debilitating stroke was obtained by specific questions, as was information concerning the presence of diabetes mellitus, smoking habits, hypertension and hyperlipidaemia treated with drugs. Details of treatments prescribed at the time of randomisation, body height and weight were also collected. Body mass index (BMI) was calculated according to the standard formula *(weight in kg)/(height in m)*^{2}.

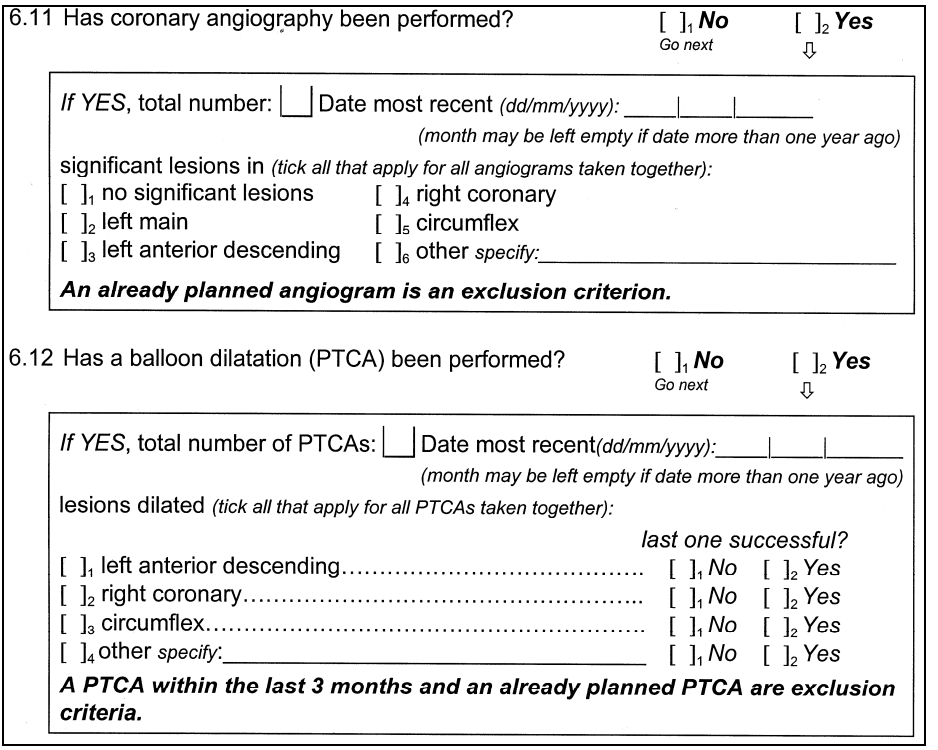


Figure 1: Items related to coronary angiography and to balloon dilatation (PTCA) as printed in the ACTION Case Report Form

Statistical methods

Proportions across frequency of anginal attack categories were compared by chi-squared test for trend. Calculations were done using least-squares linear regression, which gives identical results.[7] The dependent variable was always the frequency of anginal attacks, scored from 0 = none to 4 = daily. Dichotomous independent variables such as gender were scored as either 0 or 1. Least-squares linear regression analysis was also used to assess whether categorical variables were correlated with anginal attack frequency.

Categorical variables were scored from zero to (N – 1), where N is the number categories considered. The following variables were considered in this manner:
Number of coronary angiograms performed: 0 = none, 1 = one, 2 = more than one.
Number of vessels with clinically significant lesions for all coronary angiograms performed: The number of specifically mentioned arteries as shown in Figure 1 that were ticked by the investigator was counted. Left-main disease was considered as two-vessel disease.[1]

Table 1: Frequency of anginal attacks, NYHA class and history of cardiovascular disease

		Frequency of anginal attacks*					
	Total	None	Occasional	Regular	Frequent	Daily	
Total No. of patients	7,669	595 (8%)	4,869 (63%)	1,714 (22%)	290 (4%)	201 (3%)	*P*-value[#]
Age (years), median (range)	64 (35 – 89)	62 (35 – 88)	64 (35 – 88)	64 (35 – 89)	65 (36 – 85)	63 (39 – 89)	0·07
Male gender	6,085 (79%)	517 (87%)	3,901 (80%)	1,304 (76%)	219 (76%)	144 (72%)	<0·0001
In NYHA class I	4,137 (54%)	521 (88%)	2,825 (58%)	654 (38%)	85 (29%)	52 (26%)	<0·0001
History of MI	3,890 (51%)	307 (52%)	2,557 (53%)	802 (47%)	138 (48%)	86 (43%)	<0·0001
History of claudication	604 (8%)	41 (7%)	352 (7%)	158 (9%)	24 (8%)	29 (14%)	0·0001
History of TIA	296 (4%)	20 (3%)	180 (4%)	68 (4%)	16 (6%)	12 (6%)	0·04
History of debilitating stroke	132 (2%)	11 (2%)	81 (2%)	27 (2%)	10 (3%)	3 (1%)	0·5
Any peripheral artery disease[§]	938 (12%)	69 (12%)	559 (11%)	226 (13%)	41 (14%)	43 (21%)	<0·0001

Except for the top row, percentages are given for each category of attack frequency = 100%. NYHA, New York Heart Association; MI, myocardial infarction; TIA, transient ischaemic attack.

* Occasional, less than one attack/week; regular, 1 – 3 attacks/week; frequent, 4 – 6 attacks/week; daily, at least one attack every 24 hours.

[#] For test of trend across categories of attack frequency (c.f. *statistical methods*).

§ Claudication and/or TIA and/or debilitating stroke.

Outcome of previous PTCA: Only specifically mentioned arteries as shown in Figure 1 were considered if ticked by the investigator. The result was scored as follows: 0 = complete failure (all dilatations unsuccessful); 1 = partial success (at least one but not all dilatations successful); 2 = complete success (all dilatations successful). For both CAG and PTCA, significant lesions either seen or dilated that were documented in the section 'other' (c.f. figure 1) were ignored in this analysis.
Number of anti-anginal medications used: 0 = none, 1 = one, 2 = more than one. Only drugs registered for symptomatic relief of angina were counted as anti-anginal medication.

Multiple linear regression analysis was used to assess which variables were conditionally independent correlates of anginal attack frequency. Separate backward elimination procedures were used to determine which of the variables from Tables 1 – 4 respectively were statistically significant ($P<0{\cdot}05$) conditionally independent correlates. From Table 1, NYHA class was not considered. Indicator variables and the categorical variable *number of coronary angiograms performed* were entered in multivariate analyses coded as described earlier. To assess whether the number of lesions was an independent predictor, the categorical variable *number of coronary angiograms performed* was replaced by two indicator variables: one for history of angiography (coded as 0 = no, 1 = yes) and one for number of lesions (coded as 0 = no history of angiography, or no lesions, or one-vessel disease; 1 = left-main, two- or more than two-vessel disease). Similarly, history of PTCA and its result were represented by two indicator variables: one for history of PTCA (coded as 0 = no, 1 = yes) and one for result (coded as 0 = failure; 1 = partial or complete success as defined earlier). The variables thus identified from each table were then combined in one model. From this model, non-significant ($P>0{\cdot}05$) variables were removed. Calculations were performed using SAS®.

RESULTS

Study population

Between November 1996 and December 1998, 7,797 patients were randomised in 298 centres from 19 countries (see Appendix). Because of irregularities observed during on-site audits, 5 centres were closed and all 128 patients randomised in the centres concerned were removed from the database. Hence, this analysis concerns 7,669 patients who are currently participating in ACTION.

Anginal attack frequency and clinical history

Data concerning the frequency of anginal attacks in relation to clinical history at the time of randomisation are given in Table 1. Of 7,669 patients, 8% had no anginal attacks while 63% had occasional, 22% had regular, 4% had frequent and 3% had daily attacks. The median age was 64 years (range 35 – 89), and was not related to the frequency of angina.

Table 2: Frequency of anginal attacks and coronary angiography

		Frequency of anginal attacks*					
	Total	None	Occasional	Regular	Frequent	Daily	
History of coronary angiography (number of angiograms performed):							
Total No. of patients	7,668[§]	595 (8%)	4,869 (63%)	1,713 (22%)	290 (4%)	201 (3%)	*P*-value[#]
None	2,270 (30%)	219 (37%)	1,431 (29%)	481 (28%)	82 (28%)	57 (29%)	
One	3,229 (42%)	261 (44%)	2,093 (43%)	707 (41%)	105 (36%)	63 (31%)	<0·0001
> one	2,169 (28%)	115 (19%)	1,345 (28%)	525 (31%)	103 (36%)	81 (40%)	
Extent of coronary disease in patients with at least one coronary angiogram:							
Total No. of patients	5,398	376 (7%)	3,438 (64%)	1,232 (23%)	208 (4%)	144 (3%)	
No lesions	240 (4%)	19 (5%)	155 (5%)	50 (4%)	9 (4%)	7 (5%)	
1 VD	1,822 (34%)	145 (38%)	1,174 (34%)	400 (33%)	61 (29%)	42 (29%)	0·001
2 VD or LM	1,632 (30%)	123 (33%)	1,030 (30%)	370 (30%)	61 (29%)	48 (33%)	
>2 VD	1,704	89 (24%)	1,079	412 (33%)	77 (38%)	47 (33%)	

	(32%)		(31%)				

Except for the top rows, percentages are given for each category of attack frequency = 100%. VD, vessel disease; left-main (LM) disease was considered as 2 VD, c.f. *statistical methods.*

* Occasional, less than one attack/week; regular, 1 – 3 attacks/week; frequent, 4 – 6 attacks/week; daily, at least one attack every 24 hours.

[#] For test of trend across categories of attack frequency (c.f. *statistical methods*).

[§] History of coronary angiography unknown in one patient.

The percentage of males decreased across attack frequency categories from 87% for patients with no attacks to 72% for patients with daily attacks (*P* 0·0001). The percentage of patients in NYHA class I declined from 88% for patients with no anginal attacks to 26% for patients with daily attacks (*P*<0·0001). The percentage of patients with a history of myocardial infarction changed from 52% for patients with no anginal attacks to 43% for patients with daily attacks (*P*<0·0001). The percentage of patients with a history of claudication, transient ischaemic attacks or debilitating stroke increased from 12% for patients with no anginal attacks to 21% for patients with daily attacks (*P*<0·0001). Data for a history of claudication, transient ischaemic attacks and debilitating stroke separately is given in Table 1 also.

At least one coronary angiogram had been performed in 70% of all patients for whom this information was available (Table 2). The percentage of patients who in the past never had coronary angiography decreased with increasing attack frequency from 37% for patients without anginal attacks to 29% of patients with daily attacks. Conversely, the percentage of patients who had more than one coronary angiogram increased from 19% for those without anginal attacks to 40% for those with daily attacks (*P*<0·0001). Patients with a higher anginal attack frequency had more severe angiographic coronary disease. The percentage of patients with no lesions or one-vessel disease decreased from 43% (19 + 145 of 376) for patients without anginal attacks to 34% (7 + 42 of 144) for patients with daily attacks. The percentage of patients with more than two-vessel disease increased from 24% for patients without anginal attacks to 38% for patients with frequent and to 33% for patients with daily attacks (*P*=0·001, Table 2).

Anginal attack frequency and treatment

Data on current medical and past invasive treatment are given in Table 3. The number of anti-anginal drugs prescribed increased with the frequency of anginal attacks. Overall, one % of patients was not on any anti-anginal medication at the time of randomisation. The percentage of patients prescribed one anti-anginal drug or less decreased gradually from 48% (9 + 277 of 595) for patients without anginal attacks to 14% (5 + 23 of 201) for patients with daily attacks. Conversely, the percentage of patients prescribed two or more anti-anginal drugs increased from 52% for patients without anginal attacks to 86% for patients with daily attacks (*P*<0·0001, Table 3).

Overall, 23% of patients had a history of CABG and 26% of PTCA (Table 3). Previous CABG was less frequent for patients without anginal attacks (15%) than for those with daily attacks (28%, p < 0·0001). The same applied to a history of PTCA (20% for patients without and 28% for patients with daily anginal attacks, p = 0·03).

Table 3: Frequency of anginal attacks and treatment

		Frequency of anginal attacks*					
	Total	None	Occasional	Regular	Frequent	Daily	
Total No. of patients	7,669	595 (8%)	4,869 (63%)	1,714 (22%)	290 (4%)	201 (3%)	*P*-value[#]
Anti-anginal medication:							
None	107 (1%)	9 (2%)	77 (2%)	16 (1%)	-	5 (2%)	
On one anti-anginal Rx	2,659 (35%)	277 (46%)	1,813 (37%)	484 (28%)	62 (21%)	23 (12%)	<0·0001
On ≥ 2 anti-anginal Rx	4,903 (64%)	309 (52%)	2,979 (61%)	1,214 (71%)	228 (79%)	173 (86%)	
History of CABG	1,791 (23%)	88 (15%)	1,141 (23%)	435 (25%)	70 (24%)	57 (28%)	<0·0001
History of PTCA	2,016 (26%)	120 (20%)	1,294 (27%)	469 (27%)	77 (27%)	56 (28%)	0·03
History CABG and PTCA	378 (5%)	13 (2%)	218 (4%)	112 (7%)	16 (6%)	19 (9%)	<0·0001
Result of PTCA in patients with a history of PTCA:							
Total No. of patients	1,996[§]	116 (6%)	1,283 (64%)	466 (23%)	77 (4%)	54 (3%)	
Failure°	220 (11%)	10 (9%)	129 (10%)	64 (14%)	10 (13%)	7 (13%)	
Partial success°	67 (3%)	2 (2%)	46 (4%)	13 (3%)	1 (1%)	5 (9%)	0·02

Complete success°	1,709 (86%)	104 (89%)	1,108 (86%)	389 (83%)	66 (86%)	42 (78%)	

Except for the top rows, percentages are given for each category of attack frequency = 100%. CABG, coronary artery bypass grafting; PTCA, percutaneous transluminal coronary angioplasty.

* Occasional, less than one attack/week; regular, 1 – 3 attacks/week; frequent, 4 – 6 attacks/week; daily, at least one attack every 24 hours.

[#] For test of trend across categories of attack frequency (c.f. *statistical methods*).

[§] Result of PTCA unknown in 20 patients.

° Failure, no successful dilatation(s); partial success, one or more, but not all dilatation(s) successful; complete success, all dilatation(s) successful.

Table 4: Frequency of anginal attacks and risk factors for cardiovascular disease

		Frequency of anginal attacks*					
	Total	None	Occasional	Regular	Frequent	Daily	
Total No. of patients	7,669	595 (8%)	4,869 (63%)	1,714 (22%)	290 (4%)	201 (3%)	*P*-value[#]
BMI (kg/m^2), median (range) [§]	27 (15 – 69)	27 (19 – 40)	27 (15 – 69)	27 (16 – 57)	27 (17 – 38)	27 (19 – 39)	0·08
Hyperlipidaemia°	5,336 (70%)	407 (68%)	3,347 (69%)	1,241 (72%)	202 (70%)	139 (69%)	0·09
Hypertension°	3,025 (39%)	243 (41%)	1,885 (39%)	700 (41%)	118 (41%)	79 (39%)	0·5
Current smoker	1,345 (18%)	97 (16%)	876 (18%)	283 (17%)	51 (18%)	38 (19%)	0·9
Diabetes	1,094	77 (13%)	669 (14%)	257 (15%)	58 (20%)	33 (16%)	0·005

	(14%)						
Diabetes treated with insulin	173 (2%)	8 (1%)	102 (2%)	49 (3%)	8 (3%)	6 (3%)	0·02

Except for the top row, percentages are given for each category of attack frequency = 100%. BMI, body mass index.

* Occasional, less than one attack/week; regular, 1 – 3 attacks/week; frequent, 4 – 6 attacks/week; daily, at least one attack every 24 hours.

[#] For test of trend across categories of attack frequency (c.f. *statistical methods*).

[§] Either body height or weight not recorded for 31 patients at baseline

° Treated with drugs.

Only 5% of patients had a history of both CABG and PTCA. The percentage of patients with such a history increased gradually across attack frequency categories from 2% for patients without attacks to 9% for patients with daily attacks ($P<0{\cdot}0001$). The success rates of PTCA were not related in a gradual manner to anginal attack frequency although the percentage of patients with complete success of PTCA was lowest (78%, Table 3) for patients with daily attacks.

Data on standard risk factors for cardiovascular disease are given in Table 4. The only risk factor that showed a statistically significant trend towards a higher prevalence with increasing anginal attack frequency was diabetes. There were 13% diabetics amongst patients without anginal attacks while the corresponding percentages for patients with frequent and daily attacks were 20% and 16% respectively ($P=0{\cdot}005$).

Multivariate analysis showed that the following variables were statistically significant ($P<0{\cdot}05$) conditionally independent correlates of attack frequency: gender, history of claudication, myocardial infarction and CABG; presence of diabetes, the number of anti-anginal drugs used and the number of coronary angiograms ever performed. As expected based on the results of univariate analyses, male gender and prior history of myocardial infarction were negatively associated with attack frequency also in multivariate analyses, while the other variables were positively associated.

DISCUSSION

The present report is based on pre-randomisation data from an ongoing trial in patients with stable angina pectoris (ACTION) who were treated with other anti-anginal drugs than calcium-antagonists at the time of enrolment. Such data are cross-sectional in nature and reflect the spectrum of similar patients seen in clinical practice, rather than the natural history of disease.

The main finding for the ACTION cohort is that anginal attacks at baseline were well controlled in the majority of patients. Eight % of patients were free of attacks and a further 63% reported less than one attack per week. Any CABG had been performed in 23% and any PTCA in 26% while 5% had a history of both procedures. Although the selection criteria of ACTION define a broad spectrum of patients with stable angina,[4] our findings are limited to patients who are eligible for this study. Patients with clinical heart failure and patients who were prescribed certain drugs that could not be stopped for clinical reasons were excluded. These included calcium-antagonists, anti-arrhythmic agents other than sotalol, and amiodarone. The protocol required that calcium antagonists were stopped prior to randomisation. This explains why overall 1% patients was not on any anti-anginal treatment at the moment of randomisation. Because of these and other restrictions that are inherent in selecting patients for participation in a clinical trial, the present analysis does not necessarily reflect the fraction of patients with a history of angina who in clinical practice remain symptomatic despite treatment.

All patients in this cohort had a medical history of exertional angina requiring treatment but it was not required that patients were symptomatic. Investigators, after interviewing the patient, scored the current anginal attack frequency using a pre-defined scale. Responses to the WHO Rose questionnaire[8] completed by the patient are known to differ from what is stated in medical records.[9] Absolute attack frequencies as observed in this study can therefore not be compared to results of studies using other methods for data collection.

It is a limitation of cross-sectional data that both investigators and patients may be more likely to remember specific past events when the patient experiences daily, or almost daily, anginal attacks. For instance, a patient with daily anginal attacks may be more likely to remember that he/she had an MI in the past than a patient without attacks. Such a recollection bias would result in a higher frequency of a history of MI in patients with daily attacks than in patients without attacks. The opposite was observed.

In the present analysis, the percentage of patients classified as NYHA class I decreased progressively across categories from 88% of patients without anginal attacks to 26% of patients with daily attacks (c.f. Table 1). The NYHA classification is a four-point-scale for the assessment by the physician of a patient's degree of disability due to cardiovascular disease. The scale ranges from class I = no limitation of physical activity, to class IV = any physical activity causes discomfort.[6] That patients without anginal attacks are not always assessed as being in NYHA class I is not surprising as patients who must limit physical activity to avoid angina will be classified in NYHA class II or higher. That 27% of patients with daily anginal attacks were nonetheless classified as in NYHA class I suggests that patients who report daily anginal attacks are not necessarily hindered by these attacks in their ordinary activities, at least not in the opinion of their treating physician. The differences in anginal attack frequency observed in this study may therefore in part be due to a more active lifestyle amongst patients with more frequent attacks.

Twenty-nine percent of patients had at least one anginal attack per week (c.f. Table 1). This may suggest that in current clinical practice not all patients are treated to the extent that they become asymptomatic. Other than that none of the patients were using a calcium-antagonist, there may be several explanations for this. First, patients or even their doctors may not make an attempt to minimise symptoms as patients may chose to live within their symptoms. Second, the side effects of anti-anginal drugs may be so bothersome that a proportion of patients prefer not to be treated. Third, lack of compliance may be an issue, as some patients may not use their medication(s) as prescribed. Fourth, under dosing when multiple drugs are used may be a factor and, finally, in some patients the symptoms may be refractory to all currently available treatments.

We believe that, in part, our data reflect the latter. As shown in Tables 3 and 4, patients with more frequent anginal attacks had coronary angiograms and revascularisation procedures performed more often, had more extensive coronary artery disease and were prescribed more anti-anginal drugs. In our data there was no indi-

cation that lower dosages of these were used when multiple anti-anginal drugs were prescribed, or when anginal attacks were more frequent. The number of coronary angiograms ever performed, a history of CABG and the number of anti-anginal drugs prescribed were independently associated with attack frequency in multivariate analysis. These relationships indicate that the intensity of treatment (drugs and/or revascularisation) is at least in part symptom-driven and are compatible with the view that some patients are refractory to all currently available treatments.

As the attack frequency increases, more anti-anginal drugs are prescribed. The same applies to revascularisation. Despite this, a fraction of patients remains symptomatic. In the analysis only treatments with proven anti-anginal efficacy were considered. Thus, the association between the intensity of treatment and the frequency of anginal attacks cannot be explained by the use of ineffective treatments. Rather, the conclusion must be that in clinical practice it is apparently sometimes difficult or even impossible to achieve complete control of anginal attacks with currently available treatments.

The association between the number of drugs prescribed and the frequency of anginal attacks may in part be due to failure of revascularisation to achieve lasting symptomatic relief. This is supported by the strong association between the percentage of patients who had both CABG and PTCA and attack frequency, and by the lower partial or complete success rates reported for patients with more frequent anginal attacks who had PTCA (c.f. Table 3). At the same time, this observation reinforces the conclusion that currently available treatment has limitations.

Our analysis of treatment does not include calcium-antagonists, a class of agents with proven anti-anginal efficacy. To comply with the ACTION selection criteria, a calcium-antagonist had been withdrawn in only 1% of patients before data on anginal attack frequency was collected. This percentage did not depend on the frequency of anginal attacks (c.f. Table 3). It is therefore unlikely that withdrawal of a calcium-antagonist for the purpose of ACTION had an effect on the frequency of anginal symptoms. Nonetheless, the question whether a cross-sectional survey that also covers patients on calcium-antagonists would have given different results cannot be answered.

Data on age, gender and clinical history are given in Table 1 and on risk factors for cardiovascular disease in Table 4. The small p-values are both a reflection of the differences observed and the large number of patients in this analysis. We have no explanation why women experience more frequent attacks than men. Krogh *et al.*[10] suggested that the threshold for perceiving or reporting angina is lower amongst women than amongst males. Non-anginal chest pain is more frequent in women than in men. Whether this has contributed to the present result cannot be determined. It is unlikely that the relationship between anginal attack frequency and the number of anti-anginal medications and the number of coronary angiograms ever performed can be explained by the relationship with gender. All three were conditionally independent predictors of anginal attack frequency in multivariate analysis. Similarly, we have no explanation for the decreasing fraction of

patients with a history of myocardial infarction as the frequency of anginal attacks increases. One contributing factor may be that myocardial infarction increases the angina threshold, as there may be less residual ischaemia in post-MI patients.

CONCLUSIONS

In the majority of patients with chronic stable angina pectoris not on a calcium-antagonist, who may have had a revascularisation procedure in the past, anginal attacks can be symptomatically controlled by currently available medical treatment with other anti-anginal drugs. Invasive treatments are only needed in a small proportion where symptoms persist. A minority of patients remain symptomatic despite multiple therapies. The findings support the view that a minority of patients with chronic stable angina are refractory to treatment.

ACKNOWLEDGEMENT

Guided by an independent Steering Committee, the ACTION trial is executed by an independent clinical research organisation (SOCAR Research SA). The role of the sponsor (Bayer AG, Leverkusen, Germany) is limited to supply and packaging of nifedipine GITS or matching placebo, and to on-site monitoring. Members of the Steering and the Data Monitoring and Ethical Review committees have all declared absence of conflicts of interest before the study started.

The present analysis was performed by B-A. Kirwan and J. Lubsen. B-A. Kirwan is ACTION project coordinator and J. Lubsen is the responsible biostatistician. P.A. Poole-Wilson is chairman of the ACTION Research group.

APPENDIX

ACTION participants are:

PRINCIPAL INVESTIGATOR AND STUDY CHAIRMAN
PA Poole-Wilson, Cardiac Medicine, National Heart & Lung Institute, Dove House Street, London SW3 6LY, UK.

INVESTIGATORS
293 centres in Australia (5), Austria (2), Belgium (6), Canada (50), Denmark (24), Finland (4), France(14), Germany(25), Greece(5), Israel(18), Italy(36), The Netherlands(27), New Zealand(3), Norway(16), Portugal(7), Spain(19), Sweden(16), Switzerland(3), UK(13).

COORDINATION AND MANAGEMENT
SOCAR Research SA, PO Box 2564, 1260 Nyon 2, Switzerland.

SPONSOR
Bayer AG, Leverkusen, Germany.

STEERING COMMITTEE
PA Poole-Wilson, London, UK (chairman); H Just, Freiburg, Germany (co-chairman); M. Motro, Tel-Hashomer, Israel (co-chairman); JD Parker, Toronto, Canada (co-chairman); MG Bourassa, Montreal, Canada; T Dart, Melbourne, Australia; J-M Detry, Brussels, Belgium; K Fox, Edinburgh, Scotland; P Hildebrandt, Frederiksberg, Denmark; A Hjalmarsson, Goteborg, Sweden; JA Kragten,

Heerlen, the Netherlands; GP Molhoek, Enschede, the Netherlands, J-E Otterstad, Toensberg, Norway; P Rizzon, Bari, Italy; R Seabra-Gomes, Carnaxide, Portugal; J Soler-Soler, Barcelona, Spain; S Weber, Paris, France.

EXECUTIVE COMMITTEE
PA Poole-Wilson (chairman); H Just, M Motro, JD Parker, BA Kirwan, J Lubsen, FJ van Dalen, M Llewellyn.

DATA MONITORING AND ETHICAL REVIEW COMMITTEE
SJ Pocock, London, UK (chairman); J-P Boissel, Lyon, France; WW Parmley, St-Francisco, USA; W Rutishauser, Geneva, Switzerland; L Wilhelmsen, Gothenburg, Sweden.

CRITICAL EVENTS COMMITTEE
N Danchin, Paris, France (chairman), A Battler, Beer Sheva, Israel; A Bayes de Luna, Barcelona, Spain; P Dunselman, Breda, The Netherlands; S Glasser, Tampa, USA; P Koudstaal, Rotterdam, the Netherlands; G Sutton, Uxbridge, UK.

REFERENCES

1. Chaitman BR, Bourassa MG, Davis K *et al*. Angiographic prevalence of high-risk coronary artery disease in patient subsets (CASS). *Circulation* 1981;**64**: 360-7
2. Eagle KA, Guyton RA, Davidoff R *et al*. ACC/AHA guidelines for coronary artery bypass graft surgery: executive summary and recommendations. A report of the American College of Cardiology/American Heart Association task force on practice guidelines (Committee to revise the 1991 guidelines for coronary artery bypass graft surgery). *Circulation* 1999;**100**:1464-80
3. RITA-2 trial participants. Coronary angioplasty versus medical therapy for angina: the second Randomised Intervention Treatment of Angina (RITA-2) trial. *Lancet* 1997;**350**:461-8
4. Lubsen J, Poole-Wilson PA, Pocock SJ, van Dalen FJ, Baumann J, Kirwan B-A, Parker AB. Design and current status of ACTION: A Coronary disease Trial Investigating Outcome with Nifedipine GITS. *Eur Heart J* 1998:**19** (Suppl 1):120-132
5. Rehnqvist N, Hjemdahl P, Billing E, *et al*. Effects of metoprolol vs. verapamil in patients with stable angina pectoris: the angina prognosis study in Stockholm (APSIS). *Eur Heart J* 1996;**17**:76-81
6. The Criteria Committee of the New York Heart Association. Nomenclature and criteria for diagnosis. 9th ed. Boston; Little, Brown 1994
7. Altman DG. Practical statistics for medical research. London, Chapman & Hall 1991, p. 319
8. Rose G. The diagnosis of ischaemic heart pain and intermittent claudication in field surveys. *Bull World Health Organ* 1962;**27**:645-59
9. Lawlor DA, Adamson J, Ebrahim S. Performance of the WHO Rose angina questionnaire in post-menopausal women: are all of the questions necessary? *J Epidemiol Community Health* 2003;**57**:538-41

10. Krogh V, Trevisan M, Panico S, *et al*. Prevalence and correlates of angina pectoris in the Italian nine communities study. Research group ATS-RF2 of the Italian National Research Council. *Epidemiology* 1991; **2**:26-32

Chapter 6

SOCDAT®: a comprehensive clinical trial data and study management philosophy based on simultaneous display of scanned documents and corresponding database content

ABSTRACT

We describe a modular database and study management philosophy called SOCDAT® (Source Consultation Data Acquisition Tool) designed for clinical trials that use pre-printed case report forms (CRFs) for data capture. All CRFs and supporting documents are registered in the database upon receipt, scanned and then archived. Further database management is entirely based on scanned images. Scanned images and corresponding database content are always simultaneously displayed on the user's computer screen. The content of the database is therefore continuously validated each time the database is consulted. A standard relational database is used (ORACLE™). Tables contain links to the relevant document images, which are stored by page as separate files. Data such as names of investigators, local normal ranges, coding dictionaries, etc. are also stored within the same database. This allows applications to access both patient- and centre-specific data. The system allows for flexible implementation of applications for data checking and clarification, and for the generation of study management tools such as status reports, listings of outstanding documents, etc. Apart from continuous validation during database consultation, the system has several other advantages. Once scanned, original documents can be permanently archived. Scanning guarantees legibility of documents that fade over time, and images can be incorporated automatically into reports. Double data entry and data checking can be done simultaneously at different locations. The scanned documents and the database can be stored together on removable mass storage media, which is an extra safeguard against loss. Database content and images can also be copied to a portable computer, which facilitates on-site monitoring and audits. The system runs on any platform supported by Oracle.

Over the past decade, the tools used to collect, manage and analyse data in clinical trials have gone through revolutionary changes. Although alternatives such as data fax with or without optical character recognition, remote data entry and internet-based data management systems have become available, acceptance and implementation of such methods has been slow.[1] Hence, many clinical trials today still rely on pre-printed paper Case Report Forms (CRFs) for data capture.

Any large multi-centre trial needs to have a multi-functional and efficient data management system in place that meets international Good Clinical Practice (GCP) requirements[2] and allows for customisation while the trial is ongoing. To meet these requirements for trials that rely either in part or completely on paper documents to capture patient data, we designed a 'paperless' concurrent data and document management system called SOCDAT® (Source Consultation Data Acquisition Tool). The key features of this system are that all incoming completed CRFs and supporting documents, such as electrocardiograms, laboratory test print-outs, discharge letters, etc. are registered in the database upon receipt, then scanned and archived. All subsequent data processing activities (data entry, data cleaning and review, processing of adverse events, coding of terms, etc.) are handled via computer screens that simultaneously display the scanned image of patient data as completed by the investigator, and the corresponding content of the database. SOCDAT's system architecture is based on a standard relational database. Other data relevant to the conduct of a specific study (investigator names and addresses, local laboratory normal ranges, dictionaries to code medical terms, etc.) can be incorporated into the same database. The system can thus support the varied logistic and study management functions required to execute clinical trials efficiently. Below we describe the system's architecture and functionality, using examples from the ongoing ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) study, of which the design and methods have been described elsewhere.[3] ACTION is a large randomised trial with fatal and non-fatal clinical endpoints in almost 8,000 patients from 299 centres in 19 countries. Patients were recruited between November 1996 and December 1998. The planned mean follow up is five years and main results are expected mid 2004.

SYSTEM DESIGN AND IMPLEMENTATION

The basic design considerations for the SOCDAT system were:

1. Integration of concurrent database and document management based on scanned images of completed CRF pages and supporting documents simultaneously displayed with the corresponding database content.
2. The use of only one standard software package that runs on a standard local area network (LAN) consisting of standard personal computers or workstations.
3. User access to all functionalities required for the co-ordination and management of a study through one start-up screen.
4. Possibility to port the system to other network and/or operating system environments if necessary.

5. Modular architecture that allows for flexible addition of functions as the need arises.

The core of SOCDAT is a standard relational database model based on tables containing information which can be grouped together logically. Each record contains creator and time stamp identification, which is updated whenever the record is changed. Relations between tables specified during programming are maintained by the system. Names of investigators and co-workers, allocated patient identification and medication numbers, local laboratory normal ranges, etc. are stored by centre within the same system. Hence, centre-specific data is directly accessible to all applications, and can be linked to patient data whenever required.

The user interface consists of one start-up screen that allows access to all functions available for any one study (c.f. Figure 1).

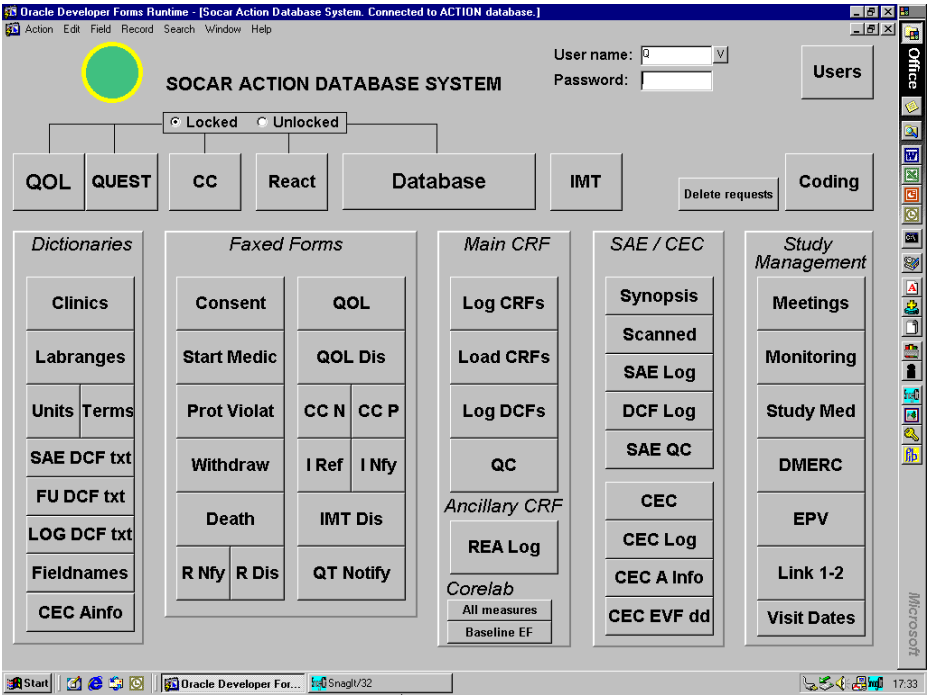


Figure 1: Start-up screen for the ACTION study

Each button allows access to a specific (set of) function(s) related to trial co-ordination and management. Buttons with related functions are grouped together.

The system is developed and implemented using standard ORACLE™ software on a low-end local area network of standard personal computers, using Microsoft Windows NT™ as operating system, but will run without changes on any platform that is supported by Oracle (Linux, Sun Solaris, Novell, HP-UX, etc.). Because of Oracle's client-server architecture, network traffic is kept to a minimum. So as a consequence, system response times both for forms and for scanned images are acceptable (less than 1 sec.) even on low-end standard 10 megabit/sec networks and with large numbers of images (>100 gigabytes).

Documents are scanned using standard scanning equipment. We found that the best results are obtained using the compressed *.tif image file format. In this format, the image size ranges from 30 KB for CRF pages to 500 KB for standard 12-lead electrocardiograms. Document images are stored by page as separate *.tif files. Data tables contain links to the document images that support the data in the table. The total disk space required to store the database for a trial as large as ACTION, including all scanned CRF pages and other documents, is 100 gigabytes. This is well within the current limits of PC technology. Because of the necessity to display image files, a 200 MHz Pentium™ PC or comparable MacIntosh™ is a minimum requirement for network client machines. Table 1 gives details on the projected magnitude of the database, including scanned documents, when the ACTION study is completed.

Table 1: Projected size of the ACTION database at study completion

Total number of patients	~8,000
Number of tables containing patient data	72
Mean number of variables per table, patient data	28
Total number of records, patient data	~2,000,000
Number of tables containing other data	76
Mean number of variables per table, other data	12
Total number of records, other data	~3,500,000
Number of scanned pre-printed pages	~1,500,000
Number of scanned supporting document pages	~500,000
Total size of the database	~100 Gigabytes

Initially, the system was implemented on one Local Area Network (LAN). More than 30 users within one building are connected simultaneously today. A subsidiary location in another country has direct access to the main database through a 256 Kbits/sec data line. To ensure an acceptable response time for images at the subsidiary location, scanned images are present at both locations, and are synchronised between the two systems each night.

APPLICATION MODULES

Below we summarise the various applications that can be distinguished within the SOCDAT system.

Registering, scanning and tracking of documents

Specially designed screens allow the user to register and track each data handling and processing step for patient-related documents. An example is shown in Figure 2.

Once registered by date of receipt and type of document, all patient-related documents are scanned. The start-up screen (c.f. Figure 1) contains buttons that allows user access to the scanning function for each type of document to be scanned.

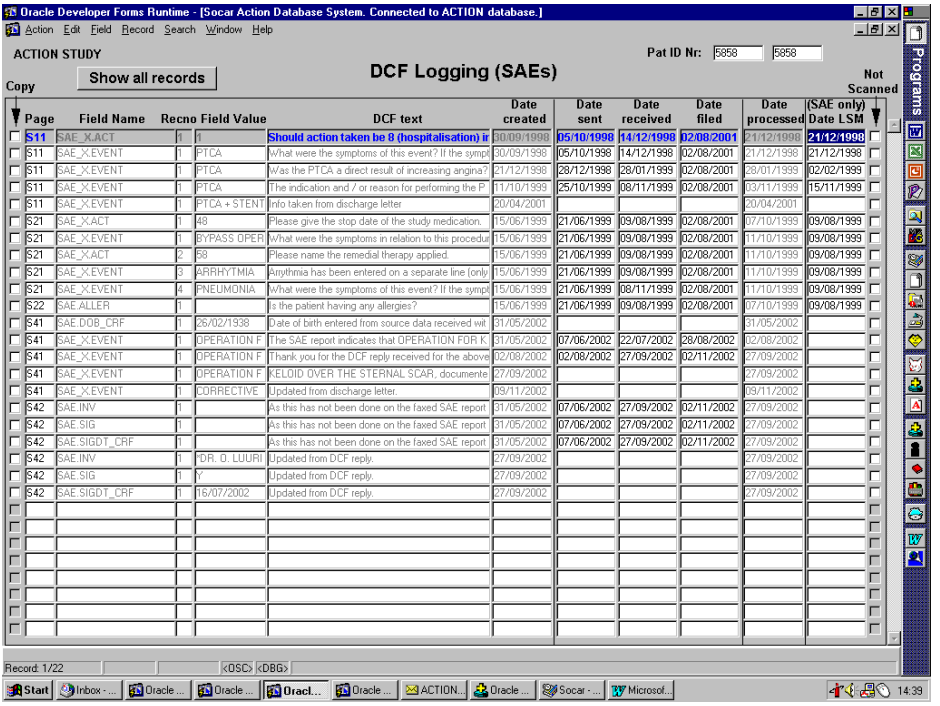


Figure 2: Typical tabular screen view used for tracking document processing

Documents that are generated from the database, such as Data Clarification Forms (DCFs), are tracked from date created to date filed. Tabular screens are patient-specific.

Data entry

The scanned image of the data to be entered appears on the left side of the screen, with a corresponding empty data entry window on the right. Data entry is performed manually using the scanned images. A completed data entry screen is shown in Figure 3.

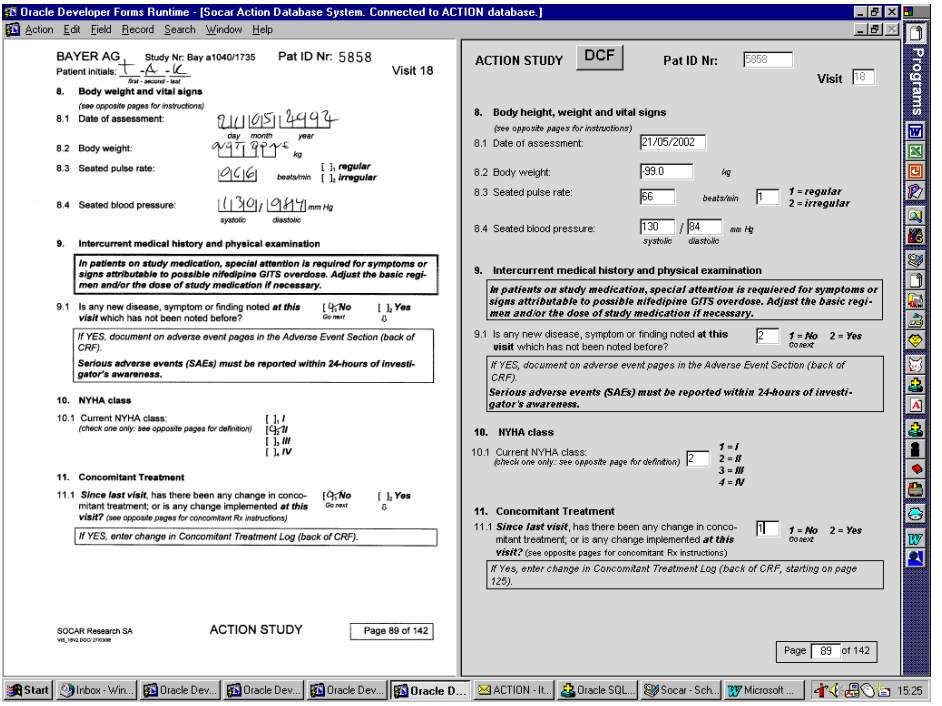


Figure 3: Screen layout for data entry and data consultation

For data entry, the scanned image of the CRF page to be entered appears on the left and a corresponding empty data entry template on the right. Repeated pressing of the *Page Down* key gives access to other scanned documents that support the database content shown on the right (such as a completed DCF, a laboratory report, an electrocardiogram, etc.), which are shown in sequence on the left side of the screen. Note that the second part of item 8.3 on the CRF page shown in the figure was not answered by the investigator. In addition, there is a discrepancy between image and database content for item 9.1. The Page Down key and the DCF button allow the user to consult the audit trail that explains such discrepancies (c.f. Figure 4 *next page*). Data entry of free text is facilitated by a dictionary of pre-defined terms, which is accessible via a specific button attached to the appropriate field. The latter prevents typing errors and ensures consistent data entry of similar terms. The user can however enter any text string.

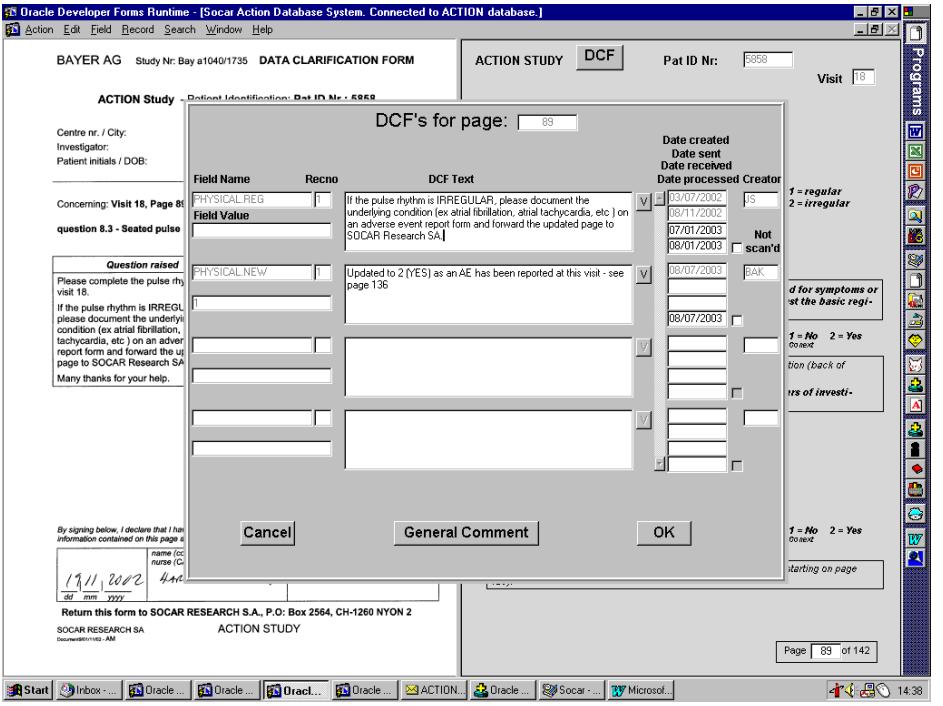


Figure 4: Data clarification system as implemented for ACTION

The same database content page as illustrated in Figure 3 is shown. After pressing *Page Down*, the DCF completed by the investigator appears on the left side and explains the discrepancy for item 8.3 (c.f. Figure 3). The DCF sub-screen shown in the figure is accessed via the DCF button. The first block shows the query text for item 8.3. The second block justifies the discrepancy for item 9.1. The DCF sub-screen can be opened at any time during data entry or consultation, and the user can enter data clarification queries as required. When the cursor is put in any field in the database content page, the field name appears to the left in the sub-screen when the DCF button is clicked. The DCF button remains red until all outstanding DCFs for a given page have been processed.

Data checking and quality control

SOCDAT allows for the implementation of any combination of data checking and quality control procedures that are compliant with GCP requirements.[2] Pre-defined checks to detect inappropriate character formats are automatically performed on initial data entry. Access to scanned images allows double data entry to be done concurrently on different workstations. Data checking programs created in Standard Query Language (SQL) are run regularly to check for completeness and plausibility. Based on the SQL output, the scanned images are used to make the appropriate changes in the database. Specific screens continuously track the checking of each data module. Standard locking procedures for checked data are implemented.

Data clarification

To query missing, incorrect or non-plausible data, Data Clarification Forms (DCFs) are generated from the database and sent to investigators for completion. Data clarification queries are field-specific and directly linked to the item that needs to be clarified. Figure 4 shows a screen view of the DCF system in SOCDAT. One DCF is generated for each data item to be clarified. The data clarification query can be inserted automatically by a data-checking program, or typed in manually. A dictionary of pre-defined queries for the data field concerned can be made available but the user can enter any text string. DCFs are tracked from date of creation to date filed, and are printed using a merge-file that is imported into standard word-processing software. When the DCF is printed, patient and investigator identification is added automatically (c.f. Figure 4).

The DCF system is also used to attach to any field a comment that explains a difference between scanned image and corresponding database content in the absence of a DCF that was completed by the investigator.

Coding of medical terms and drug names

The dictionaries used to standardise medical terms and drug names are stored as separate tables in the database. The initial code for any given term is selected manually from the coding dictionary. Once coded, if the same term is repeated in the database, the code is then automatically assigned. Each time a term is changed in the database, the code that was assigned is removed automatically. In order to ensure accurate and consistent coding, distinct coded terms are validated by an independent coder using a screen specifically designed for this purpose.

Data consultation

Within SOCDAT, different ways of consulting database content have been implemented. Page views as shown in Figure 3 allow the user to consult document images and corresponding database content – including codes that were assigned to medical terms – for any specific patient by page. Tabular views as shown in Figure 2 allow the user to consult pre-selected data in horizontal rows. More complex views are also possible (c.f. Figure 4).

Apart from using standard queries and views that are already available within the system, users can also consult the database on a read-only basis by user-defined SQL queries.

Document generation

Because both patient- and study-specific data are stored in the same database, documents can be generated for a variety of purposes. Two classes of documents can be distinguished: *single documents* and *data listings*.

Single documents (such as DCFs, c.f. Figure 4) relate to a specific patient. Selected data can be automatically combined with relevant scanned images. Single

documents can either be printed directly at the time of generation, or in bulk at regular intervals. The date of generation is inserted into the database automatically.

Listings generated by SQL show the output of data checking programmes, or selected data fields for patients that fulfil certain conditions. Listings are used to optimise tracking of patient data, and data checking. Examples of listings that can be generated from the database are study status reports, listings of documents that are still outstanding, visits that need to be done during a certain time interval, etc.

Management of study medication supplies

SOCDAT contains applications to manage study medication supplies. The medication numbers allocated to each participating centre are tracked, and a unique link between medication number and patient is established. Other functions, such as ordering study medication supplies, can be added if necessary. Because these are an integral part of the database management system, medication orders are automatically ceased for patients who die or who stop study medication prematurely. For long-term studies such as ACTION, the system can track medication expiry dates to prevent patients using study medication that has expired.

Study management tools

The system allows for the flexible addition of study management tools to support efficient trial execution as the need arises. Examples are the tracking of Investigator, CRA and Committee meetings and tracking of on-site monitoring visits.

Provisions for data analysis

The system is not primarily designed as a data analysis tool. Because the architecture is based on a standard relational database development environment, any statistical package that support ODBC (Open Data Base Connectivity) can be used to retrieve and analyse data directly from the system.

SYSTEM VALIDATION

A flexible system that evolves continually because of the addition of new applications needs continuous validation. Because of the modular structure, this can be done per module.

Before a study is started, the document logging, scanning and the data entry application modules are validated by processing a number of completed CRFs that contain dummy data. The data are then read and verified directly from the database. We rely in addition on continuous validation by always showing the user scanned images together with corresponding database content (c.f. Figure 3).

For the validation of other applications we rely on three methods. The first method is the introduction of data satisfying pre-defined conditions to ensure that the desired result is produced by the system. A second method is independent programming of the same application by different programmers. Thirdly, independ-

ently written SQL queries are used to verify that the output produced by a routine or application is correct.

DISCUSSION

In designing SOCDAT, our main goal was to bring together within one system all applications that are necessary for efficient concurrent document, database and study management for clinical trials. Our purpose was also to design a system that is both flexible and can easily be adapted to the requirements of any trial. Nonetheless, SOCDAT has two features that we consider essential components: (1) simultaneous display of scanned documents and corresponding database content, and (2) both study-specific and patient-specific data are stored within the same relational database architecture. We believe the first component to be unique to SOCDAT. Commercial systems that manage document scanning and archiving are widely available but we are not aware of a system that displays database content simultaneously with corresponding scanned documents. We consider this a major advantage over other systems. Brandt *el al*.[4] have commented that "while one can program all kinds of checks without limit, even the most over-engineered system cannot protect against wilful entry of incorrect data that still passes all the checks." Because of the simultaneous display of scanned images and corresponding database content, continuous data validation is an integral part of the system. Although difficult to quantify, we believe that errors that "still passes all the checks" are greatly reduced by this approach.

Storage of document images directly linked to the corresponding database content within the same system has a number of other advantages. We have found the possibility to incorporate scanned images into reports generated from the database particularly important. For the purpose of diagnostic adjudication by a special committee, a patient-specific report is automatically generated for selected Serious Adverse Events (SAEs) in the ACTION study. This report combines data on the patient's history with scanned images of relevant discharge letters, electrocardiograms and laboratory reports. The possibility of using scanned images for this purpose removes the need for photocopying of relevant documents, which saves time. Also, the resulting documents can be printed as many times as necessary, are always of the same quality, and can be stored on a CD. There are many other advantages of scanning. The need for physical handling of documents is reduced to a minimum. All data processing can be done concurrently, which again saves time. Legibility is preserved for documents that fade over time. The complete database including all scanned documents can be stored on removable mass storage media. This is an extra safeguard against loss, and makes it possible to consult the complete documentation and database content from a portable computer without the need for Internet access. This should make both on-site monitoring, and auditing of study sites by company auditing departments or regulatory authorities a great deal more efficient.

Simultaneous display as shown in Figure 3 for the purpose of consultation is useful only when the user can determine the explanation for any discrepancy between the left and the right side of the screen. This also determined our approach to data clarification as the Page Down key and the DCF button give the user access to the full audit trail that explains discrepancies noted on the screen while consulting data (c.f. Figure 4).

SOCDAT allows for the flexible implementation of any function that relates to standard quality control and assurance procedures for clinical trials. Identification of areas where customisation is of benefit is a continual and incremental process as the system evolves to meet the needs of diverse functions and users over time. Because incoming and outgoing documents are registered, listings can be provided of outstanding documents. Laboratory data can be compared to local normal ranges, as the latter are also stored in the system. The same applies to signatures of local study personnel authorised to sign study forms. Listings can be produced of outstanding documents, which can be used by Clinical Research Associates (CRAs) as a monitoring tool. Hence, SOCDAT is not just a data management system for use at the coordinating centre, but is also a management tool for improving quality standards of study execution in general.

While designing the basic system architecture certain deliberate choices were made. We chose to use a classical relational database design, with each variable represented by a column in a multi-column table. Within the same programming environment it would also have been possible to use an Entry-Attribute-Value (EAV) design, with all data stored in one table as described by Brandt *et al.*[4] While an EAV design would have greater flexibility, it would also have been slow on the low-end PC-based platform that we use for large databases.

Another choice was not to implement skip logic and automatic code, range, logical or completeness checks during data entry. While automatic checks during data entry may reduce errors, it may also reduce the productivity of data entry personnel, and lead to entering “data” that passes the checks rather than reflects what was entered by the investigator on the document concerned. We prefer therefore to only prevent the entry of character strings in the wrong format. Reliability of data entry is assured by double data entry. Thereafter, data quality is assured by frequently run batch data checks.

In a concurrent data processing environment, records that are already present in a database may need to be updated at any time. We chose not to store the previous version of an existing record at each update, although this would be possible within the same programming environment – albeit at the cost of a larger database and a slower response time. We believe that storing previous record versions is not necessary with the continuous visual data verification that is inherent to simultaneous display of scanned image and database content. Because each record contains a time stamp and all returned DCFs are logged upon receipt, it is possible to query the database for records that should have been updated.

The approach used in designing SOCDAT may also be used to design a system for a trial that uses a means of local data capture other than pre-printed CRFs. Amongst the options that may be considered in this regard are transmittal of completed pre-printed forms by telefax to the coordinating centre combined with optical character recognition, direct data entry on local workstations, etc. We do not use optical character recognition for two reasons: (1) the inherent constraints that apply to form design, and (2) relative to manual data entry, little efficiency is gained as character recognition must be verified by an operator. Local direct data entry poses problems of source data verification, data authentication and system security for which generally accepted tried-and-tested solutions are only partly available at present. Nonetheless, SOCDAT can be adapted for a hybrid data capture system that relies partly on transfer by telefax or internet of locally entered data, and partly on scanning of paper documents.

Since only one standard development environment is used, system implementation and maintenance does not require large amounts of specialised programming. On a Microsoft Windows NT™ platform the system is stable and system crashes do not occur. Although the current implementations of SOCDAT run on low-end PCs even for studies as large as ACTION, it is noted that actual patient data takes less than half of the total amount of data fields (c.f. Table 1). Apparently, it takes more data to manage data than the actual data itself. In the past this would have been a major disadvantage. Today, storing large amounts of data within complex relational databases structures is no longer a problem. Hence, the implementation of applications related to data checking and management is limited only by the availability of programming capacity.

CONCLUSIONS

Linking and storing scanned documents and data within the same database as is the case in SOCDAT has five major advantages:

1. Automatic display of scanned images in conjunction with corresponding database content ensures integrated continuous validation of data.
2. Scanned images can be incorporated automatically in reports generated by the system
3. After scanning, documents can be permanently archived. Scanning guarantees legibility of documents that fade over time.
4. Data entry and data checking can be done concurrently and at different locations.
5. Scanned documents and database can be stored together on removable mass storage media as an extra safeguard against loss, and to facilitate on-site monitoring or study audits.

REFERENCES

1. Rosenberg M, Mackenzie-Lawrie S. Innovation, efficiency and outsourcing. *Scrip Magazine* May 2003:45-48

2. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline – Guideline for Good Clinical Practice E6. June 10th, 1966 (www.ich.org).
3. Lubsen J, Poole-Wilson PA, Pocock SJ *et al*.. Design and current status of ACTION: A Coronary disease Trial Investigating Outcome with Nifedipine GITS: *Eur Heart J* 1998;**19** (Suppl 1):120-132
4. Brandt CA, Nadkarni P, Marenco L *et al*.. Reengineering a database for clinical trials management: lessons for system architects. *Control Clin Trials* 2000;**21**: 440-461

Chapter 7

ACTION Statistical Analysis Plan (SAP)

Table of contents for this chapter:

1 INTRODUCTION

This document describes the Statistical Analysis Plan (SAP) for the ACTION study. ACTION (**A C**oronary disease **T**rial **I**nvestigating **O**utcome with **N**ifedipine GITS) is a multinational, multicentre, randomised, double blind, placebo controlled, clinical trial designed to evaluate the effect of (long-acting) nifedipine GITS on major cardiovascular event-free survival in patients with chronic symptomatic coronary artery disease (angina pectoris) who are otherwise optimally treated. In total 7,669 patients were randomised to either nifedipine GITS or placebo between November 1996 and December 1998. Patients still under observation had a final end-of-study visit during March – September 2003.

The design and methods of the main ACTION study are described in detail in the main ACTION study protocol version 6.03, dated August 27, 1996 as revised in amendment no. 1 dated June 13, 1997 (the 'protocol'). This SAP has been finalised following the close of patient follow up but before the database was locked and the study medication code was added. The criteria for evaluation of study results and the analysis populations are defined in section 10 of the protocol and an outline of the statistical analysis plan may be found in section 11.2 of the same document. The purpose of this document is to describe the analyses that will be performed in accordance with the protocol in more detail. In addition, this document outlines changes and/or additions made by the Steering Committee to the analyses specified in the protocol.

Changes and/or additions to the analyses specified in the protocol are boxed in this document.

1.1 Study objectives

As stated in section 2 of the protocol, the primary objective for efficacy of ACTION is to assess, relative to placebo, the effect of nifedipine GITS on the combined rate of death from any cause, acute myocardial infarction, emergency coronary angiography for refractory angina, overt heart failure, debilitating stroke and peripheral revascularisation in ambulatory patients with chronic symptomatic coronary artery disease (angina pectoris) but without severely depressed left-ventricular function, who otherwise receive optimal treatment.

The primary criterion for safety is the combined rate of death from any cause, acute myocardial infarction and debilitating stroke.

1.2 Study endpoints

1.2.1 Primary endpoint for efficacy

As stated in section 10.1 of the protocol, the primary criterion for efficacy is *major cardiovascular event-free survival*; defined as the time that each patient is under observation and free of all following events:

1. Death from any cause.
2. Unequivocal acute myocardial infarction.
3. Emergency coronary angiography for refractory angina.
4. Overt heart failure requiring hospitalisation and a change in heart failure treatment.
5. Debilitating stroke.
6. Peripheral revascularisation.

The primary criterion for efficacy will be based on the diagnoses made by the Critical Events Committee (CEC, c.f. 1.2.4). As regards the second event mentioned above, it is noted that the CEC distinguishes between *Acute Myocardial Infarction (AMI)* and *Procedural or accompanying Myocardial Infarction* (PMI). Overt heart failure that occurred during hospitalisation is considered equivalent to the fourth event mentioned above.

1.2.2 Primary endpoint for safety

As stated in section 10.2 of the protocol, the primary criterion for safety is the combined rate of death from any cause, non-fatal acute myocardial infarction and debilitating stroke. This criterion will also be based on the diagnoses made by the Critical Events Committee (CEC, c.f. 1.2.4).

1.2.3 Secondary endpoints

The following secondary endpoints for efficacy are mentioned in section 10.1 the protocol:

- *The occurrence of any of the following cardiovascular events as diagnosed by the CEC:* cardiac death, acute myocardial infarction, emergency coronary angiography for refractory angina, overt heart failure, debilitating stroke or peripheral revascularisation.
- *The percentage observation time that patients are using additional cardiac medications for symptomatic relief of angina and of heart failure.* These percentages will be estimated from data about start and stop dates of concomitant medications collected at each assessment. Patients who undergo CABG or PTCA will be considered to be using additional cardiac medication from the date of the procedure onwards.

Contrary to what is stated in the protocol (c.f. above paragraph), the analysis for percentage observation time that patients are using additional cardiac medication

will be limited to medications registered for symptomatic relief of angina while CABG and PTCA will be ignored (c.f. 2.3.10).

- *The percentage of major cardiovascular event-free survival time that patients were in NYHA class I, II, III and IV respectively.* As described elsewhere (Olsson et al., *Br Med J* 1986;**292**:1491-93) these percentages will be estimated from data about NYHA class collected by assuming that transitions from one class to another occur, on the average, half-way between assessments.
- *The need for coronary angiography and revascularisation procedures as assessed by coronary angiography, revascularisation and major cardiovascular event-free survival.*

In event counts, PTCA and coronary angiography performed on the same date will be counted as PTCA only (c.f. 4.3.2).

The following secondary endpoints for safety are mentioned in section 10.2 of the protocol:

- ECG findings.
- Laboratory test results.
- Vital signs.
- Non-fatal adverse events not included in the primary criteria for efficacy and safety.

Total mortality is an additional secondary safety criterion

1.2.4 Centralised assessment of endpoints

The clinical histories of patients who have potentially sustained any of the events 1 – 6 mentioned in section 1.2.1 of this document have been assessed by the Critical Events Committee (CEC) based on the diagnostic criteria given in the ACTION study document *Diagnostic Classification of Cardiovascular Events and Cause of Death: Criteria and Critical Event Committee Procedures* (version 3.0, dated 8 May 2002). The CEC distinguishes the following diagnostic categories:

1. Cause of death (cardiovascular / non-cardiovascular / unknown).

2.1 Acute myocardial infarction (AMI).

2.2 Procedural or accompanying myocardial infarction (PMI).

3. Refractory angina requiring emergency coronary angiography without progression to myocardial infarction.
4. Overt heart failure requiring hospitalisation (or occurred during hospitalisation) which led to start of, or change in, heart failure treatment.
5. Debilitating stroke.
6. Peripheral revascularisation.

All analyses will be based on the diagnosis made by the CEC. It is noted that the CEC has assessed all clinical events for potential endpoints – not just the first one that occurred – and was required also to determine the date of each event. In the

case two or more events occurred on the same date (e.g. *emergency coronary angiography for refractory angina* and *heart failure*) the CEC also determined the order of occurrence.

1.3 Definition of analysis populations

In accordance with section 10.3 of the protocol, two analysis populations will be distinguished. All tabulations and analyses described in this document will be produced for both populations.

1.3.1 All-randomised population

Membership of the all-randomised population is defined by a valid start date of study medication in the field RANDOM.STARTDT in the ORACLE® database. Hence, this population consists of those patients who are known to have taken at least one tablet of study medication. Compliance with study medication and protocol violations are otherwise irrelevant.

1.3.2 Valid-for-efficacy population

This population consists of the subset of the all-randomised population of patients who comply with certain criteria relating to eligibility and availability of information. The field RANDOM.ELIG_CHK in the ORACLE® database defines membership of the valid-for-efficacy population. If this field contains "OK", no protocol violation was detected and the patient is a member of the valid-for-efficacy population. If this field contains "PV", one or more protocol violations were detected and the patient is a member of the all-randomised, but not of the valid-for-efficacy population.

The criteria for membership of the valid-for-efficacy population mentioned in section 10.3 of the protocol were relaxed by the Steering Committee (see minutes of meeting dated 21 August 1998).

All patients were checked for screening protocol violations based on the following criteria (as relaxed by the Steering Committee):

Eligibility criterion:	In case of violation, criterion violated (item in the CRF):
Age 35 years or older	12.2, #1
In a stable clinical condition for at least 10 days prior to randomisation	12.2, #2
Presence of qualifying angina	12.2, #3
On anti-anginal medication	12.2, #3
Not on a calcium-antagonist during the last 10 days, or any change made to anti-anginal medication	12.2, #3
Presence of at least one criterion for coronary artery disease as mentioned in item 12.2, #4 of the CRF	12.2, #4
Local EF measured and at least 40%, core lab ejection fraction (if present) at least 35%	12.2, #5

Eligibility criterion:	In case of violation, criterion violated (item in the CRF):
Ambulatory	12.2, #7
Signed Declaration of Consent	12.2, #7
Major CV event or surgery within two months	12.3, #1
PTCA or CABG planned	12.3, #2
Known intolerance to dihydropyridines	12.3, #3
Clinically significant valvular disease	12.3, #4
Severe obstructive airway disease	12.3, #5
Unstable insulin-dependent diabetes	12.3, #6
Gastro-intestinal condition limiting absorption or passage	12.3, #7 or #8
Non-CAD condition limiting life expectancy	12.3, #9
Clinically significant heart failure	12.3, #10
Orthostatic hypotension or supine SBP lower than or equal to 90 mm Hg, or SBP unknown	12.3, #11
SBP greater than or equal to 200 mm Hg and/or DBP greater than or equal to 105 mm Hg despite treatment, or BP unknown	12.3, #12
Creatinine above 2 x upper limit of normal, or creatinine unknown	12.3, #13
ALAT or ASAT above 3 x upper limit of normal, or ALAT or ASAT unknown	12.3, #14
Dose of diuretics above limit	12.3, #15
On ACE-inhibitor plus diuretic for heart failure	12.3, #16
On other incompatible medication	12.3, #17
Anticipated problems with compliance or follow up	12.3, #18
Pregnancy, breast feeding or risk of pregnancy (females only)	12.3, #19
Participation in another trial or study	12.3, #20
Time window violations:	
EF assessed more than 6 weeks (echo), or more than one year (other methods) before start of study Rx	
Lab tests assessed more than 6 weeks before start of study Rx	
12-lead ECG recorded more than 10 days before start of study Rx	
Screening medical history assessed more than 10 days before start of study Rx	
Vital signs/baseline physical assessed more than 10 days before start of study Rx	
Final check inclusion / exclusion criteria more than 10 days before start of study Rx	

1.4 **Standard descriptive statistics**

1.4.1 Continuous variables

Unless specified otherwise, the following standard descriptive statistics will always be obtained for continuous variables: number of available values (% of planned values, defined as the number of values that would have been available had there been no loss to follow up – not counting death – and no missing information), mean and standard deviation, median and 95% range, lowest and highest value. The standard table layout for tabulating descriptive statistics for continuous variables appears as Table 1 in *Appendix I – table and figure layouts*.

1.4.2 Categorical variables

For categorical values, the number of values in each category and the corresponding percentage of the total number of values available will be calculated. In addition the number of values that are missing will be given, defined as the number of values that would have been available in addition, had there been no loss to follow up (not counting death) and no missing information. The standard table layout for tabulating descriptive statistics for categorical variables appears as Table 1 in *Appendix I – table and figure layouts*.

1.5 Statistical inference

In accordance with section 11.2 of the protocol, the intention-to-treat analyses for the all-randomised population will be considered as ***primary***. P-value based claims about efficacy and safety will be based on p-values without correction for interim analysis and without adjustment for covariates for the following analyses by assigned treatment (i.e. 'intention-to-treat') for the all-randomised population:

1. Primary analysis for efficacy (c.f. 4.1).
2. Primary analysis for safety (c.f. 4.2).
3. Analysis for all-cause mortality (c.f. 4.3.1).

Data from centres that have been discontinued by the Steering Committee will not be considered in any analysis but a listing of adverse events that were reported up to the date of discontinuation will be provided by treatment group for the centres concerned in the integrated clinical trial report. In the main results paper, the number of centres excluded and the number of patients concerned will be mentioned without giving further details (c.f. Figure 1, *Appendix I – table and figure layouts*).

Although region-treatment and centre-treatment interaction will be considered (c.f. 4.4.2), no region or centre will be excluded from the primary analysis and claims about efficacy or safety will not be made for specific regions only.

It is the intention to report 95% confidence intervals for all analyses unless this is inappropriate or technically not possible; and to report p-values only for the primary analyses mentioned earlier, and for interaction tests.

1.6 Handling of missing information

1.6.1 Incomplete calendar dates

Data Clarification Forms (DCFs) have been sent for calendar dates that were incomplete on the CRF as appropriate. Whether, for example, an exercise test reported on the baseline CRF (i.e. visit 0) for which the date is (partly) missing was indeed performed before start of study medication was confirmed by a DCF. Nonetheless, the day, the month or the year may still be missing. When the day has been entered, the month and the year must be present also. When the day is missing, the month and the year must both be present unless the month is also missing. Incomplete dates where only the year was known were entered as xx/xx (day/month). Completely missing dates were entered in the data base as -99.

For the purpose of statistical analysis, the following rules apply for imputing missing values in calendar dates:

1. When the day is missing but the month and the year are present, the day is taken as 15.
2. When the day and the month are missing but the year is present, the day is taken as 15 and the month as 06 (i.e. June).
3. When the year is also missing, the date is considered as unknown.

The following rules apply for using incomplete calendar dates in *date difference calculations*:

1. When one or both calendar days are missing and the month and the year are the same, the difference will be considered as unknown.
2. When one or both calendar months are missing and the year is the same, the difference will be considered as unknown.
3. When one or both calendar days are missing and the months are different, the fifteenth will be imputed for the missing day(s).
4. When one or both calendar months are missing and the years are different, 06 will be imputed for the missing month(s).

Other rules may need to be defined as analysis programming progresses.

1.6.2 Missing data

In general only data documented by investigators will be tabulated and analysed. However, certain analyses may require that absent data are imputed for the analysis to be meaningful. The exact rules for imputations to be made and their number will be documented with each analysis that requires imputation. If absent values need to be imputed, for instance for performing an overall test of significance of blood pressure differences, the last-observation-carried-forward (LOCF) method will be used for patients who were known to be alive or were lost to follow up. Subsequent to death, either zero will be imputed or the LOCF method will be used as appropriate.

2 DESCRIPTION OF STUDY POPULATION

2.1 Analysis populations

The number of selection criteria violations that occurred in the all-randomised population will be tabulated as shown in *Appendix I – table and figure layouts*, Table 2. The valid-for-efficacy population is defined as the subset of patients for whom no selection criteria violation was noted (c.f. 1.3.2).

2.2 Patient disposition

The contribution of each centre to each of the two treatment arms for both the all-randomised and the valid-for-efficacy population will be tabulated. The tabulation

will be sorted by country and then by centre as shown in *Appendix I – table and figure layouts*, Table 3.

2.3 **Baseline characteristics**

The general format for tabulating patient characteristics at the time of randomisation is shown in *Appendix I – table and figure layouts*, Table 1. Variables to tabulate are (by treatment group to be shown in table captions):

2.3.1 Demographics

Variable label:	**CRF**		**Comment / category labels:**
	Page	**Item**	
Age	15	17.5	Take as (start date of study medication – date of birth expressed in days)/365.25. Tabulate as continuous variable (c.f. Table 1).
Gender	15	17.6	Tabulate N (% of total in analysis population) male.
Race	1	1.5	Tabulate N (% of total in analysis population) white / black / Asian / other. List free-text entries with Pat ID No.

2.3.2 Ejection fraction

Variable label:	**CRF**		**Comment / category labels:**
	Page	**Item**	
Method used	1	2.1	Tabulate N (% of total in analysis population) 2D-echo / radionuclide / contrast / not done. When the date of measurement (item 2.2 CRF) is after start date of study medication (as entered on CRF page 15, item 17.8), consider as 'not done' irrespective of what was ticked for this item.
EF core laboratory	N/A	N/A	Tabulate No. (% of total in analysis population) of patients with a valid core laboratory EF value (N_1). Valid core laboratory values are those with a recording date in the core laboratory database (NOT item 2.2 CRF) before or on the start date of study medication. Tabulate summary statistics as described in section 1.4.1.
Date of measurement (relative to start date of study medication)			Take as (start date of study medication – valid recording date in the core laboratory database) in days. Tabulate median and range. Tabulate N (% of N_1) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).
Investigator value only: 2D-echo			N (% of total in analysis population) = N_2 = number of patients with a valid investigator 2D-echo value only (see below).
Investigator value only: radionuclide			N (% of total in analysis population) = N_3 = number of patients with a valid investigator radionuclide value only (see below).
Investigator value only: contrast			N (% of total in analysis population) = N_4 = number of patients with a valid investigator contrast value only (see below).

Variable label:	CRF		Comment / category labels:
	Page	Item	
Total number of values in tabulation (do not tabulate)			*$N_1 + N_2 + N_3 + N_4$ must be equal to number of patients in tabulation for the valid-for-efficacy population.*
Investigator value only (all methods combined)	1	2.3	Use value only when there is no core laboratory value with a recording date before or on the start date of study medication and the date of measurement (CRF page 1 item 2.2) is before or on the start date of study medication. Tabulate No. (% of total in analysis population) of available values (must be equal to $N_2 + N_3 + N_4$) and summary statistics (c.f. Table 1 for layout).
Date of measurement (relative to start date of study medication)			Take as (start date of study medication – valid date of measurement in item 2.2, page 1 CRF) in days. Tabulate median and range. Tabulate N (% of $N_2 + N_3 + N_4$) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).

Core lab echocardiography data are stored in the table ACTION_ECHO_2D_ALL in the ORACLE® database. Note that this table may contain more than one value for any given patient. The core lab baseline ejection fraction value must be taken from the record that has the last core lab date (field CLDT) before or on the start date of study medication (RANDOM.STARTDT_ CRF).

Whenever possible, the core lab made three assessments of ejection fraction for each analysable echocardiogram (EF1, EF2 and EF3). When all three are available, the ejection fraction value to be used in the analysis is the median. When only two values are available, the value to be used in the analysis is the arithmetic mean. When one value is available, the value present will be used.

2.3.3 Standard laboratory tests

Baseline blood chemistry tests will be considered as having been performed if the date of sampling (item 3.1, CRF page 2) is valid and on or before the start date of study medication. The same applies to haematology. Laboratory tests with incomplete sampling dates will also be considered as having been performed (tests done before randomisation has been confirmed by a DCF).

The time elapsed between date of sampling and start date of study medication will be taken as (start date of study medication – date of sampling) in days. For chemistry and haematology separately, the median and range will be tabulated, as well as N (% of total in analysis population) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. *Appendix I – table and figure layouts*, Table 1 for layout).

Except for total cholesterol, haematocrit and haemoglobin, available baseline laboratory test data will be tabulated by treatment group as follows:

- Standard descriptive statistics for continuous variables (c.f. 1.4.1).
- N (% of No. of available values) above global upper limit of normal.
- N (% of No. of available values) below global upper limit of normal.

As the normal range for these assays depends on gender, haematocrit and haemoglobin will be tabulated separately for men and women.

For total cholesterol, the following will be tabulated:

- Standard descriptive statistics for continuous variables.
- N (% of No. of available values) below 5·0 mmol/l / 5·0 - 6·5 mmol/l / above 6·5 mmol/l.

Standard units and global limits of normal are given in *Appendix II – Standard laboratory test units* (not incorporated in this thesis).

2.3.4 Standard 12-lead ECG

The baseline ECG will be considered as having been recorded if the date of the ECG (item 5.1, CRF page 3) is valid and on or before the start date of study medication. Baseline ECGs with an incomplete date will also be considered as having been recorded (ECG made before randomisation has been confirmed by a DCF).

The time elapsed between date of recording and start date of study medication will be taken as (start date of study medication – date of recording) in days. The median and range will be tabulated, as well as N (% of total in analysis population) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. *Appendix I – table and figure layouts*, Table 1 for layout).

For ventricular rate (item 5.3, CRF page 3), PR or PQ interval (item 5.4, CRF page 3), QRS interval and QT interval (item 5.4, CRF page 3), standard descriptive statistics for continuous variables (c.f. 1.4.1) will be tabulated (ventricular rate in beats/min; cardiac cycle measurements in milliseconds, i.e. the value entered in item 5.4, CRF page 3 multiplied by 10^3).

In addition, the following will be tabulated for item 5.5, CRF page 3:

- N (% of No. of available values) within normal limits / abnormal, not clinically significant / abnormal, clinically significant / ECG recorded but item not completed.
- N (% of No. of available values) atrial fibrillation/flutter (item 5.6, CRF page 3, rhythm-conduction finding #105 or #106 ticked).
- N (% of No. of available values) pacemaker (item 5.6, CRF page 3, rhythm-conduction finding #111 ticked).

2.3.5 Cardiovascular medical history

Variable label:	**CRF**		**Comment / category labels:** (N = number of patients)
	Page	**Item**	
History assessed	4	6.1	N (% of total in analysis population) for which date in item 6.1 is on or before start date of study medication.
History of qualifying angina	4	6.2	Tabulate N (% of total with history assessed) = yes.

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
Duration of qualifying angina (days)	12	16	Take as (start date of study medication – onset date) in days. Tabulate median and range. Tabulate N (% of total with history assessed and item 6.2 = yes) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for lay-out).
Attack frequency	12	16	Tabulate as categorical variable (c.f. 1.4.2).
NYHA class	4	6.3	Tabulate as categorical variable (c.f. 1.4.2).
History of documented AMI	4	6.4	Tabulate N (% of total with history assessed) = yes AND date most recent MI is complete and on or before start date of study medication (no date check is required for incomplete date of most recent MI).
Number of MIs	4	6.4	Tabulate N (% with history of MI) for one / multiple / number unknown.
Time since most recent MI (days)	4	6.4	Take as (start date of study medication – date most recent MI) in days. Tabulate median and range. Tabulate N (% of total with history of MI) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).
History of claudication	4	6.5	Tabulate N (% of total with history assessed) = yes.
History of TIA	4	6.6	Tabulate N (% of total with history assessed) = yes.
History of any stroke	12	16	Tabulate N (% of total with history assessed) for stroke mentioned on CRF page 12 (ICD 9 code = 430, 431, 432, 434.91 or 436).
History of debilitating stroke	4	6.7	Tabulate N (% of total with history assessed) = yes.
History of claudication, TIA or stroke			Tabulate N (% of total with history assessed) = any 6.5 – 6.7 = yes or stroke mentioned on CRF page 12.
Exercise test performed	5	6.9	Tabulate N (% of total with history assessed) = yes AND date most recent test is complete and on or before start date of study medication OR date most recent test is incomplete (i.e. test done before randomisation has been confirmed by a DCF).
Time since last test (days)	5	6.9	Take as (start date of study medication – date last test) in days. Tabulate median and range. Tabulate N (% of total with valid test) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).
Result	5	6.9	Tabulate N (% of total with valid test) positive for CAD / negative for CAD / unknown.
Perfusion scintigraphy performed	5	6.10	Tabulate N (% of total with history assessed) = yes AND date most recent test is complete and on or before start date of study medication OR date most recent test is incomplete (i.e. test done before randomisation has been confirmed by a DCF).
Time since last test (days)	5	6.9	Take as (start date of study medication – date last test) in days. Tabulate median and range. Tabulate N (% of total with valid test) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).
Result	5	6.10	Tabulate N (% of total with valid test) positive for CAD / negative for CAD / unknown.

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
Coronary angiography (CAG) performed	5	6.11	Tabulate N (% of total with history assessed) = yes AND date most recent angiogram is complete and on or before start date of study medication OR date most recent angiogram is incomplete (i.e. angiogram done before randomisation has been confirmed by a DCF).
Number of CAGs	5	6.11	Tabulate N (% of total with valid CAG) one / two / three or more / unknown how many.
Time since last CAG (days)	5	6.9	Take as (start date of study medication – date most recent CAG) in days. Tabulate median and range. Tabulate N (% of total with valid CAG) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).
Result unknown	5	6.11	Tabulate N (% of total with valid CAG) for whom none of the items 1 – 6 in the box concerned was ticked.
No significant lesions	5	6.11	Tabulate N (% of total with valid CAG) 'no significant lesions' ticked AND no other lesion 2 – 6 ticked in the box concerned.
Left main (LM)	5	6.11	Tabulate N (% of total with valid CAG) ticked.
Left anterior descending (LAD)	5	6.11	Tabulate N (% of total with valid CAG) ticked.
Right coronary (RC)	5	6.11	Tabulate N (% of total with valid CAG) ticked.
Circumflex (CX)	5	6.11	Tabulate N (% of total with valid CAG) ticked.
Other	5	6.11	Tabulate N (% of total with valid CAG) ticked.
No. of lesions:			
Other only	5	6.11	N (% of total with valid CAG) other ticked; none of LM, LAD, RC, CX ticked.
One vessel (other than LM)	5	6.11	N (% of total with valid CAG) LM not ticked plus one of LAD, RC, CX ticked.
Two vessels (LM counted as two)	5	6.11	N (% of total with valid CAG) LM but not LAD, RC, CX ticked; or LM not ticked and two of LAD, RCA, CX ticked.
Three vessels (LM counted as two)	5	6.11	N (% of total with valid CAG) LM ticked plus one of LAD, RCA, CX ticked; or LM not ticked and all three LAD, RCA, CX ticked.
Four vessels (LM counted as two)	5	6.11	N (% of total with valid CAG) LM ticked plus two of LAD, RCA, CX ticked.
Five vessels (LM counted as two)	5	6.11	N (% of total with valid CAG) LM ticked plus LAD, RCA, CX all ticked.
PTCA performed	5	6.12	Tabulate N (% of total with history assessed) = yes AND date most recent PTCA is complete and on or before start date of study medication OR date most recent PTCA is incomplete (i.e. PTCA done before randomisation has been confirmed by a DCF).
Number of PTCAs	5	6.12	Tabulate N (% of total with valid PTCA) one / two / three or more / unknown how many.
Number of PTCAs with stenting	12 – 14		Tabulate N (% of total with valid PTCA) with stenting mentioned on CRF page 12 – 14.
Time since last PTCA (days)	5	6.9	Take as (start date of study medication – date most recent PTCA) in days. Tabulate median and range. Tabulate N (% of total with valid PTCA) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. Table 1 for layout).

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
Total No. of vessels dilated	5	6.12	Tabulate N (% of total with valid PTCA) one / two / three or more / unknown how many.
Total No. of vessels successfully dilated	5	6.12	Tabulate N (% of total with valid PTCA) one / two / three or more successes ticked / unknown how many successes.
Left anterior descending (LAD) dilated	5	6.12	Tabulate total N (% of total with valid PTCA) LAD dilated.
Result for LAD	5	6.12	Tabulate total N (% of total with valid PTCA and LAD dilated) that were ticked as successful = no / successful = yes / unknown
Right coronary (RC) dilated	5	6.12	Tabulate total N (% of total with valid PTCA) RC dilated.
Result for RC	5	6.12	Tabulate total N (% of total with valid PTCA and RC dilated) that were ticked as successful = no / successful = yes / unknown.
Circumflex (CX) dilated	5	6.12	Tabulate total N (% of total with valid PTCA) CX dilated.
Result for CX	5	6.12	Tabulate total N (% of total with valid PTCA and CX dilated) that were ticked as successful = no / successful = yes / unknown.
Other dilated	5	6.12	Tabulate total N (% of total with valid PTCA) other dilated.
Result for other	5	6.12	Tabulate total N (% of total with valid PTCA and other dilated) that were ticked as successful = no / successful = yes / unknown.
CABG performed	6	6.13	Tabulate N (% of total with history assessed) = yes AND date most recent CABG is complete and on or before start date of study medication OR date most recent CABG is incomplete (i.e. CABG done before randomisation has been confirmed by a DCF).
Number of CABGs	6	6.13	Tabulate N (% of total with valid CABG) one / two / three or more / unknown how many.
Time since last CABG (days)	5	6.9	Take as (start date of study medication – date most recent CABG) in days. Tabulate median and range. Tabulate N (% of total with valid CABG) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown.
Total No. of vessels bypassed	6	6.13	Tabulate N (% of total with valid CABG) one / two / three or more / unknown how many.
Left main (LM) bypassed	6	6.13	Tabulate total N (% of total with valid CABG) LM bypassed.
Left anterior descending (LAD) bypassed	6	6.13	Tabulate total N (% of total with valid CABG) LAD bypassed.
Right coronary (RC) bypassed	6	6.13	Tabulate total N (% of total with valid CABG) RC bypassed
Circumflex (CX) bypassed	6	6.13	Tabulate total N (% of total with valid CABG) CX bypassed.
Other bypassed	6	6.13	Tabulate total N (% of total with valid CABG) other bypassed.
Risk factors:			

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
History of hypertension treated with drugs at the time of randomisation	6	6.14.1	Tabulate N (% of total with history assessed) where hypertension is mentioned on CRF page 12 (ICD 9 code = 401.9 has been assigned during data entry), any drug registered for lowering blood pressure is documented on CRF pages 125-129, *date first used* of drug(s) concerned is before or on date of first study medication intake (or ticked as *prescribed at moment of randomisation* in the case date first used is incomplete) and route of administration is ticked as *PO/O*.
History of hyperlipidaemia treated with drugs at the time of randomisation	6	6.14.2	Tabulate N (% of total with history assessed) = yes.
Diabetes	6	6.14.3 6.14.4	Tabulate N (% of total with history assessed) 6.14.3 = yes OR 16.4.4 = yes.
Treated with insulin	6	6.14.3	Tabulate N (% of total with history assessed) and 6.14.3 = yes.
Smoking	6	6.14.5	Tabulate N (% of total with history assessed) never smoked / ex-smoker / current smoker / unknown.
Ex-smokers stopped			Tabulate N (% of total with history assessed and ex-smoker) within last year / 1 – 5 years ago / more than 5 years ago / unknown.
Pack-years smoked			Tabulate as continuous variable (c.f. 1.4.1).

2.3.6 Other cardiovascular medical history

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
History assessed	4	6.1	N (% of total in analysis population) for which date in item 6.1 is on or before start date of study medication.
History other CV condition			Tabulate N (% of total with history assessed) with any other CV condition listed on CRF page 12 – 14.
	12 – 14		List/count cardiovascular conditions mentioned by treatment group, sorted by total number of patients concerned.

2.3.7 Non-cardiovascular medical history

Variable label:	CRF		Comment / category labels: (N = number of patients)
	Page	Item	
History assessed	4	6.1	N (% of total in analysis population) for which date in item 6.1 is on or before start date of study medication.
History non-CV condition			Tabulate N (% of total with history assessed) with any non- CV condition listed on CRF page 12 – 14.
	12 – 14		List/count non-cardiovascular conditions mentioned by treatment group sorted by total number of patients concerned.

2.3.8 Height, weight and vital signs

Height, weight and vital signs measurements will be considered as having been performed if the date of assessment (item 8.1, CRF page 7) is valid and on or before the start date of study medication. Baseline vital signs with an incomplete date will also be considered as having been recorded (assessments done before randomisation confirmed by a DCF).

The time elapsed between date of assessment and start date of study medication will be taken as (start date of study medication – date of assessment) in days. The median and range will be tabulated, as well as N (% of total in analysis population) for 0 – 14 days / 15 – 42 days / 43 – 182 days / 183 – 365 days / more than 365 days ago / unknown (c.f. *Appendix I – table and figure layouts*, Table 1 for layout). Standard descriptive statistics for continuous variables (c.f. 1.4.1) for body weight, seated pulse rate and body height respectively will be tabulated by treatment group (c.f. *Appendix I – table and figure layouts*, Table 1 for layout). For item 8.3 (pulse rate) N (% of No. of available values) regular / irregular / unknown will be given also.

Body mass index (BMI):

BMI will be taken as weight in kilograms divided by the squared height in metres and will be tabulated as follows:

- Standard descriptive statistics for a continuous variable (c.f. 1.4.1).
- N (% of No. of available values) below 25 kg/m^2 / 25·00 – 29·99 kg/m^2 / 30·00 – 39·99 kg/m^2 / 40·00 kg/m^2 or above.

Systolic blood pressure (SBP):

In addition to standard descriptive statistics for a continuous variable (c.f. 1.4.1), N(% of No. of available values) below 100 mm Hg / 100 - 119 mm Hg / 120 - 139 mm Hg / 140 - 159 mm Hg / 160 - 179 mm Hg / 180 mm Hg or above will be tabulated (c.f. *Appendix I – table and figure layouts*, Table 1 for layout).

Diastolic blood pressure (DBP):

In addition to standard descriptive statistics for a continuous variable (c.f. 1.4.1), N(% of No. of available values) below 70 mm Hg / 70 - 79 mm Hg / 80 - 89 mm Hg / 90 - 99 mm Hg / 100 - 109 mm Hg / 110 mm Hg or above (c.f. *Appendix I – table and figure layouts*, Table 1 for layout).

Blood pressure classification (1999 WHO/ISH classification):

Tabulate as categorical variables (c.f. *Appendix I – table and figure layouts*, Table 1 for layout):

- N (% of No. of available values) optimal (SBP < 120 and DBP < 80) / normal (SBP 120 – 129, DBP 80 – 84) / high normal (SBP 130 – 139, DBP 85 – 89) / Grade 1 hypertension (SBP 140 – 159, DBP 90 – 99) / Grade 2 hypertension (SBP 160 – 179 or DBP 100 – 109) / Grade 3 hypertension (SBP > 179 or DBP

> 109). *When a patient's systolic and diastolic blood pressures fall into different categories, the more severe category will apply.*

- N (% of No. of available values) isolated systolic hypertension (SBP > 139 and DBP < 90).

2.3.9 Physical examination

Physical examination abnormalities will be listed as part of the cardiovascular or non-cardiovascular medical history (c.f. 2.3.6 and 2.3.7) as appropriate.

2.3.10 Concomitant treatment

For drugs entered on the concomitant treatment log pages (CRF page 125 – 129) with a date first used before or on the start date of study medication and a date last used that is either empty or after the start date of study medication, or for drugs for which the field CT.BEFORE_START contains "Y" (i.e. when the investigator has indicated on a DCF that the treatment concerned was used at baseline although the date first used on the CRF is incomplete), the following will be tabulated:

- N (%) on a beta-blocker.
- N (%) on an ACE-inhibitor.
- N (%) on a diuretic.
- Etc.
- N (%) using anti-anginal medication categorised as none / one / two or more, counting only drugs registered for symptomatic relief of angina as anti-anginal medication.

Cardiovascular and non-cardiovascular treatments will be tabulated separately. Tabulations will be sorted by frequency of use.

3 COMPLIANCE WITH FOLLOW UP

3.1 Person-time of follow up realised

The ORACLE® database field V_SYS_VISIT_DATES.PLANDT contains for each patient the planned date of the end-of-study visit as calculated in accordance with the ACTION study document *Close-Out Procedures: Guideline for Investigators and CRAs* (version 2 dated 14 February 2003).

Follow up for clinical events after the planned date of the end-of-study visit will be ignored in time-to-event analyses.

The total number of person-years of *intended follow up* will be calculated by treatment arm by summation of the following quantities:

1. For patients who died on or before the planned date of the end-of-study visit: *sum of years elapsed between start date of study medication and date of death.*

2. For patients who were seen for the end-of-study visit on or after the planned date: *sum of years elapsed between start date of study medication and planned date of end-of-study visit.*
3. For patients who had the last out-patient clinic visit in the time interval 42 days – 1 day (both days included) before the planned date of the end-of-study visit: *sum of years elapsed between start date of study medication and actual date of last visit.*
4. For patients who were last seen or contacted earlier than 42 days before the planned date of the end-of-study visit: *sum of years elapsed between start date of study medication and planned date of end-of-study visit.*

Note that these categories concern mutually exclusive groups of patients. Calculations will be done in days, which will be converted to years by dividing the number of days by 365·25.

Patients belonging to category 3 above (i.e. last out-patient clinic visit 42 days – 1 day before the planned date of the end-of-study visit) will not be labelled as 'lost to follow up' in trial reports.

The total number of person years of intended follow up as calculated above will be displayed by treatment arm as shown in *Appendix I – table and figure layouts*, Figure 1. In the same figure, the following quantities will also be displayed by treatment arm:

- For patients who died: *sum of years elapsed between start date of study medication and date of death* (i.e. total follow up as calculated for patients belonging to category 1 above).
- For patients who terminated follow up earlier than planned: *sum of years elapsed between start date of study medication and actual date last seen or contacted* (i.e. total follow up for patients belonging to category 4 above as calculated by using the actual date last seen or contacted, rather than the planned date of the end-of-study visit).
- For patients who completed study as planned: sum of years of follow up *as calculated for category 2 above* plus *sum of years of follow up as calculated for category 3 above.*

The percentage of intended follow up completed will be taken as (follow up for patients who died *plus* follow up for patients who terminated study earlier than planned *plus* follow up for patients who completed follow up as planned) divided by intended follow up as defined earlier.

3.2 Time windows for planned follow up visits

The planned date for each follow up visit or contact is given in years and weeks relative to the start date of study medication in Table 1 in section 9.4.2.1 of the protocol. Relative to the planned visit or contact date based on this table, the following time windows will be used as a basis for deciding whether the planned follow up

visit or contact concerned has actually been performed (c.f. summary ACTION SAP meeting dated April 17-18, 2003):

- For visit 1 (2-week clinic visit): minus seven days – plus seven days (first and last day included).
- For visit 2 (6-week clinic visit): minus 2 x 7 = 14 days – plus 3 x 7 = 21 days (first and last day included).
- For visits 3 – 26 (interim contacts, half-yearly clinic visits and end-of-study visit): minus 6 x 7 = 42 days – plus 7 x 7 = 49 days (first and last day included).

Data from visits outside these time windows will not be analysed statistically.

3.3 Planned follow up visits / contacts performed

The actual dates of (extra) visits and contacts in the database and the time windows given in section 3.2 will be used to determine whether a valid planned visit or contact is available for each time window concerned. If more than one visit or contact took place within a certain time window, the planned visit or contact closest to the planned date will be considered valid even if an extra visit closer to the planned date is available. In the case only extra visits have taken place within the time window concerned, the extra visit that is closest to the planned visit or contact date will be considered equivalent to a valid planned visit or contact. Note that this implies that follow up data may be available for analysis even though the planned visit concerned didn't take place within the time window specified in the protocol, or wasn't done at all. Note also that the end-of-study visit will be considered valid when done no later than 49 days after the planned date.

For each planned visit or contact specified in the protocol, the percentage of visits or contacts that took place will be calculated as follows. The numerator is equal to the number of patients who had a valid planned visit or contact within the time window concerned as given in section 3.2 of this document. This must be divided by the number of patients who were still under follow up on the last day of the time window for the visit or contact concerned. The latter is obtained by taking the number of patients in the analysis population concerned *minus* the number of patients who had died either before or on the last day of the time window concerned, *minus* the number of patients who were last seen or contacted either before or on the last day of the time window concerned. Results will be tabulated as shown in *Appendix I – table and figure layouts*, Table 4.

3.4 Compliance with study medication

3.4.1 Study medication continued at each visit

For each valid planned follow up visit or contact (c.f. 3.3), the following will be calculated:

1. Number of patients who were continued on any dose of study medication at the visit or contact concerned (CRF item 7.1, 10.1 or 12.1 – depending on visit

number and type (i.e. planned or extra) – ticked as either *30 mg/day* or *60 mg/day*).

2. The fraction of patients who were continued on any dose of study medication at the visit or contact concerned, using the total number of valid planned visits or contacts that were performed as denominator.
3. Number of patients who were continued on full dose of study medication at the visit concerned (CRF item 7.1, 10.1 or 12.1 ticked as *60 mg/day*).
4. The fraction of patients who were continued on full dose of study medication at the visit concerned, using the total number of valid planned visits or contacts as denominator.

Results will be tabulated in similar manner as shown in *Appendix I – table and figure layouts*, Table 4.

Continuation of study medication by visit will be displayed by plotting the following fractions for each treatment group and for each planned clinic visit or contact as specified in the protocol:

1. Fraction of patients still alive after the planned duration of follow up for the visit concerned (as obtained from the Kaplan-Meier curve for total mortality).
2. Fraction of patients still alive after the planned duration of follow up for the visit concerned, multiplied by the fraction of patients on any dose of study medication for the visit concerned.
3. Fraction of patients still alive after the planned duration of follow up for the visit concerned, multiplied by the fraction of patients on full dose of study medication for the visit concerned.

3.4.2 Time compliance

Time compliance with the study medication regimen (% of time until event or censoring on full dose / % of time until event or censoring on half dose) will be calculated based on start and stop dates documented on the study medication log pages (CRF page 130 – 134), and on the event or censoring dates concerned. Calculations must be done separately for the time-to-event analyses described below in the sections *Primary analysis for efficacy* (c.f. 4.1) and *Primary analysis for safety* (c.f. 4.2) respectively, and for the secondary time to event analyses described in section 4.3.1. Results will be tabulated for each time-to-event analysis mentioned as shown in *Appendix I – table and figure layouts*, Table 5.

3.4.3 Tablet compliance

Based on tablet count data documented on the study medication log pages, study medication tablet compliance will be calculated as the percentage of predicted tablets that was actually used.

4 EVALUATION OF EFFICACY AND SAFETY

4.1 Primary analysis for efficacy

The primary analysis for efficacy (c.f. 1.2.1) is a time-to-event analysis specified as follows:

Start date: Date of first study medication intake.

Event date: Date of the first of the following: death (all causes), acute myocardial infarction (AMI), procedural or accompanying myocardial infarction (PMI), refractory angina requiring emergency coronary angiography, overt heart failure, debilitating stroke, peripheral revascularisation (as diagnosed by the CEC, c.f. 1.2.4).

Censoring date: Date that the patient was last seen or contacted (patient lost to follow up, end-of-study visit took place before planned date), or planned date of end-of-study visit.

The planned date of the end-of-study visit in the database has been calculated as described in the ACTION study document *Close-Out Procedures: Guideline for Investigators and CRAs* (version 2 dated 14 February 2003).

Events occurring after the planned date of the end-of-study visit will not be considered.

The primary analysis variable is *major cardiovascular event-free survival*, defined as (date of first study medication intake *minus* date of event or censoring) in days. It follows that in the case the date of the first event is the same as the date of first study medication intake, the number of days until first event is equal to zero.

Using the SAS® version 8 proc LIFETEST or equivalent, Kaplan-Meier curves will be plotted that show major cardiovascular event-free survival as a function of observation time by treatment group. These Kaplan Meier plots will be presented 'going up', i.e. starting at 0% experiencing any of the events concerned at the time of start of first study medication intake. Below the horizontal axis the number of patients still 'at risk' of a first major cardiovascular event will be presented at each half year (i.e. 365·25/2 days) of follow up for each treatment arm (c.f. *Appendix I – table and figure layouts*, Figure 2)

The primary statistical analysis testing the null hypothesis that there is no treatment difference will be a log-rank test (unadjusted for any covariates) accompanied by the estimated hazard ratio and its 95% confidence interval. Hazard ratios for nifedipine will be taken relative to placebo. Hence, a hazard ratio less than 1 will indicate that patients assigned to nifedipine GITS have a decreased risk.

The primary endpoint is a composite of several events. The occurrence of each component event (first event only), the number of patients with any combined event, the number of patient years at risk for the combined event and the event rate (hazard) per 100 patient years 'at risk' will be tabulated by treatment group as shown in *Appendix I – table and figure layouts*, Table 5. The hazard ratio (relative

to placebo), its 95% confidence interval and the p-value for the log-rank test will be given as shown. Note that these tabulations also list the number of patients withdrawn permanently from study medication before the combined endpoint concerned occurred, and the applicable percentage of time 'at risk' for this combined endpoint that full- or half-dose study medication were prescribed.

4.2 Primary analysis for safety

The primary analysis for safety (c.f. 1.2.2) is a time-to-event analysis similar to the primary analysis for efficacy (c.f. 4.1), but considering only the first of the following events: death (all causes), acute myocardial infarction (AMI), procedural or accompanying myocardial infarction (PMI), debilitating stroke (as diagnosed by the CEC, c.f. 1.2.4).

The primary variable for safety will be analysed and documented in the same manner as the primary analysis variable for efficacy (c.f. 4.1).

Events occurring after the planned date of the end-of-study visit will not be considered.

4.3 Secondary Analyses

4.3.1 Time-to-event analyses

Secondary time-to-event analyses will be performed in the same manner as described for the primary analysis for efficacy (c.f. 4.1) for the following combined endpoints:

1. Death from any cause.
2. Any major cardiovascular event as diagnosed by the CEC (i.e. the primary endpoint for efficacy minus non-cardiovascular death as diagnosed by the CEC – secondary efficacy criterion according to protocol).

Deaths with cause classified by the CEC as 'unknown' will be considered as cardiovascular.

3. Death or any cardiovascular event as diagnosed by the CEC or cardiac intervention (i.e. the primary endpoint for efficacy plus coronary angiography, PTCA, CABG – secondary efficacy criterion according to protocol).
4. Any major vascular event or revascularisation (i.e. the primary endpoint for efficacy minus non-cardiovascular death as diagnosed by the CEC, minus overt heart failure; plus PTCA, CABG – additional pre-specified secondary efficacy criterion).

Each of these time-to-event analyses will be tabulated in the same manner as the primary analysis variable for efficacy (c.f. 4.1). For tabulation purposes, events which occur on the same day must be ordered. For events diagnosed by the CEC with the same date, the sort order for determining the event that occurred first may

be found in the CEC database. For time-to-event analyses that also consider the procedures coronary angiography, PTCA and CABG, the following rules apply:

1. If the first event includes any event diagnosed by the CEC (c.f. 1.2.4), the event terminating event-free follow up will be taken from the CEC database, using the sort order from the CEC database in the case of multiple events diagnosed by the CEC on the same day.
2. If the date of the first event concerns *one* of the procedures *coronary angiography, PTCA,* or *CABG* in the absence of an event diagnosed by the CEC on the same day, the event terminating event-free follow up will be taken as the procedure concerned (note that *coronary angiography* is counted as event only in the third secondary time-to-event analysis mentioned earlier, while PTCA and CABG are counted only in the fourth).
3. If the date of the first event concerns *any two or three* of the procedures *coronary angiography, PTCA,* or *CABG* in the absence of an event diagnosed by the CEC on the same day, the sort order which determines the event terminating event-free follow up will be taken as PTCA, CABG, coronary angiography (note that *coronary angiography* is counted as event only in the third secondary time-to-event analysis mentioned earlier).

Events occurring after the planned date of the end-of-study visit will not be considered.

4.3.2 Analyses for individual events

The total occurrence of each individual event or procedure mentioned in the previous section, the number of patients with at least one event, the number of patient years at risk of event and the event rate (hazard) per 100 patient years 'at risk' will be tabulated by treatment group as shown in *Appendix I – table and figure layouts*, Table 6. For each event, the hazard ratio (relative to placebo), its 95% confidence interval and the p-value for the log-rank test will be given as shown. Patient years at risk will be taken as the sum of the observation time from start of study medication until the first occurrence of the event concerned, or until censoring because of death, planned end of observation or date last seen or contacted.

For tabulating coronary angiography and PTCA as individual events, the following rule applies: if a PTCA is documented as an SAE and coronary angiography is documented in the SAE table on the same date as PTCA, only PTCA will be counted.

Events occurring after the planned date of the end-of-study visit will not be considered.

4.3.3 NYHA Class

The principal analysis of NYHA class is based on the following information:

- Start date of study medication.

- Date of occurrence of the first of the following: death (all causes), acute myocardial infarction (AMI), procedural or accompanying myocardial infarction (PMI), refractory angina requiring emergency coronary angiography, overt heart failure, debilitating stroke, peripheral revascularisation occurred (i.e. the primary endpoint for efficacy, based on diagnoses made by the CEC, c.f. 1.2.4).
- Actual date that the patient was last seen or contacted (patient lost to follow up, end-of-study visit took place before planned date).
- Planned date of end-of-study visit.
- NYHA class values for valid planned follow up visits and for the end-of-study visit (c.f. 3.3).

The analysis will first calculate the length of follow up in days until death or censoring (as used in time-to-event analyses, not counting days after the planned date of the end-of-study visit). For patients who had one of the events mentioned earlier, this time will then be subdivided into event-free survival time, and survival time after first event. Next, event-free survival time will be subdivided into the time spent in each of the four NYHA classes. This will be done as follows. First, the start date and the end-date of event-free survival time will be determined. Next, the *actual* dates of available NYHA class values recorded for baseline and for valid planned visits (c.f. 3.3) done while the patient was event-free will be determined. For patients who were event-free at the end-of-study visit, the end-of-study NYHA class value date will be taken as the date the patient was last seen or contacted (for end-of-study visits done before the planned date), or as the planned date of the end-of-study visit (for end-of-study visits done on or after the planned date). The end-of-study visit NYHA class will be considered missing in the case the end-of-study visit was done more than 49 days after the planned date.

Based on valid planned NYHA class assessments and their actual dates as defined earlier, event-free survival time will be subdivided into the number of days spent in each NYHA class, assuming that changes in NYHA class occurred at the mid-point between the dates that NYHA class was recorded. In the case there is no valid end-of-study NYHA class value, NYHA class will be considered not to have changed since the last valid assessment (which may be the baseline assessment). Note that this analysis uses only recorded NYHA class values and is not affected by missing follow up observations. In the case the baseline NYHA class value is missing, the mode (not the mean) of the overall distribution at baseline of available assessments will be imputed.

Having obtained the amount of event-free survival time spent in each NYHA class for each patient, and if applicable survival time after first event, the mean event-free survival time spent in each NYHA class, and the survival time after first event, will be calculated and tabulated for each treatment group.

An overall significance test for the differences between treatment groups with respect to mean time in each NYHA class will be based on the following ranking scale:

1 = Alive, free of any event mentioned earlier, in NYHA class I.

2 = Alive, free of any event mentioned earlier, in NYHA class II.

3 = Alive, free of any event mentioned earlier, in NYHA class III.

4 = Alive, free of any event mentioned earlier, in NYHA class IV.

5 = Alive but first event has occurred.

6 = Died.

For each patient, the time elapsed between start of study medication and censoring will be calculated assuming that no patient died. This time will then be subdivided by the time spent in each class of the ranking scale 1 – 6 just given. For the time elapsed between start of study medication and first event, this will be done as described above for NYHA class. Ranks 5 and 6 on the other hand will be considered as 'absorbing states'. From the results, the weighted average rank will be calculated for each patient, using the time spent in each rank as weights. Weighted average ranks will be compared between treatment groups using the following model:

Weighted average rank = c_0 + c_1·(NYHA class at baseline) + c_2·(duration of follow up assuming no death) + c_3·(indicator of treatment) + random error.

For each planned visit the availability of NYHA class data will be tabulated by treatment group as for other information collected at routine follow up visits. This will be done both including and excluding valid planned visits performed subsequent to the occurrence of the primary endpoint for efficacy. The following data structure will be the basis for displaying the evolution of NYHA class over time for each treatment group (imaginary data):

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	NYHA class (% in patient free of primary efficacy endpoint):				Points to plot by month:			
month	I	II	III	IV	S(t)	IV	III	II
0	0	70	30	0	1.00	1.00	0.70	0.00
6	10	62	24	4	0.97	0.93	0.70	0.10
12	20	55	18	7	0.94	0.88	0.71	0.19
18	25	44	22	9	0.91	0.83	0.63	0.23
24	30	42	17	11	0.89	0.79	0.64	0.27
30	33	40	12	15	0.86	0.73	0.63	0.28
36	30	38	17	15	0.83	0.71	0.57	0.25
Etc.								

By column, the following is tabulated:

[1] Planned timing of subsequent clinic visits.

[2] – [5] For each planned follow up visit and the end-of-study visit: percentage of patients who are free of the primary efficacy endpoint, and who are known to be in class NYHA I, II, III or IV respectively. The denominator for calculating the percentages is equal to the total number of NYHA class assessments available for the valid planned follow up visit concerned (c.f. 3.3) amongst patients who are free of the primary efficacy endpoint *at the actual date of the visit concerned*.

[6] Kaplan-Meier estimates of major cardiovascular event-free survival for the time points corresponding to planned visits.

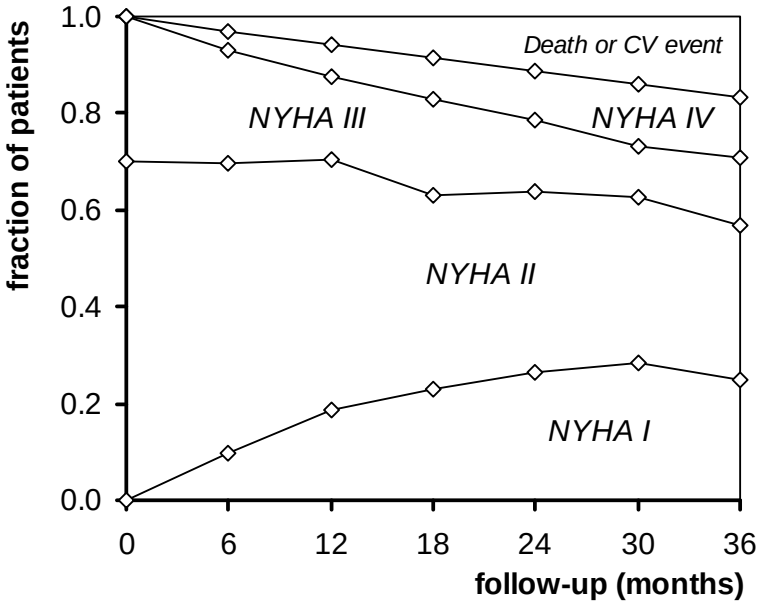
[7] For each subsequent follow up visit: estimated fraction of patients who are free of the primary efficacy endpoint and who are in either NYHA class I, II or III, as obtained by multiplying the entry in column [6] with the sum of the entries in column [2] – [4], divided by 100 (for example for 12 months: 0·94 x [(20 + 55 + 18)/100] = 0·88).

[8] Same as in column [7] but now for either NYHA class I or II (for example for 12 months: 0·94 x [(20 + 55)/100] = 0·71).

[9] Same as in column [8] but now for NYHA class I only (for example for 12 months: 0·94 x (20/100) = 0·19).

Note that in columns [2] – [5], only actually available NYHA class values are considered and percentages are taken relative to the total number of patients free of the primary efficacy endpoint for whom NYHA class is available. This solves the problem of missing NYHA class values.

The points in column [1] and columns [6] – [9] can now be plotted as shown below:



The Kaplan-Meier estimates of survival will also be plotted in the above figure.

A secondary analysis for NYHA class will be performed as follows. Suppose that valid NYHA class data for planned visits (c.f. 3.3) and hospitalisation data for a certain patient have the following form:

Date(0): Start of study medication, baseline NYHA class = II.

Date(1): Valid planned out-patient clinic visit, NYHA assessed as class I.

Date(2): Valid planned out-patient clinic visit, NYHA assessed as class IV.

Date(3): Admitted to hospital (for any reason).

Date(4): Discharged from hospital.

Date(5): Valid planned out-patient clinic visit, NYHA assessed as class I.

Date(6): End of follow up, end-of-study NYHA (c.f. 4.3.3) assessed as class III.

This patient contributes [date(6) – date(0)] in days to total survival time, subdivided as follows:

- From date(0) to ½·[date(1) – date(0)]: no. of days in NYHA class II (assuming as before that transitions between NYHA classes occur, on the average, halfway between assessments).
- From ½·[date(1) – date(0)] to date(1): no. of days in NYHA class I.
- From date(1) to ½·[date(2) – date(1)]: no. of days in NYHA class I.
- From ½·[date(2) – date(1)] to date(2): no. of days in NYHA class IV.

- From date(2) to date(3): no. of days in NYHA class IV (assuming that there is no change in NYHA class between the last assessment before hospitalisation and the date of admission).
- From date(3) to date(4): no. of days in hospital.
- From date(4) to date(5): no. of days in NYHA class I (assuming that the patient is discharged in the same state as the first NYHA class assessment subsequent to discharge).
- From date(5) to 1/2·[date(6) – date(5)]: no. of days in NYHA class III.

In the case the end-of-study NYHA class value is missing (i.e. patient died or was lost to follow up), the assumption will be made that NYHA class doesn't change after the last assessment, and the last quantity mentioned above will be replaced by:

- From date(5) to date(6): no. of days in NYHA class I.

In addition, the following rules apply:

1. If the baseline NYHA class is missing, impute the mode (not the mean) of the overall distribution at baseline of available assessments.
2. If there are no valid follow up NYHA class assessments before hospitalisation, take the NYHA class up to the day of admission as the baseline NYHA class value.
3. If the patient was in hospital at the planned date of the end-of-study visit (or on the date last seen or contacted for patients lost to follow up), count only the days up to this date.
4. If there are no follow up NYHA class assessments after discharge from hospital, assume that the patient was discharged in the same NYHA class as recorded for the last known assessment before hospitalisation (which may be the baseline value).

For tabulation and analysis, the same methods will be used as described earlier in this section for the analysis that considers NYHA class only up to the first of the events 1 - 6 mentioned in section 1.2.1 of this document.

4.3.4 Vital signs at planned follow up visits

Available data on vital signs, body weight, body mass index and blood pressure classification for valid planned follow up visits (c.f. 3.3) will be tabulated by treatment group for each visit in the same manner as at baseline (c.f. 2.3.8). Data to be used include data from end-of-study visits done after the planned date, provided that these were done no later than 49 days after the planned date (c.f. 3.2). Changes from baseline will not be tabulated but graphs will be provided showing median changes from baseline with 95% ranges by treatment group for each planned follow up visit (c.f. *Appendix I – table and figure layouts*, Figure 3). A standard test of the overall difference between treatment groups will be performed.

4.3.5 Blood pressure control over time

For the purpose of assessing blood pressure control over time, available blood pressure data for valid planned follow up visits (c.f. 3.3) will be categorised as follows:

I = Optimal (SBP < 120 and DBP < 80).
II = Normal (SBP 120 – 129, DBP 80 – 84).
III = High normal (SBP 130 – 139, DBP 85 – 89).
IV = Grade 1 hypertension (SBP 140 – 159, DBP 90 – 99).
V = Grade 2 hypertension (SBP 160 – 179, DBP 100 – 109).
VI = Grade 3 hypertension (SBP > 179, DBP > 109).

When a patient's systolic and diastolic blood pressures fall into different categories, the more severe category will apply. Mean time spent in each of the above categories will be analysed and displayed in the same manner as described elsewhere for NYHA class (c.f. 4.3.3).

4.3.6 Standard laboratory tests and ECG at planned follow up visits

Available standard laboratory test and ECG data for valid planned follow up visits (c.f. 3.3) will be tabulated by treatment group for each visit in the same manner as at baseline (c.f. 2.3.3 and 2.3.4). Data to be used include data from end-of-study visits done after the planned date provided that these were done no later than 49 days after the planned date (c.f. 3.2). Changes from baseline will not be tabulated but graphs will be provided showing median changes from baseline with 95% ranges by treatment group for each planned follow up visit (c.f. *Appendix I – table and figure layouts*, Figure 3). A standard test of the overall difference between treatment groups will be performed.

4.3.7 Concomitant medication at planned follow up visits

Using the *actual* date of the visit for each valid planned visit (c.f. 3.3) as reference date, the following will be tabulated based on data entered on concomitant treatment log pages in the CRF:

- N (%) on a beta-blocker.
- N (%) on an ACE-inhibitor.
- N (%) on a diuretic.
- Etc.
- N (%) using anti-anginal medication categorised as none / one / two or more, counting only drugs registered for symptomatic relief of angina as anti-anginal medication.

Cardiovascular and non-cardiovascular treatments will be tabulated separately. Tabulations will be sorted by frequency of use.

Concomitant medication at the *end of the study* will be tabulated in the same manner, using as reference date the following as applicable:

- For patients who died: date of death.
- For patients who were lost to follow up or had the end-of-study visit before the planned date: date patient last seen or contacted.
- For patients who had the end-of-study visit on or after the planned date: planned date of the end-of-study visit.

4.3.8 Use of anti-anginal drugs over time

The number of anti-anginal drugs used in addition to study medication (none / one / two or more) will be analysed in the manner as described earlier for categories of NYHA class (c.f. 3.3.4). Only drugs registered for symptomatic relief of angina will be counted as anti-anginal medication.

4.3.9 Number of days spent in hospital

For each treatment group, the number of patients ever hospitalised, the total number of hospitalisations, and mean duration of hospital stay will be tabulated. Days in hospital after the planned end of follow up will not be counted.

4.4 **Adjusted and subgroup analyses**

4.4.1 Covariate adjustment

The main covariate-adjusted proportional hazards model planned in the protocol for the primary endpoints for efficacy and safety, and for all-cause mortality, will be adjusted for the following variables: age, sex, NYHA class, ejection fraction, history of MI, history of diabetes, systolic and diastolic blood pressure. Quantitative covariates will be fitted continuously. Binary variables will be fitted by indicator variables coded as 0 = characteristic absent, 1 = characteristic present. Categorical variables will be represented by (1 – number of categories) binary variables. All variables mentioned will be included irrespective of the statistical significance of the coefficient. From this analysis, the primary feature of interest is the hazard ratio for the covariate-adjusted treatment comparison and its 95% confidence interval and p-value. In addition, the proportional hazard model will be displayed for each covariate. Model coefficients will be converted into hazard ratios and their 95% confidence intervals (based on the standard error of the coefficient); for quantitative variables this will be the hazard ratio per a suitable unit of increase in the covariate.

In order to develop a more extensive proportional hazards model predicting risk for the primary efficacy and safety outcomes and for mortality for the purpose of risk stratification (an analysis pre-specified in point 9, section 11.2 of the protocol), the following baseline variables will be considered in addition to those listed above: serum creatinine, glucose, cholesterol, uric acid, haematocrit, left-ventricular hypertrophy, QTc interval, atrial fibrillation or flutter (c.f. CRF page 3), history of claudication, history of TIA, history of stroke, previous coronary intervention (CABG, PTCA), previous coronary angiogram (yes with left main stem

and/or multi-vessel disease, single vessel disease but not left main stem, no significant disease or no angiogram), smoking (current, ex, never), body mass index, height, pulse rate, blood pressure, angina attack frequency, angina at rest (as ticked on CRF page 12), treated for hypertension, treated for hyperlipidaemia with drugs, treated with: beta-blocker, ACE-inhibitor and/or A2 receptor antagonist, long and/or short acting nitrate, aspirin and/or other anti-platelet agents, digoxin, anti-arrhythmic drugs, diuretics, previous treatment with calcium antagonist, number of anti-anginal drugs (none, one or two).

For each outcome, from this large set of potential predictors a subset of significant independent predictors will be selected by a forward stepwise selection procedure based on proportional hazard models with increasing numbers of covariates that always also include treatment assignment. In this forward stepwise sequence, the most significant predictor not in the model so far will be included next at each step. At the same time, the need for interaction terms with treatment assignment will be considered. Also, certain variables may be combined (i.e. *history of peripheral artery disease*, combining history of claudication, TIA or stroke) or replaced by indicators representing categories (i.e. two indicators representing three categories of baseline blood pressure). The final model will be such that all variables and interaction terms are independently predictive at $p<0{\cdot}05$, but the next most significant variable not in the model fails to achieve $p<0{\cdot}05$. In addition, the presence of each variable in the final model must be supported by biological plausibility or relevant findings about risk factors from other studies.

For each of the three outcomes concerned, the final model will be displayed with hazard ratios and their 95% confidence intervals for each covariate. The linear predictive function in this final proportional hazards model will then be used as a risk score to categorise patients into five levels of risk. The risk score will be calculated assuming that the patient was assigned to placebo. The five risk categories will be chosen such that they contain equal numbers of patients experiencing the outcome (both treatment assignments combined). For each of these five risk categories the following will be tabulated:

- Number of patients, number of events and number of patient years by treatment group.
- Event rate per 100 patient years by treatment group.
- Hazard ratio and its 95% confidence interval relative to placebo.

In addition, a p-value for interaction will be calculated. Results will be displayed as shown for age in Figure 4, *Appendix I – table and figure layouts*.

4.4.2 Subgroup analyses

Pre-defined subgroup analyses for the primary endpoints for efficacy and safety, and for all-cause mortality, will be performed for the following baseline variables (as defined in tabulations of baseline characteristics, c.f. 2.3):

- Age (in tertiles of the age distribution for patients with event).
- Sex (male / female).

- NYHA class (class I / II / III and IV combined).
- Ejection fraction (in tertiles of the ejection fraction distribution for patients with event).
- History of document AMI (no / yes item 6.4 CRF page 4).
- History of diabetes (no / yes).
- On a beta-blocker.
- On lipid-lowering drug therapy.
- Use of a calcium-antagonist before entry.
- On an ACE-inhibitor or an AII-antagonist.
- On digoxin.
- Diastolic blood pressure (DBP), in tertiles of the DBP distribution for patients with event.
- Systolic blood pressure (SBP) in tertiles of the SBP distribution for patients with event.
- Blood pressure classification categorised as described in section 2.3.8.

For each outcome/subgroup variable combination, a statistical test of interaction will be performed to test whether there is evidence of heterogeneity in hazard ratios between subgroups. For quantitative subgroup variables, these interaction tests will take into account the ordering of the subgroup categories. Subgroup analyses will be displayed as shown in *Appendix I – table and figure layouts*, Figure 4

In addition, subgroup analyses will be performed for the following (regional groupings of) countries:

- Canada.
- Israel.
- Netherlands and Belgium.
- Germany, Austria and Switzerland.
- Denmark, Finland, Norway and Sweden.
- France, Spain, Portugal, Italy, Greece.
- UK, Australia, New Zealand.

For each region and for each outcome the number of patients, number of events and hazard by treatment group, and the hazard ratio with its 95% confidence interval will be tabulated. For each outcome a statistical test of interaction will be performed to see if evidence of heterogeneity in hazard ratios between regions is present.

Subgroup analyses will also be tabulated by country and by centre but due to the larger number of categories, many with small numbers, only the numbers of patients and events by treatment group will be displayed for each country, and each centre.

4.5 **(Serious) adverse events**

The occurrence of coded (serious) adverse events after start of study medication (AEs with an onset date = start date of study medication included) will be tabulated by treatment group as follows:

4.5.1 For each system organ class

- Total number of events belonging to system organ class.
- Total number of events belonging to system organ class on study medication.
- Total number of events belonging to system organ class that were reported as *serious*.
- Total number of events belonging to system organ class on study medication that were reported as *serious*.
- Number of patients with at least one event belonging to system organ class.
- Number of patients with at least one event belonging to system organ class on study medication.
- Number of patients with at least one event belonging to system organ class on study medication that was reported as *serious*.

4.5.2 For each code within any given system organ class

- Total number of occurrences.
- Total number of occurrences on study medication.
- Total number of occurrences that were reported as serious.
- Total number of occurrences on study medication that were reported as serious.
- Number of patients with at least one occurrence.
- Number of patients with at least one occurrence on study medication.
- Number of patients with at least one occurrence on study medication that was reported as *serious*.

5 **PROCEDURES**

5.1 **Where the programming will be done**

Baseline comparisons and univariate time-to-event analyses for the primary and secondary efficacy and safety criteria (including the analysis for time spent in each of four NYHA classes) will be independently programmed at SOCAR Research SA and at the Medical Statistics Unit of the London School of Hygiene and Tropical Medicine (LSHTM), based on analysis data sets created independently. Analysis data sets must be documented as shown in *Appendix III – Documentation of analysis data sets* for the SAS data set PATINF. Programming at SOCAR will be done in such manner that the output (from SAS®: SAS font 10, no more than 75 characters per line) can be inserted directly into the integrated clinical report to be prepared by SOCAR. Subgroup and covariate-adjusted analyses using Cox regression will

be programmed at the LSHTM. Output for final analyses and models must be formatted such that direct insertion into the report mentioned earlier is possible. Programming and output generation of descriptive follow up data by visit will be done only at SOCAR.

5.2 Validation

For the purpose of validation, data for the first 500 patients randomised will be stored separately on CDs that contain identical files. A dummy treatment code will be added to allow tabulations by treatment group and statistical analyses comparing treatment groups can be tested.

All summary statistics tabulated will be validated independently. Acceptable methods are:

- Summary statistics obtained by SQL queries performed directly on the ORACLE® database.
- Independent analysis using any standard statistical package (SAS®, STATA®, Statistica®, etc.) of datasets obtained by SQL queries performed directly on the ORACLE® database.
- Independent analysis of data sets created within any standard statistical package (SAS, STATA, Statistica, etc.).

All analyses must be fully documented by either a printout of the SQL queries performed directly on the ORACLE® database, or a declaration of variables used from the ORACLE® database.

APPENDIX I – TABLE AND FIGURE LAYOUTS

Table 1: Tabulation layout for descriptive statistics

Table <#>: <*Type of data*> at <*time point*> (<all-randomised> or <valid-for-efficacy>)*

	Nifedipine§ N=# ###	Placebo§ N=# ###
Continuous variables (e.g. age):		
Age (years)		
No. available (% of N in column head)	N (%)	N (%)
Mean (sd)		
Lowest		
2.5th percentile		
Median		
97.5th percentile		
Highest		
Age group (years)		
35 – 44 (% of available values)¶	N (%)	N (%)
45 – 54 (% of available values)	N (%)	N (%)
55 – 64 (% of available values)	N (%)	N (%)
65 – 74 (% of available values)	N (%)	N (%)
75 – 84 (% of available values)	N (%)	N (%)
Over 84 (% of available values)	N (%)	N (%)
Binary variables (e.g. gender)		
No. available (% of N in column head)	N (%)	N (%)
Male (% of available values)	N (%)	N (%)
Female (% of available values)	N (%)	N (%)
Categorical variables (e.g. NYHA class)		
No. available (% of N in column head)	N (%)	N (%)
NYHA class I (% of available values)	N (%)	N (%)
NYHA class II (% of available values)	N (%)	N (%)
NYHA class III (% of available values)	N (%)	N (%)
NYHA class V (% of available values)	N (%)	N (%)

* The table heading must contain the type of data tabulated (e.g. *demographics*), the time point (e.g. *baseline*) and – between brackets – the analysis population concerned (either *all-randomised* or *valid-for-efficacy*).

§ The column heading must show the number of patients for whom the data tabulated should have been available, taking death and loss to follow up into account.

¶ Percentages by category must be based on the total number of available values = 100.

Table 2: Admission protocol violations

Total number of patients randomised (%)	N (100)
PVs of inclusion criteria:	
Age less than 35 years (CRF 12.2.1)	N (%)
Not in a stable clinical condition and/or no symptomatic angina (CRF 12.2.2)	N (%)
Not on anti-anginal treatment, Ca-antagonist not washed out, or on Ca-antagonist (CRF 12.2.3)	N (%)
Criteria for coronary artery disease not met (CRF 12.2.4)	N (%)
Core lab LV EF less than 35%, or local value and core lab value missing (CRF 12.2.5)	N (%)
Not ambulatory or no informed consent signed (CRF 12.2.7)	N (%)
Patients with at least one PV of inclusion criteria	N (%)
PVs of exclusion criteria related to medical history:	N (%)
Recent major CV event or surgery (CRF 12.3.1)	N (%)
PTCA or CABG planned (CRF 12.3.2)	N (%)
Known intolerance to diphydropyridines (CRF 12.3.3)	N (%)
Clinically significant valvular disease (CRF 12.3.4)	N (%)
Severe obstructive airway disease (CRF 12.3.5)	N (%)
Unstable insulin-dependent diabetes (CRF 12.3.6)	N (%)
Gastro-intestinal condition limiting absorption or passage (CRF 12.3.7 or 12.3.8)	N (%)
Non-CAD condition limiting life expectancy (CRF 12.3.9)	N (%)
Patients with at least one PV related to medical history	N (%)
PVs of exclusion criteria related to current symptoms or findings:	
Clinically significant heart failure (CRF 12.3.10)	N (%)
Orthostatic hypotension or supine SBP lower than or equal to 90 mm Hg, or SBP unknown (CRF 12.3.11)	N (%)
SBP greater than or equal to 200 mm Hg and/or DBP greater than or equal to 105 mm Hg despite treatment, or BP unknown (CRF 12.3.12)	N (%)
Creatinine above 2 x upper limit of normal, or creatinine unknown (CRF 12.3.13)	N (%)

ALAT or ASAT above 3 x upper limit of normal, or ALAT or ASAT unknown (CRF 12.3.14)	N (%)
Patients with at least one PV related to current symptoms or findings	N (%)
PVs of exclusion criteria related to current treatment:	
Dose of diuretics above limit (CRF 12.3.15)	N (%)
On ACE-inhibitor plus diuretic for heart failure (CRF 12.3.16)	N (%)
On other incompatible medication (CRF 12.3.17)	N (%)
Patients with at least one PV related to current treatment	N (%)
PVs of miscellaneous exclusion criteria:	
Anticipated problems with compliance or follow up (CRF 12.3.18)	N (%)
Pregnancy, breast feeding or risk of pregnancy (females only, CRF 12.3.19)	N (%)
Participation in another trial or study (CRF 12.3.20)	N (%)
Patients with at least one PV related to miscellaneous exclusion criteria	N (%)
Other protocol violations:	
Any time window violation (CRF 9.4.1)	N (%)
Valid-for-efficacy population:	
Number of patients with any protocol violation	N (%)
Valid for efficacy population	N (%)

Table 3: Patient disposition by country and by centre

		Nifedipine	Placebo
Total number of patients: all-randomised		N (100)	N (100)
Total number of patients: valid-for-efficacy		N (100)	N (100)
Australia:			
AUS-001:	all-randomised	N (%)	N (%)
	valid-for-efficacy	N (%)	N (%)
AUS-002:	all-randomised	N (%)	N (%)
	valid-for-efficacy	N (%)	N (%)
Next centre, etc.			
Total Australia:	all-randomised	N (%)	N (%)
	valid-for-efficacy	N (%)	N (%)
Next country, etc.			

Table 4: Completeness of follow up visits (<all-randomised> or <valid-for-efficacy>)

	Nifedipine N_T=# ###	Placebo N_T=# ###
Visit 1 (7 - 21 days)		
Died before or on last day of window (% of N in column head)	N_1 (%)	N_1 (%)
Lost to f-up before or on last day of window (% of N in column head)	N_2 (%)	N_2 (%)
Under f-up at end of window (% of N in column head)	N_3 (%)	N_3 (%)
Any visit within window (% of N under f-up at end of window)	N (%)	N (%)
	Cave: N_3 must be equal to $N_T - N_1 - N_2$	
Visit 2 (6 wks – 14, + 21 days)		
Died before or on last day of window (% of N in column head)	N_1 (%)	N_1 (%)
Lost to f-up before or on last day of window (% of N in column head)	N_2 (%)	N_2 (%)
Under f-up at end of window (% of N in column head)	N_3 (%)	N_3 (%)
Any visit within window (% of N under f-up at end of window)	N (%)	N (%)
	Cave: N_3 must be equal to $N_T - N_1 - N_2$	
Visit 3 (13 wks – 42, + 49 days)		
Died before or on last day of window (% of N in column head)	N_1 (%)	N_1 (%)
Lost to f-up before or on last day of window (% of N in column head)	N_2 (%)	N_2 (%)
Under f-up at end of window (% of N in column head)	N_3 (%)	N_3 (%)
Any visit within window (% of N under f-up at end of window)	N (%)	N (%)
	Cave: N_3 must be equal to $N_T - N_1 - N_2$	
Etc.		

f-up, follow up

Table 5: Tabulation layout for combined endpoints

The example below shows the table layout for the primary endpoint for efficacy for the all-randomised population. Tables for other combined endpoints differ only as regards the individual endpoints considered.

	Nifedipine		Placebo		
	Number	Event rate*	Number	Event rate*	HR (95% CI)
Total patients (all randomised)	N NNN	-	N NNN	-	
Total patient years 'at risk'	N NNN	-	N NNN	-	
On full dose study medication (%) §					
On half dose study medication (%) §					
No of patients permanently withdrawn from study medication while 'at risk'	N (%)		N (%)		
No of patients with first event¶					
Death (all causes)	N	-	N	-	-
Acute MI	N	-	N	-	-
Procedural MI	N	-	N	-	-
Refractory angina	N	-	N	-	-
Heart failure	N	-	N	-	-
Stroke	N	-	N	-	-
Peripheral revascularisation	N	-	N	-	-
Any of the above	N	h	N	h	HR (h_l - h_u) (p=value)#

* Number of events per 100 patient years 'at risk' for the event concerned.
§ Percentage of total patient years of follow up.
¶ As diagnosed by the Critical Events Committee (CEC).
# p-value from log-rank test without adjustment for co-variates.
CI, confidence interval; HR, hazard ratio (nifedipine relative to placebo); MI, myocardial infarction.

Table 6: Tabulation layout for individual endpoints

	Nifedipine		Placebo		
Event	Total No. of events	Pts with event (event rate*)	Total No. of events	Pts with event (event rate*)	
Death (all causes)	-	N (h)	-	N (h)	hr (h_l - h_u)
Non-cardiovascular§	-	N (h)	-	N (h)	hr (h_l - h_u)
Cardio-vascular§	-	N (h)	-	N (h)	hr (h_l - h_u)
Unknown§	-	N (h)	-	N (h)	hr (h_l - h_u)
Acute MI§	N	N (h)	N	N (h)	hr (h_l - h_u)
Procedural MI§	N	N (h)	N	N (h)	hr (h_l - h_u)
Refractory angina§	N	N (h)	N	N (h)	hr (h_l - h_u)
Heart failure§	N	N (h)	N	N (h)	hr (h_l - h_u)
Stroke§	N	N (h)	N	N (h)	hr (h_l - h_u)
Peripheral revascularisation§	N	N (h)	N	N (h)	hr (h_l - h_u)
PTCA	N	N (h)	N	N (h)	hr (h_l - h_u)
CABG	N	N (h)	N	N (h)	hr (h_l - h_u)
Coronary angiography¶	N	N (h)	N	N (h)	hr (h_l - h_u)

* Number of events per 100 patient years ‘at risk’ for the event concerned.
§ As classified by the Critical Events Committee (CEC).
¶ Without PTCA on same day.
MI, myocardial infarction; PTCA, percutaneous coronary angioplasty; CABG, coronary artery by-pass grafting.

Figure 1: Trial profile

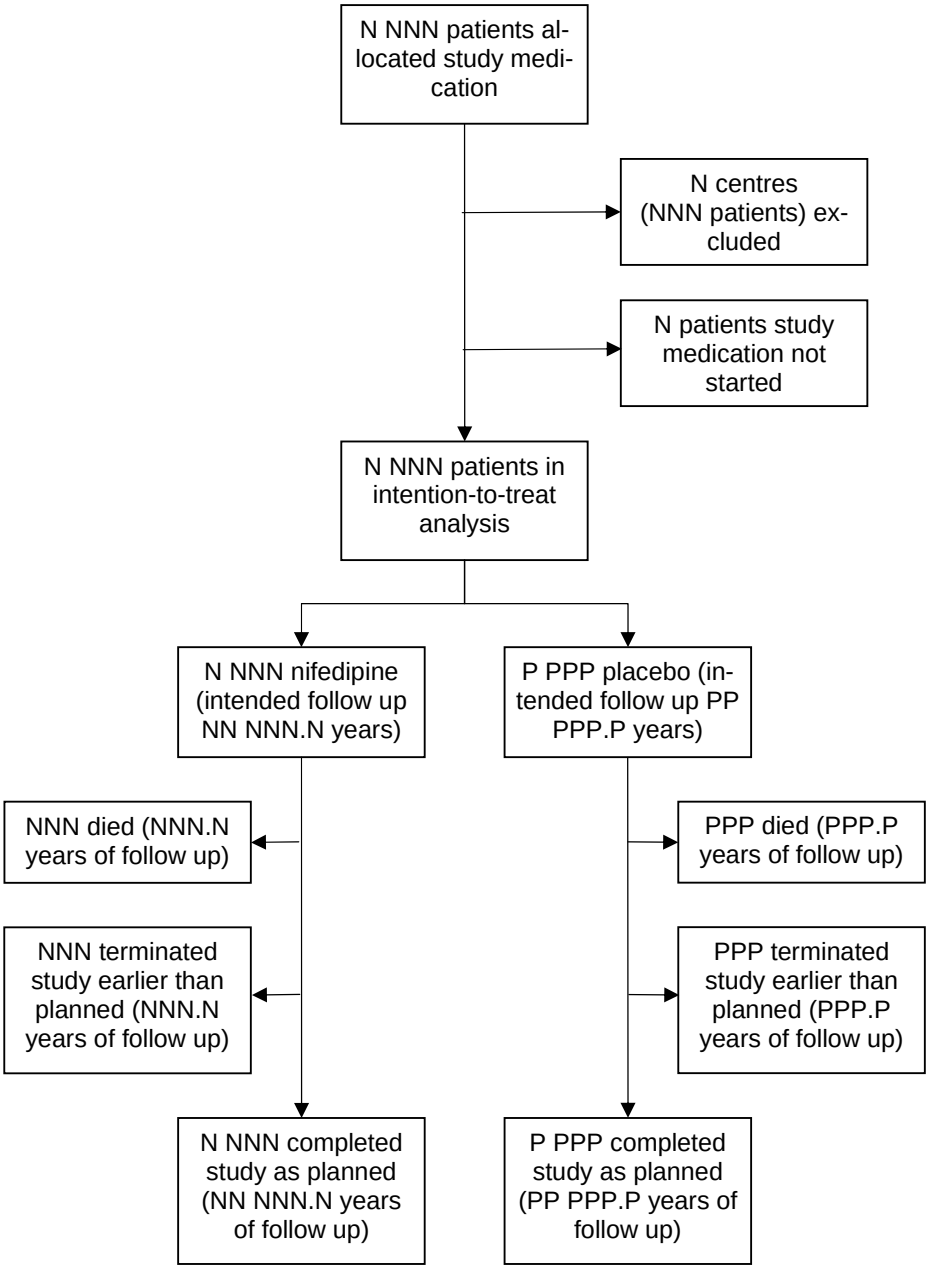


Figure 2: Display of time to occurrence of clinical outcomes

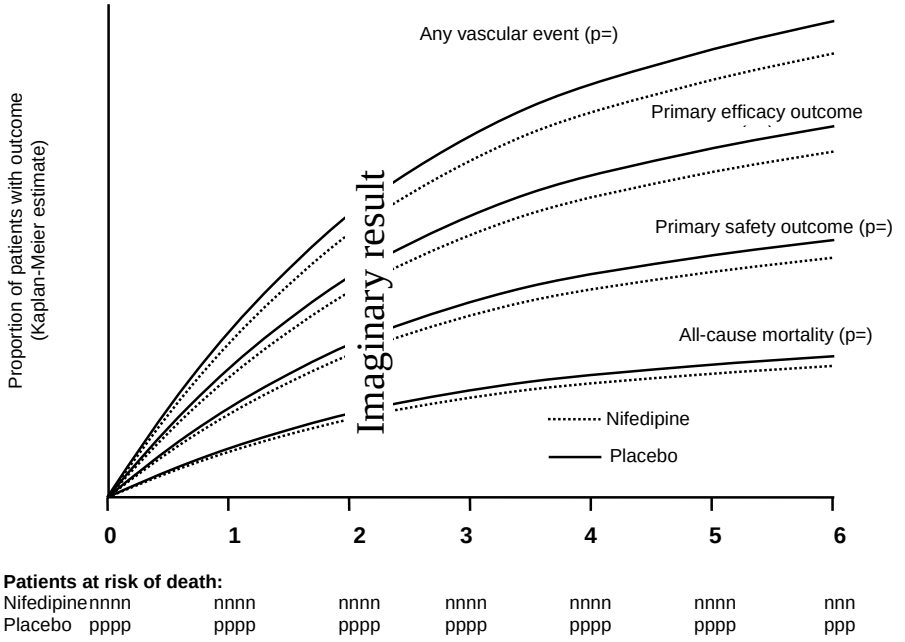
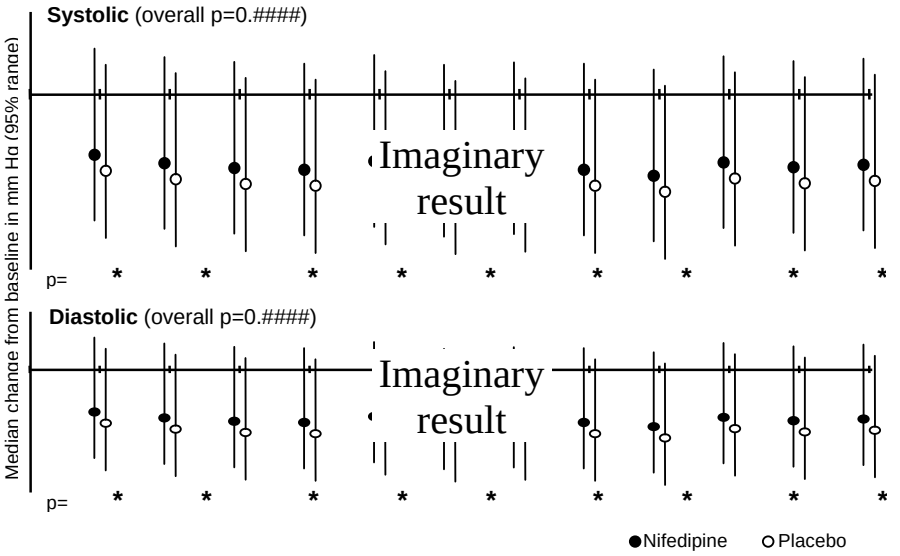


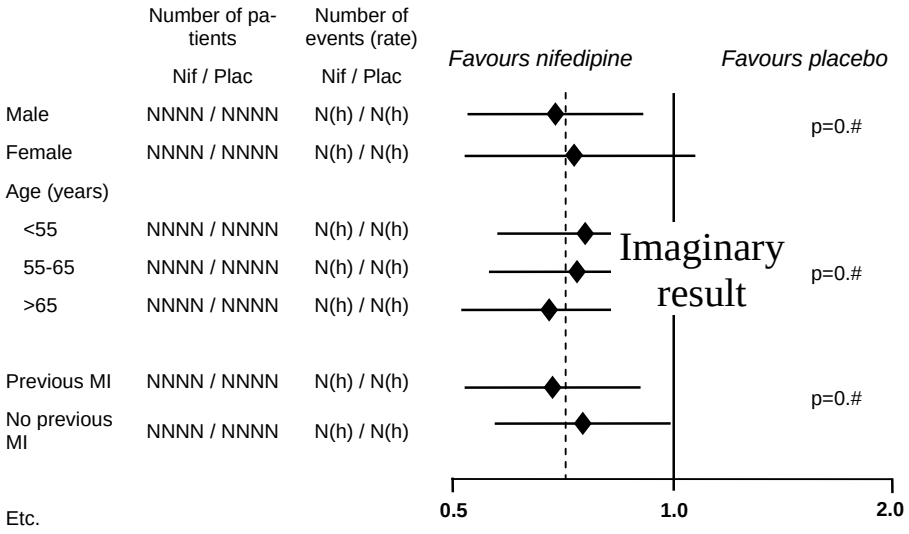
Figure 3: Display of evolution over time of repetitive observations

E.g. for systolic and diastolic blood pressure and blood pressure control:



Time (yrs)	**0.5**	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**	**5.5**	**6.0**
Number of values:												
Nifedipine:	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn	nnnn
Placebo:	pppp	pppp	pppp	pppp	pppp	pppp	pppp	pppp	pppp	pppp	pppp	pppp
% blood pressure controlled (systolic < 140 and diastolic blood pressure < 90):												
Nifedipine	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n	nn.n
Placebo	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p	pp.p

Figure 4: Display of subgroup analyses



MI, myocardial infarction. Rates in number of events per 100 patient years of follow up 'at risk'. P-values for interaction or trend.

APPENDIX II – STANDARD LABORATORY TEST UNITS

Not incorporated in this thesis.

APPENDIX III – DOCUMENTATION OF ANALYSIS DATA SETS

Analysis Dataset: PATINF
Description: Patient Information.
Source Tables (ORACLE): C_DD_DDA, CT, DEATH, HISTORY1, HISTORY2, HISTORY3, MEDREC, PATID, PHYSICAL, RANDOM, STUDYLOG, V_SYS_VISIT_DATES, WITHDRAW
Sorted by: PATID, one record per patient

Variable Name	**Label**	**ORACLE Source** (TABLE.*var1, var 2*)	**Type**	**Format**	**Notes**
PATID	Unique Patient Identifier	RANDOM.*patid*	Num	8.	= RANDOM.*patid*
SCRNR	Screening Number	RANDOM.*scrnr*	Num	8.	= RANDOM.*scrnr*
INIT	Patient Initials	RANDOM.*init*	Char	$3.	= RANDOM.*init*
CENTNO	Centre	RANDOM.*centno*	Char	$9.	= RANDOM.*centno*
COUNTRY	Country	Derived	Char	country.	Derived from *centno*. Take letter(s) preceding '-' to identify country
DOBDT	Date of Birth	RANDOM.*dob_crf*	Num	date9.	RANDOM.*dob_crf* converted to SAS date format. The general imputation rule (see Note 1 below) is used for partial dates
RANDDT	Date of Randomisation	RANDOM.*startdt_crf*	Num	date9.	RANDOM.*startdt_crf* converted to SAS date format. If RANDOM.*startdt_crf* is missing or incomplete then RANDDT is missing (no imputation done)
AGE	Age at Baseline	Derived	Num	8.	Age in years on day of randomisation, calculated using *randdt* and *dobdt*.
SEX	Sex	RANDOM.*sex*	Num	sex.	= PATID.*sex*
RACE	Race	PATID.*race*	Num	race.	= PATID.*race*
RACEOTH	Race (other)	PATID.*race, race_o*	Char	$40.	= PATID.*race_o* for patients where RACE = 4
EF_DT	EF – Date Performed	ECHO_2D_ALL.*cldt* PATID.*efdt_crf*	Num	date9.	Last core lab date (ECHO_2D_ALL.*cldt*) that is on or before the start date of study medication (i.e. RANDDT). If no such value exists, then EF_DT = PATID.*efdt_crf*. EF_DT must be in SAS date format

Variable Name	Label	ORACLE Source (TABLE.*var1, var 2*)	Type	Format	Notes
EF_MTH	EF – Method Used	PATID.*efmeth, efdt_crf*	Num	efmeth.	= PATID.*efmeth* **if** PATID.*efdt_crf* (converted to SAS date format) is on or before the start date of study medication (i.e. RANDDT) = missing otherwise
EF_IO	EF – Investigator value only	Derived	Num	noyes.	= 2 (i.e. 'yes') if EF_DT = . **and** EF_MTH is not missing = 1 (i.e. 'no') otherwise
EF_B	Ejection Fraction (%) at Baseline	ECHO_2D_ALL.*ef1, ef2, ef3* PATID.*ef*	Num	8.2	If ECHO_2D_ALL.*ef1, ef2, ef3* are all present, then EF_B is equal to the median of the three values. If only two are present, then EF_B is equal to the mean of the two values. If only one value is present, then EF_B is equal to this value. If there are no values, then EF_B = PATID.*ef*
Remainder omitted from this thesis					

Notes:

1. *Rules for imputing partial dates:* The general rule for imputing partial dates is described in section 1.6.1 of the SAP. If the month for the partial date is missing then it is imputed to be '06'; if the day is missing it is imputed to be '15'. For certain variables, additional rules must be used to prevent implausible values being imputed; these rules are discussed in the individual specifications for these variables.
2. *Analysis Populations:* Populations currently assigned by running program P:/ACTION/SAS/REPORT/2_SET UP POPULATIONS.SAS

Formats for data set PATINF:

Format	Type	Values
country.	Character	'AUS' = 'Australia' 'A' = 'Austria' 'B' = 'Belgium' 'C' = 'Canada' 'DK' = 'Denmark' 'FIN' = 'Finland' 'F' = 'France' 'D' = 'Germany' 'GR' = 'Greece' 'IL' = 'Israel' 'I' = 'Italy' 'NL' = 'Netherlands' 'NZ' = 'New Zealand' 'N' = 'Norway' 'P' = 'Portugal' 'E' = 'Spain' 'S' = 'Sweden' 'CH' = 'Switzerland' 'UK' = 'United Kingdom'
efmeth.	Numeric	1 = '2D-echocardiography' 2 = 'Radionuclide scanning' 3 = 'Contrast ventriculography' . = 'Not Done'
noyes.	Numeric	1 = 'No' 2 = 'Yes'
race.	Numeric	1 = 'White' 2 = 'Black' 3 = 'Asian' 4 = 'Other' 99 = 'Unknown'
sex.	Numeric	1 = 'Male' 2 = 'Female'

Notes:

1. In SAS, missing numeric values are denoted by '.' (quotes not included)
   e.g. *if x = . then x = 1* means "If x is missing, then let x=1"
2. In SAS, Missing character values are denoted by '' (quotes included)
   e.g. *if x = '' then x= 'one'* means "if x is missing, then let x = 'one' "

Chapter 8

General discussion

The section on *Trial Conduct Considerations* in the ICH *Note for Guidance on Statistical Principles for Clinical Trials* (ICH E9)[1] states on page 20: "For the purpose of overseeing the quality of the trial the checks involved in trial monitoring may include whether the protocol is being followed, the acceptability of data being accrued, the success of planned accrual targets, the appropriateness of the design assumptions, success in keeping patients in the trials, etc.". On page 4, the same document defines *bias* as: "the systematic tendency of any factors associated with the design, conduct, analysis and interpretation of the results of clinical trials to make the estimate of treatment effect deviate from its true value". Monitoring means to watch, keep track of, or check.[2] The purpose of trial monitoring as defined in ICH E9 cannot merely be to watch passively, but surely it must also be to prevent errors during trial conduct pro-actively. We prefer therefore the term management over the term monitoring, and define *trial management* as all activities directed at ensuring appropriate trial conduct while avoiding bias as defined earlier.

Based on our experience with managing the PICO (Chapter 2) and ACTION (Chapter 4) trials, this chapter puts trial management into perspective. In an ideal trial investigators make no mistakes with patient selection, follow all patients according to protocol, report events as required and complete Case Report Forms (CRFs) without error or omission. Ideal patients undergo follow up assessments exactly as planned, and take study treatment as instructed. In practice the ideal is rarely if ever achieved however and contingencies that compromise the ideal do occur. Management of an ongoing trial requires therefore that such contingencies are spotted as soon as possible, corrected as appropriate, and prevented from reoccurring. Every individual contingency detected requires a trial management decision as to how it should be handled. In the interest of consistency, these decisions must be based on pre-defined policies based on the scientific principles underlying randomised clinical trials.

In this chapter we consider the policies implemented to deal with the most important contingencies that occurred during the conduct of the two trials mentioned earlier. Our purpose is to show that there is an intricate relationship between how a trial is managed and the underlying scientific principles. We will distinguish between *validity of comparison* and *generalisability*, and will consider a treatment comparison valid "if it is based on comparable groups of persons treated and observed in such a way so as to make treatment assignment the most likely explanation of the result observed".[3] Importantly, validity of comparison "derives from the design of the trial and the way it is carried out".[3] Generalisability on the other hand

has to do with applicability to clinical practice. As such it is "largely a matter of judgement"[3], and is an issue only when a trial result can be considered valid. The chapter ends with a list of recommendations.

PREREQUISITES FOR TRIAL MANAGEMENT

Realising the complexities of multi-centre trials

The proper conduct of a single-centre study or trial requires care and attention to detail. This is true a fortiori for medium to large multi-centre trials such as PICO and ACTION. Hence, the complexity of managing such trials needs to be appreciated from the outset and nothing can be left to chance. One cannot assume that investigators interpret a certain medical term, or execute a certain measurement in the same way. While managing the ACTION trial, we noted that the same medication trade name was used in different countries for different drugs. Despite efforts to standardise units of measurement, variation between centres remains. The same applies to normal ranges for even the most frequently used laboratory tests. In ACTION, we started by using local normal ranges when evaluating laboratory tests for each participating centre, but found this unmanageable because of frequent changes. In the same study, 19 different countries from various geographical and linguistic regions were involved. While selecting centres, it is therefore necessary not just to consider whether the availability of patients, the facilities and the experience with performing trials are adequate. One must also ensure that the command of oral and written English by all research personnel involved is sufficient for this to be the official language for all study documents and communication *(recommendation 1)*. Nonetheless, documents such as the patient information sheet and the declaration of informed consent need translation in local languages, and need to comply with local regulations. In some countries, a version of the protocol and CRF must be provided in the local language in order to obtain approval of local ethics committee or other competent authorities.

Protocol and case report form

No trial can start before complete trial documentation has been approved by all parties involved. Within a trial, the protocol has the same role as the law in a constitutional state. A well-developed protocol is therefore a prerequisite for effective trial management. In the ICH *Note for Guidance on Good Clinical Practice* (ICH E6), the protocol is defined as "a document that describes the objective(s), design, methodology, statistical considerations and organisation of a trial".[4] ICH E6 also comprehensively defines the topics that the protocol must cover. Importantly, the protocol is not only the basis for investigators and other participants (committee members, coordinating centre, etc.) to judge whether they can fulfil the requirements for participation, but also for judging whether the conditions are acceptable. To mention just two issues as regards the latter: (i) the protocol needs to specify a policy that allows, if this is intended, publication regardless of the trial's outcome; and (ii) the protocol is generally also the basis for contracts that set out "any ar-

rangements on delegation and distribution of tasks and obligations, and, if appropriate, on financial matters".[4]

Apart from a well-developed protocol, trial management and investigators alike need a workable CRF. ICH E6 defines the CRF as "a printed, optical, or electronic document designed to record all of the protocol-required information to be reported … on each trial subject".[4] One reason that a good CRF is so important is that few investigators will consult the protocol each time a patient is seen. The CRF should therefore not just be a document on which data can be recorded but also a checklist that instructs the investigator what to do when and how. In this regard, the CRF should follow the sequence in which examinations, laboratory tests, etc. are normally carried out in clinical practice. To ensure standardisation both within and between centres, relevant instructions for carrying out examinations and tests should be printed on the opposite pages so that these are visible during CRF completion. Relevant instructions are thus always directly available for consultation by investigators and their staff while collecting the data required without the need to flip through the CRF, let alone consult the protocol. We emphasize that procedures for generally-known measurements such as blood pressure, or assessment of NYHA class, should also be specified in detail in the CRF in this manner *(recommendation 2)*.

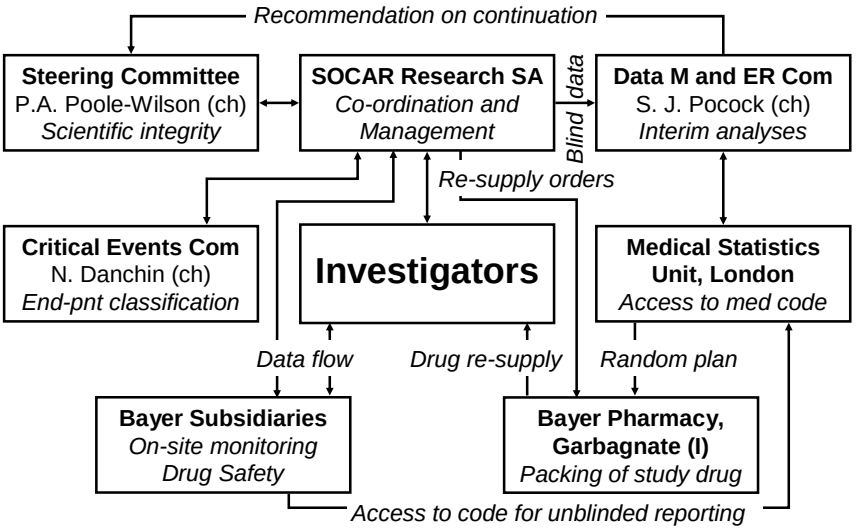


Figure 1: Organisational structure ACTION study

ch, chair; M, monitoring; ER, ethical review; Com, committee; med, medication. The random plan was prepared at the Medical Statistics Unit of the London School of Hygiene and Tropical Medicine, and was made available only for packing purposes to the pharmacy of the sponsor. Pre-specified interim analyses were performed at the Medical Statistics Unit, based on blinded data supplied by the coordinating centre (SOCAR Research SA, Nyon, Switzerland).

Organisational structure

A well-defined and workable organisational structure must be in place before the first patient is recruited. Roles, responsibilities and lines of communication must be known and no protocol can be finalised without delineating these *(recommendation 3)*. Of particular importance within the organisational structure is the role of the funding source. Today, the majority of clinical trials are industry sponsored. PICO and ACTION were no exception. Bodenheimer has outlined the many ways in which industry sponsoring can affect the trial conduct.[5] The organisational structure of ACTION is reproduced here as Figure 1 *(previous page)*. While this type of organisational structure is fairly standard for multi-centre trials, exact definitions of tasks and responsibilities may vary considerably. The initiative to perform the ACTION trial was taken by the Steering Committee chairman and our institution. The sponsor accepted that the study's credibility would be enhanced by reducing its own role to a minimum. The role and responsibility of each component part in the ACTION organisational structure as shown in Figure 1 was defined with this in mind, the objective being to set up an organisational structure which allows sensible decision making about all matters relating to trial conduct while at the same time retaining independence from the sponsor. Recently, editors of major medical journals have stressed the need to retain scientific independence.[6] Nonetheless, it must be recognised that funding sources, drug companies included, have legitimate interests that must be accommodated without jeopardising scientific integrity. Sponsoring a trial such as ACTION is useful to the drug company concerned only when the trial is conducted in such a manner that regulatory authorities and the scientific community at large will accept the results. This necessitates that the study is conducted in agreement with the many rules and regulations that nowadays apply. Also, the sponsor must be in a position to comply with its own adverse drug reaction reporting obligations, which vary from country to country.

Database management system

Apart from a workable protocol, CRF and organisational structure, trial management requires a database management system that serves its needs. These go beyond concurrent processing of data for interim and final statistical analysis. Equally important is its role as a trial management tool *(recommendation 4)*.

In Chapter 6 we have outlined the design philosophy and essential features of the database management system used to support trial management for the ACTION study. The amount of paperwork generated by such a study is enormous. To handle this efficiently, all ACTION CRFs and supporting documents were registered in the database upon receipt, scanned and then archived. Because of this, further database management could be based entirely on scanned images. Scanning also makes it possible to always display the scanned image and corresponding database content simultaneously on the user's computer screen. The content of the database is therefore continuously validated each time the database is consulted.

As described in Chapter 6, the system also supports all functions that are relevant to data management, such as generation of Data Clarification Forms (DCFs), tracking of implausible values, etc. In addition, the needs of trial management are supported. Particularly relevant in this regard is the possibility to flexibly implement tracking functions for listings that support on-site monitoring *(see below)*.

Data management for the PICO trial was handled by a traditional relational database application. In retrospect, it is difficult to imagine how a trial as large as ACTION could have been managed without the system described in Chapter 6.

Investigators' meetings

All too often, protocols, CRFs and other study documents are finalised without involving the investigators who will actually perform the trial. Both the protocol and CRF need a "reality check" before finalisation however, not just by investigators, but also by study nurses, coordinating centre staff and those who will be involved in on-site monitoring. For smaller trials such as PICO, we recommend organising an investigators' meeting in order to finalise the protocol and the CRF. For large trials such as ACTION this is usually impractical and the "reality check" must be left to the steering committee. Because of this, steering committees should not just consist of opinion-leaders, but also of practicing clinicians with clinical trial experience who represent the various geographical and linguistic regions involved.

While conducting both trials described elsewhere in this thesis, regular investigator meetings were held, which were also attended by study nurses, CRAs, etc. We believe that these meetings were needed while the study was ongoing not just to get feed back on problems that had occurred or to explain study procedures, but also to promote continued interest and motivation, the basis for successful completion of both PICO and ACTION *(recommendation 5)*.

On-site monitoring

That drug trials performed for regulatory purposes by the pharmaceutical industry need to comply with Good Clinical Practice (GCP) standards,[4] and that on-site monitoring is required to ensure this, is generally accepted. For studies done outside the industry, or non-drug studies, the situation is less clear because of the high cost of on-site monitoring.

On-site monitoring for industry-sponsored trials is the task of Clinical Research Associates (CRAs). CRAs are the main link between the coordinating centre and the clinical sites and robustness of the study data depends directly on the quality of their work. To perform their tasks successfully, CRAs must have the expertise required. Hence, trial management needs to ensure that CRAs are well trained, not just in GCP compliance in general, but also in all the procedures and policies that have been defined for a particular study. During monitoring visits, investigators often ask CRAs what to do in certain cases. If an inappropriate reply is given, an unwanted situation may occur that could otherwise have been avoided.

For the ACTION study, CRA meetings and special training courses were held repeatedly. This was necessary because of the trial's long duration and the high turn-over of CRAs. CRAs were also required to attend investigator meetings. Well-trained CRAs cannot compensate however for inadequate investigators or lack of local facilities. When there is no source documentation, verification of entries in the CRF is not possible. Centres who are unable to provide CRAs with the facilities required to perform their task should not participate in any trial *(c.f. recommendation 1)*.

Both the PICO and ACTION trials used classical CRFs printed on triplicate NCR paper. PICO investigators were instructed to mail the relevant CRF pages to the coordinating centre immediately after the patient was seen for a study visit. At the coordinating centre, each CRF was data entered upon receipt and a listing was made of missing or implausible items. Also, compliance with the protocol was checked. The CRA then took the results of these checks back to the centre during the next monitoring visit for clarification with the investigators. In this manner, time delays for data checking can be reduced to a minimum. More importantly, this approach allows the coordinating centre to notify investigators of any errors made without delay *(recommendation 6)*.

The sponsor of ACTION insisted that its standard on-site monitoring procedures be used. Completed CRF pages were checked on-site by CRAs and the necessary changes and additions were made by the investigator. Only then were CRFs sent to the coordinating centre. This caused time delays in data processing and checking as receipt of data was dependent on the frequency of monitoring visits. A temporary lack of monitoring capacity may cause major delays.

The tasks of CRAs can be made easier by supplying them with listings of CRFs that should have been received, data queries still needing to be resolved, documents such as discharge letters to be collected, etc. In fact, managing a study of the size and duration of ACTION is almost impossible without the possibility of generating such listings from the database (see also Chapter 6). For these listings to be useful, concurrent database management is essential. This is one more reason why the on-site monitoring frequency should not be allowed to determine the delay with which CRFs are received at the coordinating centre.

On-site Electronic Data Capture (EDC) is changing the way clinical trial data is collected. The considerations involved in CRF preparation and handling mentioned in this chapter also apply to EDC however. Hence, it remains to be seen whether on-site EDC will prove to be a step forward.

MANAGING RANDOMISATION AND BLINDING

Randomisation

That treatment allocation for a given trial was truly random can never be proven by statistical analysis. What matters is that a validated procedure, usually a computer program, is used to prepare a random plan based on choices that have been made about blocking, etc. and that this plan is then used to allocate treatment in such a

manner that the assumptions underlying randomisation are not violated. Trial management has to ensure that this is indeed the case. Essential to procedurally correct randomisation is that the investigator cannot know the treatment the patient is going to receive when the decision to include the patient in the study is taken. In an open study using sequentially numbered envelopes that contain randomly allocated treatments, this can usually not be guaranteed as it may be difficult to check that an investigator doesn't open the next envelope first, and then starts to look for a suitable patient for the treatment allocation contained in the envelope. Open studies therefore generally require a central treatment allocation system that allows the investigator to obtain a treatment allocation after the patient has been registered.

Study medication for double blind studies must be packed in such a manner that it is impossible to distinguish between treatments. It is important to verify before the study starts that this was done properly and that, for example, it is not possible to open coding envelopes without damaging them permanently, or to determine what is inside by holding them against a light source. The possibility to taste tablets needs to be considered also. We know of patients comparing the taste of tablets in waiting rooms. However, provided that double blinding of study medication is guaranteed, there is no need however to set up a central system for allocating treatment numbers. If blocking was used when preparing the random plan, it is only necessary to ensure that treatment numbers are used sequentially in order to prevent incompletely used blocks as much as possible.

Blinding of investigators

Whether investigators need to be blinded is a matter of trial design. It is sometimes argued that blinding is not needed for trials with 'hard endpoints'. Such trials often need blinding however in order to ensure validity of comparison for other outcomes, such as routine blood pressure measurements made by investigators using a standard office measuring device. The recent popularity[7,8] of the so-called PROBE (Prospective Randomised Open Blinded Evaluation) design[9] indicates in our opinion that the need for blinding of investigators may be underestimated. In the PROBE design, the committee that adjudicates the clinical events of interest is blinded to treatment allocation but the local investigator is not. Because of this, the decision of an investigator to send a patient to hospital for chest pain for example may very well be influenced also by knowing the treatment to which the patient was allocated. While the point that open treatment allocation can simplify clinical trials considerably[9] is well taken, blinding only the adjudication committee to study treatment allocation does not ensure validity of comparison based on our understanding of the principle involved *(recommendation 7)*.

For ethical reasons, the investigator or any other treating physician must be able to break the code when this is necessary to decide on further treatment. Rather than setting up a telephone system to monitor code breaks by investigators, we relied on traditional coding envelopes provided with study medication for both PICO and ACTION. Monitoring code breaks was left to on-site monitors and code breaks

were infrequent. We did explain extensively during investigator meetings that breaking the code is in fact rarely necessary to decide on further treatment when study medication is stopped for clinical reasons. We found that this is generally understood by investigators with experience in performing clinical trials.

Blinding the study structure

While there is general agreement about the indications for blinding investigators and patients, approaches to blinding the study structure seem to vary considerably. In our opinion, managing a double blind study requires that there is no possibility for anyone involved in data management or trial operations to access the code. We prefer therefore that the random plan is not accessible to coordinating centre staff before database lock. However, this has consequences for trial management. One consequence is that the coordinating centre is unable to perform interim analyses as specified in the protocol. For the ACTION trial, interim analyses were therefore performed at another institution by statisticians not otherwise involved in the trial (c.f. Figure 1). This institution also prepared the random plan before the study started, which was then made available only to the sponsor's pharmacy responsible for packing study medication. When a code break was required by the sponsor in order to comply with its adverse drug reaction reporting obligations, the sponsor's drug safety officer had to call the institution responsible for preparing the random plan. Because of this, it was possible to track the frequency and purpose of code breaks.

We believe that procedures for maintaining blinding of the study structure are more important for safeguarding scientific integrity than is generally realised *(recommendation 8)*. Scientific integrity requires that only pre-defined interim analyses are performed based on explicit stopping rules in order to avoid a trial being stopped early for inappropriate reasons, either by investigators, or by sponsors. The regulatory authorities have realised this for many years. Instances of a sponsor stopping a trial early for commercial reasons have in the past been surrounded by publicity and debate.[10-12] We know of instances where the investigator was prevented from publishing because the sponsor was unwilling to provide the random plan. Conflicts of interest affecting trial conduct are a cause for concern.[13] Although preparing the random plan for industry-sponsored drug trials is usually done by the sponsor, the latter should preferably not be involved in this. Ideally, access to the random plan should be controlled by an independent institution, as was the case in the ACTION trial.

FORMATION OF THE TRIAL COHORT

Centre-specific considerations

Timely patient recruitment is a prerequisite for any trial's success. To ensure the deadlines for patient recruitment are met, trial management needs to be up to date on a daily basis on the status of patient recruitment in each participating centre. In both the PICO and ACTION trials, investigators were required to notify the coor-

dinating centre immediately by telefax once a patient had given informed consent or had started study medication. This allowed continuous monitoring of progress and early identification of slow and fast recruiting centres alike. Today, the internet may be used for this purpose.

It is customary to define in the protocol or in the investigator contract, the minimum number of patients that each centre is expected to contribute. Blocked randomisation and stratification for centre being standard practice, this number is usually based on the minimum number of blocks to be used in each centre. Often, the maximum number of patients each centre can contribute is also defined. For the ACTION trial, the minimum was set in the protocol to 24 (4 blocks of six) and no maximum was set. Whether the minimum and maximum should be enforced is questionable. One could exclude from the analysis all patients contributed by centres that did not reach the minimum, and/or stop recruitment in centres when the maximum is reached. While neither of these would compromise validity of comparison, excluding centres that didn't reach the minimum is difficult to justify ethically and scientifically since patients in the centres concerned have been exposed to study treatment. Hence, no centre has been excluded from either the PICO or the ACTION trial because the number of patients recruited did not reach the block size. As far as maxima is concerned, faster recruiting centres compensated for slower recruiting ones in both trials, which ensured that the overall recruitment remained on target. This required that faster recruiting centres were re-supplied with sufficient study medication and other materials required, such as CRFs, etc. Managing the timely shipment from central stocks of trial supplies was therefore an important task in both trials *(recommendation 9)*.

A consequence of not enforcing centre minima or maxima may be that not all blocks are used completely. Centre contribution to each treatment arm will therefore tend to deviate from the allocation ratio used during blocked randomisation. This may decrease statistical power but should not compromise validity of the trial as a whole. When the number of centres is large, the effect of incomplete blocks will tend to 'randomise out' and treatment groups for all centres combined will be very similar in size if the random plan was based on blocks of an equal number of allocations to each treatment.

While deciding on trial cohort membership, the question of what to do when trial conduct appears to be unsatisfactory in a certain centre cannot be avoided. Exclusion of centres for such reasons is rarely mentioned in trial reports. An exception is the INSIGHT report, which mentions that "254 patients from centres withdrawn for misconduct" were excluded from intention-to-treat analysis.[14] Nonetheless, we find it difficult to believe that unsatisfactory local conduct is as rare as the absence of mentioning this in published trial reports suggests. As long as *all* patients from the centres concerned are excluded, validity of comparison in the remainder of the trial should not be compromised *(recommendation 10)*. While this is by itself a simple and methodologically sound policy, its actual execution is nonetheless complex and requires time, judgement and evidence. The trial coordinating centre

and/or the on-site monitors must determine which centres need to be audited, and these audits must then be performed with care *(recommendation 11)*.

The ACTION protocol contained a provision that allowed exclusion of centres. In a small fraction of centres, five in total, local trial conduct was found to be unsatisfactory by an independent auditor. We favour exclusion of centres that fail to deliver data of verifiable quality.

Who has actually been allocated treatment?

While recruitment is ongoing, trial management must first of all identify the patients that belong to the trial cohort based on an unequivocal operational definition of treatment allocation that can be applied in practice. The CONSORT group has written extensively about the need to describe clearly how participants were allocated to interventions. CONSORT has published a template for a diagram that is today part of many trial reports.[15] In the template, the terms 'randomised', 'allocated to intervention' and 'received allocated intervention' are used.

The distinction between 'allocated to intervention' and 'received allocated intervention' is essential in an open comparison between, for example, coronary bypass surgery and (continued) treatment with drugs. In such situations, 'allocated to intervention' should be taken as being equivalent to telling the patient which treatment he/she is going to receive.

Double blind trials such as PICO or ACTION are different from open trials. Double blind medication is packed in sequentially numbered containers and the actual tablet composition is concealed. Allocating a specific medication number to a specific patient does not necessarily mean that the first tablet from the allocated study medication pack is actually used. In many situations, a certain amount of time may pass between allocation of a study medication pack and actual start of treatment. This poses the problem as to what should be done in the case an allocated medication pack remains unused. One solution is to consider the patient in question a member of the trial cohort nonetheless. The logical consequence of this is that such patients must be followed as planned in the protocol. While we know of instances where this solution has been used, we consider it generally unjustifiable. Continuing patients who never started study medication in a trial poses ethical problems and may influence willingness to undergo follow up assessments. In accordance with ICH E9[1] it is our policy to omit from trial cohorts patients who failed to start double blind study medication even if a study medication number had been allocated. This should not compromise validity of comparison as the circumstances that caused this cannot have been influenced by a treatment that was never taken. It is our policy therefore to define treatment allocation in a double blind study as equivalent to taking the first tablet of study medication as this provides an unequivocal definition of trial cohort membership for such trials.

Dealing with violations of selection criteria

Ideally, all patients selected by investigators comply with the selection criteria in the protocol. These criteria often define a subgroup of patients free from other major clinical conditions or laboratory test abnormalities. The ACTION selection criteria (c.f. Chapter 4, Table 1 on page 43 of this thesis) are a case in point. In total 20 different reasons for exclusion were defined when the study was designed. Limiting recruitment to eligible patients only based on such complex criteria is rarely achievable in practice, in particular when exclusion criteria such as *creatinine above twice the local upper limit of normal* apply. From a trial management perspective, this is better replaced by *clinically relevant elevation of creatinine* as this leaves judgement to the investigator *(recommendation 12)*.

Unless ineligible patients can be detected before study treatment is started (which is difficult to implement in large multi-centre trials), trial management must determine eligibility status for each patient, and consider the steps to be taken if selection criteria have been violated.

One approach in this regard is to insist that patients who were allocated to treatment in violation of the selection criteria are considered members of the trial cohort nonetheless, and are hence also accounted for in the analysis. As noted in ICH E9, this policy implies "complete follow up of all randomised subjects for study outcomes".[1] This may be ethically unjustifiable, and also raises the question whether trial managers should advise investigators who started study treatment in violation of a selection criterion to withdraw study treatment in the interest of patient safety. An alternative approach is therefore to exclude ineligible patients from the trial cohort, notify the investigator accordingly as soon as the violation is detected, and advise the investigator to stop study medication at the same time.

There seems to be no agreement in the literature about the approach to be taken. The interpretation of the intention-to-treat principle is at issue here also. An inclusive interpretation is that ineligible patients in the trial cohort are apparently also candidates for the treatment investigated in the opinion of investigators and must therefore be included in a true intention-to-treat analysis. In addition, it is argued that removing ineligible patients could introduce bias, and is therefore inappropriate.[16] A drawback of including ineligible patients in the trial cohort is that the overall result of the study becomes a weighted average of the result in eligible patients and in patients who were not eligible. Depending on the percentage of ineligible patients and the effect of treatment in this category, the overall result may differ substantially from the result that would have been obtained had the trial cohort consisted only of eligible patients. Hence, an alternative exclusive interpretation is that the intention-to-treat principle applies only to eligible patients. Essential to this interpretation is the argument that there is no reason why excluding ineligible patients should compromise validity of comparison for those who were eligible, provided that the distinction between ineligible and eligible is made strictly based on patient characteristics established before study treatment was started.

Although proponents of the exclusive interpretation are vindicated by ICH E9 – which states that "failure to satisfy major entry criteria (eligibility violations)" is one of the "limited number of circumstances that might lead to excluding randomised subjects"[1] – we note that it is difficult to implement a policy of exclusion that is indeed strictly based on patient characteristics that were established before allocation to treatment. An example illustrating this problem could be the following. Suppose that a patient in a clinical trial was operated for prostate cancer one year after study medication was started, that the discharge letter stated that the patient was known to have prostate cancer for two years, and that the presence of this condition was not reported on the baseline CRF by the physician who recruited the patient. CRAs may be inclined to ask the investigator to add prostate cancer to the patient's baseline medical history. This added information is however biased because the presence of prostate cancer at baseline would in this case have never come to light, had the patient died suddenly six months after start of study medication. Obviously, it must be prevented that information which emerges after study medication was started influences the decision to exclude a patient from the trial cohort. A related issue is that the analysis of comparability at baseline of treatment groups as presented for the PICO trial (c.f. Chapter 2, Table 2, page 8), and analyses based on baseline characteristics (c.f. Chapter 5), are also affected by changing baseline data because of information that became available after start of follow up. A clear policy to prevent exclusion of ineligible patients on inappropriate grounds and to ensure validity of comparison for baseline data should be in place *(recommendation 13)*. This is just one example of the notion that maintaining validity of comparison in clinical studies is sometimes more important than complete representation of all that is known about a particular subject in the database.

The exclusive interpretation of the intention-to-treat principle mentioned earlier is, in our opinion, scientifically the only tenable one. Nonetheless, for reasons just mentioned, we have found it not easy to implement strict policies based on this interpretation. The protocol of the PICO trial required the coordinating centre to instruct the investigator to stop study medication in the case the investigator committed a major error of selection. The results section of the PICO trial report (c.f. Chapter 2, page 8) describes that 14 patients were excluded for violations of selection criteria before the medication code was broken. The trial design section of the same report states: "While the trial was ongoing, patients were withdrawn when a clinically relevant violation of the selection criteria was detected at the coordinating centre. The decision to do so was taken … before the medication code was broken." (c.f. Chapter 2, page 6). Taking such decisions before breaking the code in a double blind trial does not ensure validity of comparison however because excluding the just mentioned case of prostate cancer not reported at baseline compromises validity no matter when the code is broken. In the PICO trial, only a small number of ineligible patients were started on study medication. We are tempted to assume that one reason for this was the immediate notification of the investigator combined with an instruction to stop study medication. There was no room for argument about this as it was based on the provision in the protocol men-

tioned earlier and withholding remuneration for patients who don't comply with the selection acted as a deterrent *(recommendation 14)*.

Because it was impossible to reach consensus on a policy concerning exclusion of ineligible patients from the trial cohort after start of study medication, the design of ACTION was based on the inclusive interpretation of intention-to-treat. In the protocol, in the design paper (c.f. Chapter 4, page 51) and in the statistical analysis plan (SAP, c.f. Chapter 7, section 1.3, page 100) two analysis populations are defined, both to be analysed by assigned treatment. The *all-randomised population* consists of all patients who took at least one tablet of study medication irrespective of eligibility. The *valid-for-efficacy population* on the other hand is confined to eligible patients only, and is thus based on the exclusive interpretation of intention-to-treat. Because of the different views on the exact meaning of intention-to-treat, this term was not used in naming analysis populations. The inclusive view prevailed however as the ACTION protocol states in section 1.6: "The primary analysis will focus on all patients who were randomised (conventional 'intention-to-treat' analysis)."

The fact that the inclusive interpretation prevailed when the study was designed had considerable consequences for the conduct and management of the ACTION trial. The protocol required investigators to follow eligible and ineligible patients in the same manner and to maintain patients on study medication unless contra-indicated. It was therefore not necessary for trial management to distinguish between eligible and ineligible patients while recruitment was ongoing unless patient safety was compromised. In such instances, investigators were instructed to stop study medication while continuing follow up.

Although eligibility was as such inconsequential for trial conduct, it was nonetheless required to assess eligibility for each ACTION patient as this was the basis for defining the analysis populations mentioned earlier. Because of the large number of patients involved and the complexity of the selection criteria, this was no easy task which for several reasons could not be automated. All patients were therefore checked for eligibility by a data analyst at the coordinating centre. To avoid that minor violations led to exclusion from the valid-for-efficacy population, the criteria used for this purpose were less stringent than those mentioned in Chapter 4. For instance the criterion *in stable clinical condition for at least one month* (c.f. inclusion criterion no. 2, Table 1, Chapter 4, page 43) was relaxed by the Steering Committee for the purpose of eligibility checking to *in stable clinical condition for at least 10 days* (c.f. Chapter 7, section 1.3.2, page 100). ACTION was intended to focus on stable angina, and nifedipine is not indicated for acute coronary syndromes. The selection criterion concerned is a reflection of this. Nonetheless, one month of stability is as arbitrary as 10 days. This type of arbitrariness in clinical trial selection criteria is unavoidable. The result is a 'grey zone' in trial cohorts between patients who are perfectly eligible and those who are clearly outside the clinical spectrum at issue *(recommendation 15)*.

At the time of writing it is unknown whether the results of ACTION are different between eligible and ineligible patients. What can be said is that the sub-

group analysis that addresses this question will have little statistical power as only 6% of patients did not comply with the relaxed inclusion criteria as listed in section 1.3.2 of Chapter 7 (c.f. page 100). While we remain as a matter of principle committed to an exclusive interpretation of intention-to-treat combined with immediate removal from the trial cohort of patients who were started on study medication in violation of the selection criteria, it would not have been possible in retrospect to implement such a policy for the ACTION study. That certain selection criteria were not always perceived as clinically relevant by investigators didn't help. Another reason was that eligibility checking was a time consuming process, if only because the relevant sections of CRFs arrived at the coordinating centre weeks or even months after study medication had been started due to the frequency of on-site monitoring visits (c.f. *On-site monitoring,* page 149).

ENSURING INTEGRITY OF FOLLOW UP INFORMATION

The importance of defining the time span of follow up

Ensuring integrity of follow up requires firstly a practical definition of the planned time span of follow up. Deciding on the time point when follow up starts is closely related to deciding if the patient is a member of the trial cohort. In an open study, the moment that the patient is told which treatment he/she is going to receive should then also be considered as the start of follow up since everything that happens thereafter might have been influenced by the patient's knowledge of future treatment. In double blind studies, the logical starting point of follow up is the moment when the first dose of study medication is taken. To define this moment exactly, we insisted in both PICO and ACTION that the first tablet of study medication be taken in the presence of the investigator or the study nurse and that the clock time that this was done be recorded in the CRF. To illustrate why this is important for trial management, suppose that study medication was handed out on the day of treatment allocation and that the patient was instructed to start study medication the next morning. If such a patient was found dead the morning after study medication was handed out, it may be impossible for trial management to ascertain whether the patient had started study medication or not. Ambiguities of this kind can only be resolved without potentially compromising validity of comparison by considering all patients who were handed out study medication as belonging to the trial cohort even when this was never started.

Another important reason to document the exact time of start of follow up is that this allows data managers to determine which information was obtained before start of study medication and which information thereafter. To illustrate, suppose that the first tablet of study medication was documented in the CRF as having been taken at noon on a certain day. An ECG recorded before noon on the same day is then indeed a baseline ECG while an ECG recorded after noon is not, as the latter may already have been influenced by study treatment. An exact definition of the time that study medication was started is therefore a useful tool for management of baseline data *(recommendation 16).*

Apart from an unequivocal definition of the start of follow up, all trials require an exact and workable definition of its planned end. For clinical outcome trials such as ACTION, it is common to follow all patients until an arbitrary 'common stopping date'. However, choosing one common calendar date for this purpose has practical drawbacks. Patients still using study medication on the common stopping date will continue to do so until they can be seen by the investigator and the need for further treatment can be determined. Inevitably, clinical events will occur during continuation of study medication beyond the planned end-date. This may potentially lead to a discussion concerning inclusion of such events in the analysis. In an attempt to limit the number of such events, one can instruct investigators to see all surviving patients for a close-out visit as soon as possible after the common stopping date. This however leads to peak work loads for investigators, CRAs and data management personnel. To avoid such drawbacks in ACTION, we defined a six-month close-out period rather than a common stopping date and determined for each patient the planned date of the first regular ACTION out-patient clinic visit within this period. We then instructed investigators to consider this visit as the close-out visit for the patient concerned and to perform the visit as closely to the planned date as possible *but not before that date*. This approach limited continuation of study medication beyond planned end, and work loads tapered off rather than peaked.

Dealing with loss to follow up

Ensuring complete follow up that allows true intention-to-treat analysis (c.f. Chapter 3) is an important objective. What must be avoided is that time-to-event analyses have to be censored for any reason other than planned end of follow up so as to keep censoring non-informative. Similarly, it is necessary to ensure, clinical status permitting, that all planned measurements are available when comparing treatment groups for outcomes such as exercise capacity, blood pressure or NYHA class at specified time points during follow up. Otherwise, validity of comparison may be compromised *(recommendation 17*). In this paragraph, we describe some practical problems encountered while executing the PICO and ACTION trials and the policies implemented to maintain validity of comparison for follow up information.

For exercise capacity trials such as PICO, full follow up is not generally accepted standard practice. The primary analysis of such trials is often a per-protocol analysis of treatment effect on exercise capacity in patients who actually completed the study as planned on study medication – or any pre-defined part of it. In such protocols, one often finds criteria for considering a patient who belongs to the trial cohort as a 'drop-out', usually resulting in exclusion from analysis completely.

We believe that the very concept of 'drop-out' should not exist in any trial and that follow up assessments should always be fully completed as planned unless the patient is incapable of undergoing assessments for valid clinical reasons *(recommendation 18)*. In Chapter 3 we distinguish between pseudo and true intention-to-treat analysis. Data that haven't been collected can't be analysed. True intention-to-

treat analyses for PICO as presented in Chapter 2 (c.f. Figure 2 on page 12) and Chapter 3 (c.f. Table 3 on page 30) are possible only when the protocol required collection of the data needed. These analyses were crucial to the proper understanding of the effect of pimobendan on exercise duration as analysed per-protocol (c.f. Chapter 2, Figure 1 on page 11) because they showed that the positive effect of pimobendan on exercise duration was not completely negated by its negative effect on mortality.

Alleviating symptoms by treatment often implies an increased risk of morbidity and mortality that is associated with the treatment concerned. This confronts the patient with a choice between a shorter life combined with symptomatic improvement and a longer life combined with continuing or even more severe symptoms. Because of the risks associated with surgery, hip replacement is an example of such a treatment. Whether it is reasonable to even consider the choice can only be shown by true intention-to-treat analysis as performed for the PICO trial. Such analyses should therefore always be performed for trials that focus on symptoms or paraclinical measures as primary outcome (c.f. Chapter 3, Figure 1 on page 25). Such analyses may also be highly relevant to elucidate treatment effects in trials with clinical events or mortality as primary endpoint. In Chapter 3 we have outlined our thinking about how effects on symptoms and on clinical events can be analysed jointly in clinical outcome trials. The secondary analysis for NYHA class and mortality as described in section 4.3.3 of the ACTION SAP (c.f. Chapter 7, page 121) is a case in point. Such an analysis allows for the possibility that treatment reduces mean survival time, while increasing at the same time mean survival in NYHA class I (i.e. without cardiac symptoms). As in the case of PICO, this would indicate that taking the risk of reduced survival due to treatment might nonetheless be reasonable because of a positive effect on symptoms. Such analyses have been proposed a long time ago.[17] It is surprising that they haven't become more popular since they might have lead to different conclusions about the clinical value of heart failure treatments such as pimobendan.

Apart from allowing for elucidating trade-offs between positive and negative effects of treatment, there is another important reason for collecting true intention-to-treat data on morbidity and mortality in trials that focus on symptoms or paraclinical measures as primary outcome. Drug development is a sequential process with decision points that can be defined beforehand. Again, the PICO trial is a case in point. A compound such as pimobendan will not be approved by the regulatory authorities unless data from a large mortality trial are available. The decision to perform a new trial should also depend on an appropriate evaluation of the data that have already been collected during earlier trials. Evaluation of the effect on mortality using the combined data from earlier trials by a statistical meta-analytic method presupposes that the data on mortality have been collected based on intention-to-treat in all trials considered.[18] When this is not the case, the meta-analysis will not be relevant for the expected results from a large mortality trial.

For a number of reasons, complete follow up is not always easy to achieve even when required by the protocol. One reason may be that the protocol itself is

not entirely clear in this regard, or is later amended. For example, the initial version of the PICO protocol covered only the 24-week efficacy phase (c.f. Chapter 2 *Trial design*, page 6). For this phase follow up was almost complete; only 10 out of 317 patients didn't have an exercise test at 24 weeks for reasons that didn't relate to their clinical status (c.f. Chapter 2, Table 3, page 10). As the extended follow up phase (c.f. Chapter 2 *Trial design*, page 6) was added later, patients had to be asked consent for continued participation at the end of the efficacy phase. In PICO this was without consequence as only one patient was lost to follow up during extension. Nonetheless, the need to ask informed consent again because of study extension is better avoided as this can lead to problems *(recommendation 19)*. A case in point is a trial for which the published report states that "1588 (13%) of patients did not take part in the extension of the trial made necessary by the redefined primary endpoint (unable, or refused to restart or continue study medication)".[19] This in fact creates censoring at the start of extension that may not necessarily be non-informative. Hence, validity of comparison for the whole trial may be compromised.

Even when there are no later protocol amendments, problems with follow up occur. Patients may move to another city or country, or withdraw informed consent. In addition, investigators may equate refusal to continue study medication with refusal to comply with follow up *(c.f. recommendation 18)*. To minimise loss to follow up for such reasons, several policies were implemented for ACTION. Whenever a patient moved, an attempt was made to enlist the help of another physician in completion of follow up. That refusal to continue study medication is not the same as withdrawal of consent was explained at investigator meetings and to individual investigators as cases of refusal were reported. We also sent CRAs a monthly list of patients currently identified in the database as lost to follow up, and asked them to review the cases concerned with investigators during the next monitoring visit. Based on this, it was possible to bring several patients back under follow up and in some cases even to restart study medication. This was made possible by the protocol, which stipulated that study medication could always be restarted after interruption unless a contra-indication had developed. Finally, we encouraged investigators to invite patients who previously did not want to come back regularly for planned follow up visits for the end-of-study visit. If this was possible and included a review of the clinical history combined with SAE reporting for the purpose of ACTION endpoint detection, we considered such patients no longer as lost to follow up.

Despite all such efforts, loss to follow up can never be completely avoided as patients always have the right to withdraw informed consent. In a trial with total mortality as the primary evaluation criterion, loss to follow up for this endpoint can sometimes be limited because, for instance, a national registry can be accessed to ascertain the vital status of patients who withdrew consent. While loss to follow up can, and should be limited in this manner as much as possible in a mortality trial, this is not an option in a trial with a combined endpoint such as ACTION. The reason is that non-fatal events that terminate event-free follow up can usually not be

ascertained in this manner. For the ACTION trial, loss to follow up was therefore defined as the inability of the investigator to ensure further SAE reporting for the purpose of detection of study endpoints, even if it was possible to ascertain vital status at the planned end of follow up for the patient concerned. Because of this, it seems unavoidable that trials with a primary endpoint that combines fatal and non-fatal events should generally have a somewhat higher loss to follow up rate than mortality-only trials.

How should loss to follow up be accounted for in analysis and reporting? Our approach is not to accept loss to follow up as a reason for excluding the patient entirely from the trial cohort under any circumstance. As said earlier, there is no place for 'drop-outs' in this regard. Rather, the approach should be to always include all patients with less-than-complete follow up in time-to-event analyses, using the date that the patient was lost to follow up as censoring date. This is the approach that is outlined in the SAP for the ACTION trial (c.f. Chapter 7, section 4 starting on page 115). This de facto equates termination of follow up earlier than planned with follow up as planned, and assumes that censoring earlier than planned is non-informative. In a randomised trial, this assumption always holds under the null-hypothesis of 'no difference between treatments'. If the treatments compared are different however, non-informativeness of early censoring cannot be assumed even when treatment allocation was randomised. We note in passing that loss to follow up will also affect analyses for outcomes that are assessed repeatedly at certain pre-specified time points, such as NYHA class, blood pressure, etc. For such outcomes, our general approach is to include patients who were lost to follow up in the analysis with the data that have actually been collected, and to impute missing values based on a reasonable rule (c.f. the methods described in Chapter 2 for imputing missing exercise duration values).

Simultaneous participation in other studies

A problem rarely considered in trial reports is participation of patients in more than one study at the same time. Ancillary or 'side-arm' studies are a feature of many large trials as this may render participation more attractive to investigators with specific scientific interests. Whether such ancillary studies can be justified is another matter however.

In general, a distinction must be made between including a patient simultaneously in two different studies on the same day and including a patient already participating in one study in a second study at a later date. An example that illustrates the problems that can be caused by including a patient in two studies on the same day is as follows. Suppose that a group of investigators participating in a study such as ACTION decides to set up a study that requires baseline and follow up coronary angiography in 500 patients. If this was done as an ancillary study to ACTION, its overall results will be affected. Complications of angiography in the ancillary study become clinical events in the ACTION study that must be counted. While this would not compromise validity of comparison of ACTION as such,

generalisability may be affected because patients who undergo repeated angiography for no other reason than participation in the ancillary study will be treated differently. Provided that the main trial was sufficiently large, the question whether generalisability was affected by simultaneous participation in an ancillary study can be addressed by standard subgroup analysis because participation in the ancillary study is a baseline characteristic in the main study. This type of double participation therefore has a methodological solution at least in principle.

Including a patient who is already participating in one study in a second study later is infinitely more problematic. For the second study, participation in the first study is irrelevant for validity of comparison as this is just a baseline characteristic. Within the first study there is a problem however, as patients entered in the second study cannot be excluded without affecting validity of comparison. Considering patients who are included in the second study as 'lost-to-follow-up' is no option either, as this introduces potentially inappropriate censoring.

Based on these considerations, *participation in another trial or study* was an exclusion criterion in the ACTION study (c.f. Chapter 4, Table 1 on page 43). The protocol also contained a provision that while participating in ACTION, patients were not allowed to participate in any other trial, ancillary or side-arm study unless this other trial or study had been approved by the Steering Committee beforehand. The purpose of this was to prevent participation in any other study that would interfere with either the practicality of ACTION, or the eventual interpretation of its results. The rationale for this was repeatedly explained in detail at ACTION investigator meetings. In fact, several ancillary studies compatible with the main ACTION study were approved by the Steering Committee.

Notwithstanding this, at least 20 patients were included in other trials after inclusion in ACTION. One reason appeared to be that some investigators failed to see the problem, as was evidenced by explanations such as "there is no scientific reason why patients can't participate in another trial or study at the same time", or "this patient wasn't taking ACTION study medication anymore in any case" or "it wasn't me who did it, a colleague of mine entered the patient concerned during an admission to our coronary care unit without telling me". While 20 cases may seem an insignificant number, it must be taken into account that this concerns only the cases that we were able to detect. It cannot be excluded therefore that participation in other trials may have been more prevalent than we were able to assess *(recommendation 20*).

Participation in two trials at the same time is not possible for ethical and legal reasons. Hence, we were forced to define a policy for dealing with this even though the number of cases concerned in ACTION was small. As soon as we became aware of a case of inclusion in another study, the investigator was requested to remove the patient from the other study if possible. The only exception was a newly diagnosed cancer for which treatment was not available other than within a trial comparing several cancer treatments. If the investigator didn't agree to remove the patient from the other study, the latter was withdrawn from ACTION until the other study was completed. Thereafter we insisted that ACTION be resumed, including

restarting study medication unless contra-indicated. If this was not possible, we were left with no other option than to consider the patient as lost to follow up as of the moment of inclusion in the other study. As far as we know, less than 10 patients were lost to follow up for this reason.

ASSESSMENT OF OUTCOME

In clinical trials, effect estimation is based on comparing pre-specified outcomes between treatment groups by appropriate statistical methods. Valid assessment of the outcomes that will be analysed must therefore be ensured while the study is ongoing.

In Chapter 3 we distinguished four hierarchical levels of outcome information (c.f. Chapter 3, Figure 1 on page 25). Levels 1 – 3 concern clinical events that may occur at any time point during follow up. Level 4 outcomes on the other hand concerns observations that can be repeated over time in the same patient, such as blood pressure, NYHA class etc. The distinction between timed clinical events and observations that can be repeated is important in this context too. Effect estimates for clinical events will be diluted when, given validity of comparison, there is random misclassification.[20] Lack of precision because of random error in observations that can be repeated has the same effect. Reducing misclassification and random error in measurements is therefore an important concern. Standardisation is the key concept in this regard. In addition, generalisability of trial results must be kept in mind. A valid effect estimate for a certain outcome is of little scientific value unless the outcome concerned was obtained by a validated and generally-accepted method.

Timed events

There are several approaches to clinical outcome detection. One is to rely completely on the investigators to classify outcomes, based on uniform definitions in the protocol. Another approach, used in the majority of trials, is to rely on event adjudication by a special committee. While the names of the members can usually be found in the appendix of a trial report, the actual methods and procedures used are rarely described in detail. In this regard a distinction must be made between an adjudication procedure that is driven by the diagnosis of the investigator, and a procedure which is not. This distinction is important because of the potentially different event rates that may result. To illustrate the distinction, suppose that details of hospitalisations are only sent to the adjudication committee when the investigator's diagnosis was acute myocardial infarction. While this ensures that all cases of hospitalised infarction counted in the analysis fulfil the criteria that the committee used, the total number can only be less than the number diagnosed by investigators. This was observed in the TRIM trial, one of the few trials that published details on the adjudication procedure.[21] Hospitalised infarcts that were missed by investigators can be detected by sending to the adjudication committee details on *all* hospitalisations, and having the committee determine by standard criteria if acute myo-

cardial infarction occurred during hospitalisation. Such a procedure is not driven by the investigator's diagnosis.

For the ACTION trial, a Critical Events Committee (CEC) was formed to adjudicate events in a manner that was investigator diagnosis-independent as much as possible. Briefly, the procedure was as follows. Standard forms for reporting SAEs were used to allow the sponsor to comply with its reporting obligations of adverse drug reactions. In order to limit the number of forms that investigators had to complete, the same SAE form was also used to document clinical events and procedures that were ACTION study endpoints. This by itself caused a major trial management problem *(recommendation 21)*. Essentially the same SAE form is used by most if not all industry-sponsored drug trials. Hence, investigators are used to completing such forms whenever an adverse experience or event is noted that fulfils the standard criteria for seriousness. That in the case of ACTION completion of the same SAE report form was required also to document ACTION endpoints therefore needed to be explained in detail not just to investigators, but also to the CRAs and to personnel of the sponsor's drug safety department.

In ACTION, data on SAE reports was also the starting point for event adjudication. Unless it was certain that an event did not require adjudication by the CEC (for example removal of a wart under local anaesthesia), a so-called *Request for Additional Information* (RAI) was generated by means of a special database management module. This document requested the investigator to supply standard items of information and documentation (discharge letters, laboratory reports, ECGs, etc.) as required by the CEC. The information supplied by the investigator was scanned and entered in the ACTION database management system (c.f. Chapter 6). This allowed the automatic generation of documentation that was sent to the CEC for adjudication based on standard diagnostic criteria.

ACTION is a large study with a relatively long follow up in patients who all have coronary disease. Hence, the total number of SAEs to be expected is large and investigator diagnosis-independent adjudication becomes a major task. Whether this results in a higher event rate is another matter however. For example, in a preliminary analysis there were 656 investigator-diagnosed episodes of acute or procedural myocardial infarction. Of these, 80% (527 out of 656) were confirmed by the CEC. In addition, the CEC diagnosed 108 cases of infarction not diagnosed as such by the investigator concerned. Thus, the total number of episodes diagnosed as infarction by the CEC was 527 plus 108 or 635, which is similar to the total number of infarcts diagnosed by the investigator. Based on this, the need for a procedure as used in ACTION may be questioned. It must be stressed however that myocardial infarction is a clearly delineated diagnosis on which few cardiologists will err. Other diagnoses of interest in ACTION, such as *refractory angina requiring emergency coronary angiography without progression to myocardial infarction*, or *overt heart failure requiring hospitalisation (or occurred during hospitalisation) which led to start of, or change in, heart failure treatment* (c.f. Chapters 4 and 7), can in our opinion only be diagnosed reliably by a procedure that is essentially in-

vestigator diagnosis-independent, and is based on clearly defined standard diagnostic criteria that are applied by a committee of experienced cardiologists.

Many recent trials, ACTION included, use a combined primary endpoint and one could argue the relevance of adjudicating further events after termination of primary endpoint-free follow up. Such an argument ignores the fact that there may be other combined endpoints, with different component clinical events. For example, a patient who had a peripheral revascularisation procedure is not 'at risk' anymore for the primary efficacy endpoint of ACTION, but is still 'at risk' of the combined safety endpoint death, myocardial infarction and stroke. Hence, adjudication of further events after peripheral revascularisation is required. This is not the only point however. The total number of each individual event will generally also need to be reported. Consistency requires that this is done based on the same criteria that were used for diagnosing the first component event of a combined endpoint.

Other observations and measurements

Repeatable assessments such as exercise capacity, blood pressure or NYHA class must by definition be left to the investigator, who generally must be blinded for treatment allocation to ensure validity of comparison unless a special measurement method is used (such as a random-zero blood pressure measuring device). There are a variety of ways to measure exercise duration or blood pressure however, and investigators will differ concerning the exact interpretation of the NYHA classification. The generally accepted solution to ensure precision and generalisability is to standardise the methods of investigation across centres as much as possible and to perform the trial only in those centres that are familiar with the standard chosen. A case in point is exercise testing in the PICO trial. Only centres routinely using bicycle ergometry were invited to participate. The ergometry protocol to be used was extensively discussed with the investigators before the trial started and is summarised in Table 1 of Chapter 2 (page 5). The CRF developed for the PICO trial contained detailed step-by-step instructions for exercise testing based on these discussions *(c.f. recommendation 2)*.

Sometimes it is possible to ensure validity of comparison, achieve better standardisation and presumably also improved precision, by re-analysing locally made recordings at a central facility by uniformly applied standards. Examples are central analysis 24-hour ECGs (as done in PICO, c.f. Chapter 2), and central assessment of ejection fraction (as done in ACTION for the baseline ejection fraction, c.f. Chapter 4). Central coding of abnormalities in standard 12-lead ECGs has been used in many studies. The same is true for central quantitative analysis of coronary angiograms, as used for example in the MAAS trial[22] that was coordinated by our institution.

An issue related to standardisation is choosing which assessments to use in the analysis when comparing treatment groups for the evolution over time of outcomes such as NYHA classification, blood pressure, etc. Just as it is necessary to define the end-date of follow up in such a manner that censoring in time-to-event analyses

is non-informative (c.f. *The importance of defining the time span of follow up*, page 158), it is necessary to use only planned assessments of blood pressure, etc. Using non-planned assessments may compromise validity of comparison because effects of study medication on any of the outcomes to be compared may induce extra visits to the investigator. To ensure validity of comparison for repeated follow up observations at pre-specified time points, the ACTION SAP specifies the assessments to be used in detail (c.f. Chapter 7, sections 3.2 and 3.3 on page 113). The purpose is to favour the use of observations obtained at planned rather than at extra visits. A minor but nonetheless important comment in this regard is that investigators must be instructed to always calculate the planned date of a follow up visit relative to start of study medication, rather than the previous visit. This avoids that time windows allowed around planned visit dates accumulate, or are influenced by study medication.

ANALYSIS AND REPORTING

Analysis and reporting starts when the database has been declared 'clean' and is locked so that no further changes or additions are possible. It is a trial management task to plan in detail for all activities related to database lock. Event adjudication must be complete. All data queries must have been resolved. For those still outstanding, a decision must be taken if database lock must wait until the data concerned has been clarified. Coding of medical terms must be complete and accuracy of data entry must have been validated. For a double blind study, the study medication code should only be added after database lock. Otherwise validity of comparison may be compromised during data cleaning.

Even when analysis and reporting is not amongst its tasks and responsibilities, trial management must be aware of what analysis and reporting will require. Any group of investigators will want to publish the results as soon as the trial has been completed. The statistical analyses to be done, and how the trial will be reported, derive from its objective and design, not from its results. A well developed and detailed statistical analysis plan should therefore be available well before database lock and can play a major role in finalising the database. Trial management can make use of analysis programming before database lock to recheck that there is no unexplained missing or inconsistent data, that there are no implausible values in the database which have been overlooked, etc.

The ACTION SAP is reproduced in Chapter 7, starting on page 93. This document contains detailed specifications of most analyses to be done and the variables to be used, and determines how ACTION will be reported. It is the Steering Committee's role to decide how this will be done and Steering Committee input while preparing a SAP must therefore be ensured.

Influenced by the CONSORT statement, trial reporting in major journals follows nowadays more or less the same general pattern. CONSORT provides little guidance however as to how results should be tabulated and considerable variation in this regard remains.[23] In the same issue of a major journal one may find two

similar studies published back-to-back,[9,14] using a different way of reporting essentially the same combined endpoint.

Treatment effects for chronic disease trials are nowadays almost uniformly reported as hazard ratios even if this term is not mentioned. Few trials however also report the absolute underlying hazards (Wing et al.[24] provide a recent example), thus making it impossible for instance to calculate meaningful numbers-needed-to-treat.[25] The ACTION SAP specifies that this will be done.

Not all trials using combined endpoints report endpoint-free survival or which individual endpoint occurred first.[23] For ACTION both will be reported, as well as the total number of each individual event. That there are patients for whom the ACTION adjudication committee diagnosed refractory angina, acute myocardial infarction, stroke, heart failure and cardiac death in sequence reflects clinical reality. This reality will be fairly represented in the way ACTION will be analysed and reported.

Analyses of mean event-free survival subdivided by time spent in each NYHA class are currently not part of standard trial reporting, but will be reported for ACTION (c.f. Chapter 7, section 4.3.3, page 118).

While the ACTION SAP specifies the customary univariate subgroup analyses that are a standard feature of trial reports (c.f. Chapter 7, section 4.4.2, page 126), an analysis stratified by absolute risk at baseline based on a multi-attribute risk score is also included (c.f. Chapter 7, section 4.4.1, page 124). We know of just one report of a major trial in cardiovascular medicine that has included this in the past.[26] Risk stratification is important because a treatment that has a positive effect at high risk may have no effect, or even a negative effect, at low risk.[26,27]

Ideas such as these are best developed while the trial is still ongoing so that trial management can ensure that the necessary data are represented appropriately in the data sets upon which the analysis will be based.

RECOMMENDATIONS

1. Multi-centre trials that span several linguistic areas should be conducted in English and centre selection for such trials must take this into account. For industry sponsored studies, centres should be selected based on quality considerations (facilities required, compliance with GCP, availability and accessibility of source date, etc.) only, and should not be influenced by marketing.
2. Case Report Forms (CRFs) should be formulated in an unambiguous and clear manner, and should follow normal clinical procedures. Relevant instructions for carrying out examinations and tests should be printed on the opposite pages so that these are visible during CRF completion.
3. Trial management requires an unequivocal mandate based on tasks, responsibilities and lines of communication defined in the protocol.

4. Database management systems should be set up in such a manner that study management tools, such as listings for CRAs and investigators of outstanding study documents, dates of visits to be performed, etc. are also supported.
5. Regular meetings attended by all research staff involved should be a standard feature of multi-centre trials, in particular for trials which take several years to complete.
6. The completed CRF for each visit should be checked at the coordinating centre with the shortest possible delay so that investigators can be notified promptly errors in trial conduct. This should not be the task of on-site monitors (CRAs).
7. The need for double blinding must also be considered when the primary objective of a trial concerns 'hard endpoints'. Validity of comparison for clinical diagnoses requires that investigators are blinded even when adjudication of events is done by a committee that is blinded to treatment allocation. The same applies to outcomes such as routine blood pressure measurements.
8. Access to the study medication code should be limited to the statistician who performs interim analyses. Access to the code for adverse drug reaction reporting purposes should limited as much as possible and be monitored centrally.
9. Faster recruiting centres should be allowed to make up for slower recruiting ones and no limit should be imposed on the maximum number of patients that each centre can recruit. Trial supply management should take this into account.
10. All patients from centres where irregularities in trial conduct were observed should be removed from the trial.
11. Auditing of centres while a trial is ongoing is essential.
12. Selection criteria should either be enforced verbatim, or should refer to conditions that are present in the opinion of the investigator.
13. No additions should be made to sections of the case report form that concern baseline characteristics based on information that becomes available after start of follow up. Updates to baseline characteristics for this reason should be made on a separate CRF page and be kept separately in the database. Updated information should not be used in statistical analyses that use baseline characteristics, and should not be a ground for exclusion from intention-to-treat analysis.
14. In the case investigators are remunerated on a per-case basis, withholding remuneration for patients who were started on study medication in violation of the selection criteria should be considered.
15. A policy must be defined for handling ineligible patients who were nonetheless allocated study treatment. This policy must define whether such patients are a member of the trial cohort, will be followed-up in the same manner as eligible patients and will be maintained on study treatment.

16. In a double blind trial, the first dose of study medication should be taken in the presence of the investigator or study nurse. The exact date and clock time of ingestion should be documented in the case report form. In an open trial, the moment that the patient is informed about the treatment allocated should be similarly documented. Actual ingestion of the first dose of study medication (double blind trial), or actual communication of treatment allocated (open trial) should define trial cohort membership. The exact times mentioned earlier define the moment that follow up starts.
17. Validity of comparison for outcomes requires that the maximum duration of follow up for clinical events or the timing of repeated assessments to be used in treatment comparisons, does not depend on treatment allocated.
18. There is no justification for allowing 'drop-outs' during follow up. Clinical status permitting, patients who belong to the trial cohort should undergo all planned assessments irrespective of withdrawal of study medication. Refusal to continue study medication need not imply withdrawal of consent. The reasons for missing assessments should be recorded and information on clinical events should be collected based on intention-to-treat. This policy should apply also to trials that focus primarily on the evolution of symptoms or paraclinical measures while study medication is being administered.
19. Unless this contingency is excluded at the outset, the protocol should state that treatment and follow up may be prolonged for specified reasons (such as a lower endpoint rate than expected) and the need to ask for informed consent for this should be avoided. Hence, patients should be informed about possible extension during the initial informed consent procedure.
20. Patients cannot participate in two trials at the same time unless this is approved in advance by the steering committees involved.
21. To avoid confusion among investigators, standard forms that are used also in other trials (e.g. serious adverse event reports) should be used only for their intended purpose.

REFERENCES

1. The European Agency for the Evaluation of Medicinal Products. Human Medicines Evaluation Unit. Note for guidance on statistical principles for clinical trials. ICH topic E9, CPMP/ICH/363/96, London, 1998 (www.emea.eu.int/pdfs/human/ich/036396en.pdf, accessed on February 4, 2004)
2. Merriam-Webster's Online Dictionary. www.britannica.com (accessed on February 4, 2004)
3. Meinert CL. Clinical trials, overview. In: Biostatistics in clinical trials (Redmond C, Colton T, eds). John Wiley & Sons Ltd, Chichester, 2001

4. The European Agency for the Evaluation of Medicinal Products. Human Medicines Evaluation Unit. Note for guidance on Good Clinical Practice. ICH topic E6. CPMP/ICH/135/95, London, 2002 (www.emea.eu.int/pdfs/human/ich/013595en.pdf, accessed on February 25, 2004)
5. Bodenheimer T. Uneasy alliance – clinical investigators and the pharmaceutical industry. *N Engl J Med* 2000;**342**:1539-44.
6. Davidoff F, DeAngelis CD, Drazen JM *et al*. Sponsorship, authorship, and accountability. *N Engl J Med* 2001;**345**:825-6
7. Hansson L, Hedner T, Lund-Johansen P *et al*. Randomised trial of effects of calcium antagonists compared with diuretics and beta-blockers on cardiovascular morbidity and mortality in hypertension: the Nordic Diltiazem (NORDIL) study. *Lancet* 2000;**356**:359-65
8. Dahlof B, Devereux RB, Kjeldsen SE *et al*. for the LIFE Study Group. Cardiovascular morbidity and mortality in the Losartan Intervention For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol. *Lancet* 2002;**359**:995-1003
9. Hansson L, Hedner T, Dahlöf B. Prospective Randomized Open Blinded Endpoint (PROBE) study: a novel design for intervention trials. *Blood Press* 1992;**1:**113-9
10. Lievre M, Menard J, Bruckert E *et al*. Premature discontinuation of clinical trial for reasons not related to efficacy, safety, or feasibility. *BMJ* 2001;**322**:603-5
11. Evans S, Pocock S. Societal responsibilities of clinical trial sponsors. Lack of commercial pay off is not a legitimate reason for stopping a trial. **BMJ** 2001;**322**:569-70
12. Ashcroft R. Responsibilities of sponsors are limited in premature discontinuation of trials. *BMJ* 2001;**323**:53
13. Horton R. The Dawn of McScience. The New York Review of Books. Volume 51, Number 4; March 11, 2004 (www.nybooks.com/articles/16954, accessed on February 25, 2004)
14. Brown MJ, Palmer CR, Castaigne A *et al*. Morbidity and mortality in patients randomised to double-blind treatment with a long-acting calcium-channel blocker or diuretic in the International Nifedipine GITS study: Intervention as a Goal in Hypertension Treatment (INSIGHT). *Lancet* 2000; **356:** 366-72
15. Altman DG, Schulz KF, Moher D *et al*. for the CONSORT (Consolidated Standards of Reporting Trials) group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;**134**:663-94.
16. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;**319**:670-74.
17. Olsson G, Lubsen J, van Es GA *et al*. Quality of life after myocardial infarction: effect of long term metoprolol on mortality and morbidity.*Br Med J* 1986;**292**:1491-3.

18. Lubsen J. Mega-trials: is meta-analysis an alternative? *Eur J Clin Pharmacol* 1996;**49**(Suppl 1):S29-33
19. EURopean trial On reduction of cardiac events with Perindopril in stable coronary Artery disease Investigators. Efficacy of perindopril in reduction of cardiovascular events among patients with stable coronary artery disease: randomised, double-blind, placebo-controlled, multicentre trial (the EUROPA study). *Lancet* 2003;**362**:782-88.
20. Copeland KT, Checkoway H, McMichael AJ *et al*. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977; **105**:488-95
21. Näslund U, Grip L, Fischer-Hansen J *et al*. The impact of an end-point committee in a large multicentre, randomized, placebo-controlled clinical trial: results with and without the end-point committee's final decision on endpoints. *Eur Heart J* 1999;**20**:771-77.
22. MAAS investigators. Effect of simvastatin on coronary atheroma: the Multicentre Anti-Atheroma Study (MAAS). *Lancet* 1994;**344**:633-38. Erratum in: *Lancet* 1994;**344**:762
23. Poole-Wilson PA, Lubsen J. Losartan for cardiovascular disease in patients with and without diabetes in the LIFE study. *Lancet* 2002;**359**:2199
24. Wing LM, Reid CM, Ryan P *et al*. for the Second Australian National Blood Pressure Study Group. A comparison of outcomes with angiotensin-converting enzyme inhibitors and diuretics for hypertension in the elderly. *N Engl J Med* 2003;**348**:583-92
25. Lubsen J, Hoes A, Grobbee D. Implications of trial results: the potentially misleading notions of number needed to treat and average duration of life gained. *Lancet* 2000;**356**:1757-9
26. The MIAMI Trial Research Group. Metoprolol in acute myocardial infarction (MIAMI). A randomised placebo-controlled international trial. Eur Heart J 1985;**6**:199-226
27. Hoes AW, Grobbee DE, Lubsen J. Does drug treatment improve survival? Reconciling the trials in mild-to-moderate hypertension. *J Hypertens* 1995; **13**:805-11

Chapter 9

Summary / Samenvatting

Summary

In a clinical trial, patient groups assigned to different treatments are compared to assess the effects of a new treatment or to evaluate the difference between existing treatments. To avoid systematic distortion by other factors, patients are randomly allocated to the treatments compared. Both patients and treating physicians are also blinded for treatment allocation when required by the trial's objective. All patients are followed using standardised methods of observation.

As such, a clinical trial is a procedural concept. Whether randomisation (and blinding when used) was done correctly cannot be answered by statistical analysis. Neither is it possible to correct for missed or unreliable patient observations. Hence, published trial results can only be trusted when the trial was conducted using appropriate methods and procedures. This requires appropriate trial management from beginning to end. Using examples of procedures implemented for one completed and one ongoing study, this thesis aims to put trial management into perspective.

By way of example of a completed study, the main results publication of the PICO (PImobendan in COngestive heart failure) trial is reproduced as **Chapter 2**. Pimobendan is a positive inotropic agent that might potentially benefit patients with chronic heart failure. PICO was a randomised, double blind, placebo controlled trial to compare two different dosages of pimobendan and placebo. The 317 patients involved were followed for at least 24 weeks. The effects of pimobendan were evaluated by exercise capacity and clinical outcome. Both dosages of pimobendan improved exercise duration by a similar amount as observed in earlier trials, but mortality was higher in patients receiving pimobendan.

A treatment that increases mortality may nonetheless be clinically useful if it improves well-being to such an extent that the increased mortality risk is acceptable. By a special analysis of exercise capacity data that took mortality into account also, it could be shown that the positive effect of pimobendan was not completely negated by the higher mortality risk. This analysis was possible because, based on the protocol, trial management implemented a policy of continued data collection on exercise testing and clinical follow up for patients who had stopped PICO study medication earlier than planned (a non-standard design feature for this type of trial).

The thinking on which the analysis of PICO exercise data was based is further developed in a publication on combined endpoints reproduced as **Chapter 3**. The use of combined endpoints has become widespread. Combining all-cause mortality with selected non-fatal events is useful because event-free survival can be ad-

dressed in this manner. Further examples are given as to how data on outcomes such as exercise capacity can be combined with data on mortality.

As an example of an ongoing trial, the publication on the design of the ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) study is reproduced as **Chapter 4**. Its purpose is to assess the effect of long-acting nifedipine (a calcium channel blocker) on major cardiovascular event-free survival in patients with stable angina pectoris. Data on baseline characteristics are given in a publication reproduced as **Chapter 5**.

Although follow up is relatively long (4 – 6 years), ACTION is not a particularly difficult trial for the investigators because only routine clinical follow up is required. Managing ACTION is however a highly complex task as almost 7,000 patients from just under 300 centres in 19 countries are participating. The ACTION database management system is described in **Chapter 6**. All documents containing patient data (case report forms, laboratory test reports, electrocardiograms, data clarification forms, etc.) are logged, scanned and archived when received at the coordinating centre. Data entry and cleaning is done using scanned images. A novel feature is that scanned image and database content always appear simultaneously on the computer screen. In addition, the system supports the special tracking functions and management tools needed to successfully complete this study to high quality standards.

The approved ACTION statistical analysis plan is reproduced as **Chapter 7**. This plan describes in detail the variables to be used, the statistical analyses to be done and how the results will be tabulated. Standard statistical methods for clinical outcome trials will be used but several of the pre-specified analyses are not commonly reported (mean survival subdivided by symptomatic state, stratification by absolute risk at baseline based on a multi-attribute risk score).

The general discussion (**Chapter 8**) describes how the most important problems trial management encountered during the execution of the PICO and ACTION trials were solved. Each problem required a solution based on the scientific principles underlying clinical trials that could also be implemented practically. The chapter ends with a list of recommendations for handling selected aspects of trial conduct, chosen because there is no general agreement in the literature concerning the approach to be followed (i.e. who actually belongs to the trial cohort), because of their particular importance for scientific integrity (i.e. how to deal with loss to follow up), or because we found them to be controversial (i.e. simultaneous participation in several trials).

Perfection may be difficult to achieve and may be even utopia. Because of their impact on patient treatment and health care costs, methodological diligence and an eye for detail is a must when it comes to the design and conduct of clinical trials.

Samenvatting

Teneinde na te gaan wat het effect is van een behandeling, of om verschillen tussen behandelingen te evalueren, worden in een experimenteel klinisch onderzoek (klinische trial) groepen patienten met elkaar vergeleken die zijn toegewezen aan verschillende behandelingen. Om systematische verstoring door andere factoren te voorkomen wordt de behandeling van iedere patient door het toeval aangewezen (randomisatie). Wanneer de doelstelling van de trial dit vereist worden zowel de behandelend arts als de patient onkundig gelaten ten aanzien van de toegewezen behandeling (dubbele blindering). Alle aan de trial deelnemende patienten worden bestudeerd met behulp van gestandaardiseerde methoden voor klinisch onderzoek.

Als zodanig is een klinische trial een procedureel concept. Of randomisatie (en dubbele blindering indien van toepassing) correct is uitgevoerd kan niet door statistische analyse van de bevindingen worden vastgesteld. Ook is het niet mogelijk om te corrigeren voor afwezige of onbetrouwbare observaties. Gepubliceerde resultaten van klinische trials zijn daarom alleen betrouwbaar wanneer het onderzoek van begin tot eind effectief werd geleid en correct werd uitgevoerd. Doel van dit proefschrift is het leiden van de uitvoering van trials te plaatsen in het perspectief van de grondslagen ervan. Daarbij wordt gebruik gemaakt van voorbeelden uit de praktijk afkomstig van één inmiddels afgerond, en één nog aan de gang zijnd onderoek.

Als voorbeeld van een afgerond onderzoek worden in **hoofdstuk 2** de eerder gepubliceerde resulaten van de PICO (PImobendan in COngestive heart failure) studie beschreven. Voor patienten met chronisch hartfalen is pimobendan een mogelijk heilzaam positief inotropicum. In een gerandomiseerde, dubbel-blinde trial werden twee verschillende pimobendan doseringen vergeleken met placebo. In totaal werden 317 patienten gedurende tenminste 24 weken geobserveerd. De werkzaamheid van pimobendan werd geëvalueerd aan de hand van inspanningsproeven en het klinisch beloop. Beide pimobendan doseringen hadden een positief effect op de inspanningstolerantie vergelijkbaar met eerdere bevindingen. De sterfte was echter hoger bij patienten behandeld met pimobendan.

Een behandeling die de sterfte verhoogt kan desondanks van waarde zijn als deze het welbevinden van de patient in dusdanige mate verbetert dat het hogere sterfte risico acceptabel is. Middels een speciale analyse van de inspanningsproefbevindingen die tevens rekening houdt met sterfte, kon worden aangetoond dat het positieve effect van pimobendan op de inspanningstolerantie niet geheel teniet werd gedaan door de hogere sterfte. Deze analyse was mogelijk omdat de onderzoeksleiding, zoals voorzien in het onderzoeksprotocol, er op toe zag dat de in het onderzoeksplan opgenomen inspanningsproeven en de observatie van het klinisch beloop werden gecompleteerd óók bij patienten die de onderzoeksmedicatie voortijdig moesten staken (niet standaard voor dit type trial).

De gedachtengang waarop de analyse van de PICO inspanningstolerantiegegevens is gebaseerd wordt verder ontwikkeld in een publikatie over ‘gecombineerde eindpunten’ (**hoofdstuk 3**). Kinische trials maken tegenwoordig veelal

gebruik van gecombineerde eindpunten. Het combineren van totale strefte en geselecteerde ziektegebeurtenissen in één eindpunt is zinvol omdat dit evaluatie van overleving vrij van de betreffende ziektegebeurtenissen mogelijk maakt. Meerdere voorbeelden betreffende het combineren van gegevens verkregen uit observaties zoals inspanningsproeven, en gegevens over sterfte, worden beschreven.

Als voorbeeld van een nog niet afgerond onderzoek is de publikatie over de opzet van de ACTION (A Coronary disease Trial Investigating Outcome with Nifedipine GITS) studie opgenomen als **hoofdstuk 4**. Doel van dit onderzoek is het bepalen van het effect van een langwerkende formulering van de calciumantagonist nifedipine op de overleving zonder belangrijke cardiovasculaire ziektegebeurtenissen van patienten met stabiele angina pectors. Gegevens over de kenmerken van de patientenpopulatie bij toelating tot het onderzoek worden beschreven in een publikatie opgenomen als **hoofstuk 5**.

Alhoewel dit onderzoek wordt gekenmerkt door een relatief lange observatieduur (4 – 6 jaar) is ACTION niet bijzonder veeleisend voor de deelnemende onderzoekers omdat uitsluitend routine observaties nodig zijn. Het leiding geven aan de uitvoering ervan is daarentegen zeer complex vanwege het grote aantal betrokken patienten (bijna 7,000) en centra (bijna 300 in 19 landen). Het ACTION databeheerssysteem wordt beschreven in **hoofdstuk 6**. Ontvangst op het coördinatiecentrum van patient-documentatie (standaard formulieren met gegevens, uitslagen van laboratoriumonderzoek, electrocardiogrammen, wijzigingsformulieren, etc.) wordt geregistreerd in het databestand. Vervolgens wordt ieder document gescand en opgeslagen in het databestand, en tenslotte gearchiveerd. Gescande documenten dienen als basis voor het invoeren en opschonen van gegevens. Vernieuwend is de wijze waarop gescande documenten en gevens aanwezig in het databestand automatisch tegelijkertijd op het computerscherm van de gebruiker worden weergegeven. Het systeem maakt het tevens mogelijk de zoekfuncties te implementeren die de onderzoeksleiding nodig heeft om aan de hoge kwaliteitseisen die worden gesteld te kunnen voldoen, en om dit onderzoek met succes te kunnen afronden.

Het goedgekeurde ACTION statistische analyse plan is opgenomen als **hoofdstuk 7**. In dit op voorhand vastgestelde plan worden de gegevens die zullen worden geanalyseerd, de te gebruiken statistische methoden en de wijze waarop de resultaten zullen worden weergegeven in detail beschreven. Standaard statistische methoden voor het analyseren van het klinische beloop in trials zullen worden gebruikt. Verscheidene beschreven analyses zijn echter niet algemeen gebruikelijk (gemiddelde totale overlevingsduur onderverdeeld naar symptomatische toestand, stratificatie voor het absolute risico bij toelating tot het onderzoek aan de hand van een op meerdere kenmerken gebaseerde risico score).

In de algemene discussie (**hoofdstuk 8**) wordt beschreven hoe de belangrijkste problemen waarmee de onderzoeksleiding zich tijdens de uitvoering van PICO en ACTION geconfronteerd zag, werden opgelost. Voor ieder probleem moest een praktisch uitvoerbare oplossing worden gevonden die recht doet aan de wetenschappelijke basis van klinische trials. Dit hoofdstuk wordt afgesloten met een

reeks aanbevelingen hoe de beschreven problemen op te lossen of the voorkomen. Daarbij werden keuzes gemaakt, hetzij gebaseerd op de overweging dat in de literatuur geen eenstemmingheid bestaat over de aanbevolen aanpak van een bepaald probleem (zoals de vraag welke patienten feitelijk deel uitmaken van het onderzoekscohort), hetzij omdat het probleem van bijzonder belang is voor de betrouwbaarheid van een klinische trial (zoals de aanpak van incomplete observatie van het ziektebeloop), hetzij omdat het desbetreffend probleem controversieël bleek te zijn (zoals deelname door dezelfde patient aan meerdere klinische trials tegelijkertijd).

Perfectie is moeilijk te bereiken en is veelal een utopie. Resultaten van klinische trials hebben belangrijke gevolgen voor de behandeling van patienten en bepalen mede de kosten van de gezondheidszorg. Oog voor detail en aandacht voor methodologie zijn daarom noodzakelijk bij het opzetten en uitvoeren van klinische trials.

A word of thanks

I would like to express my gratitude to all those who made it possible for me to accomplish this thesis. In particular, I am greatly indebted to my promoters, Prof. Koos Lubsen and Prof. Bert Hofman. Prof. Lubsen's many valuable comments, input and continuous guidance have been of major importance for the completion of this thesis. I also wish to express my thanks to Prof. Hofman for his continued interest and helpful advice, and his many contributions to my MSc in Clinical Epidemiology degree at the Netherlands Institute for Health Sciences.

Special gratitude is due to Fred J. van Dalen whose unfailing, incessant support and patience has contributed to the completion of this thesis.

I wish to thank Pieter Jonkers for his never-ending patience, help and IT support during the conduct of both the PICO and ACTION trials. I would also like to thank Nicolas Danchin, Peter Dunselman, Philip A Poole-Wilson and Gilbert Wagener for their support and their input on clinical issues. In addition I wish to thank John and Annemiek van Rossum for their permanent assistance in all logistic aspects.

For the final co-ordination of the many details involved in obtaining a doctorate at the Erasmus University, I wish to express my particular appreciation to Το Κυμα. I thank my colleagues Thierry Pauchard and Louise Pollard-Lichtsteiner for their diligent reviews, helpful comments and never-ending search for typos, and Gisèle Lubsen for making the cover photographs.

Finally, I am indebted to the patients, the site research teams (investigators, co-investigators, study nurses and co-ordinators) who participated in both the PICO and ACTION trials. Without them, there would have been no thesis.

About the author

Bridget-Anne Kirwan was born on January 17th, 1963 in Waterford, Ireland. After secondary school education at the Convent of the Sacred Heart, Ferrybank, Waterford, she graduated in 1982 from the Waterford Regional College with a degree in Health Science (distinction), and in 1985 as a Registered Nurse from the University Hospital, Cork. From 1985 to 1987, she held various nursing positions in the University Hospital, Cork. In 1987, she moved to Lausanne, Switzerland to work at the University Hospital, where she held various senior nursing positions. In 1993, she joined SOCAR Research, a clinical research institute, in Nyon, Switzerland, as a Clinical Research Associate. While working at SOCAR, she obtained a MSc in Clinical Epidemiology at the Erasmus University, Rotterdam in April 2000. Currently, she is Director of Clinical Operations at SOCAR and Project Director of the ACTION study, a clinical trial in which 300 centres in 19 countries participate.