Efficacy of a homeopathic *Crataegus* preparation compared with usual therapy for mild (NYHA II) cardiac insufficiency: results of an observational cohort study

D. Schröder[a], M. Weiser[b,*], P. Klein[c]

[a]*Börsenstrasse 17, Frankfurt/M, Germany*
[b]*Biologische Heilmittel Heel GmbH, Baden-Baden, Germany*
[c]*Datenservice Eva Hönig GmbH, Rohrbach, Germany*

Abstract

Objectives: To compare the efficacy of the homeopathic *Crataegus* preparation Cralonin for non-inferiority to standard treatment for mild cardiac insufficiency. *Methods:* Multicentre non-randomised cohort study in patients aged 50–75 years in New York Heart Association class II. Patients received Cralonin ($n=110$) or ACE inhibitor/diuretics ($n=102$) for 8 weeks. To adjust for confounding by baseline factors, populations were stratified according to propensity score. After adjusting, there were no statistically significant differences between treatment groups. Treatment efficacy was assessed on 15 variables. A stringent non-inferiority criterion for the upper limit of the 97.5% one-sided confidence interval of the treatment difference was set to $0.2\times$ the standard deviation (S.D.). *Results:* Both treatment regimens improved scores on most variables studied, with the greatest effect on double product after exercise (average score reduction 15.4% with Cralonin vs. 16.0% for the control group). Stringent non-inferiority of Cralonin was demonstrated on 7 variables. Medium-stringent ($0.5\times$S.D.) non-inferiority was indicated by 13 variables (exceptions: systolic blood pressure (BP) during exercise and diastolic BP at rest; for these, differences between treatments were not significant). Both treatments were well tolerated. *Conclusion:* The *Crataegus*-based preparation Cralonin is non-inferior to usual ACE inhibitor/diuretics treatment for mild cardiac insufficiency on all parameters except BP reduction.
© 2003 European Society of Cardiology. Published by Elsevier Science B.V. All rights reserved.

Keywords: Cralonin; Cardiac insufficiency; Homeopathy; Double product

1. Introduction

Complementary medicine is widely used in the developed world [1,2]. In particular, the use of and belief in the principles of homeopathy are widespread both in the US and in Europe [3–8]. However, the issue of whether there are real benefits from homeopathic treatment has not been conclusively resolved to date. Several reviews and meta-analyses of clinical trials agree that there seem to be benefits over placebo generally, but that more, rigorous and systematic, research is warranted [9–11]. However, many of the trials conducted to date have been of low quality and a general increase in the standards of trials would be beneficial to practitioners and patients alike.

The current study evaluates the efficacy of the homeopathic preparation Cralonin in mild cardiac insufficiency, New York Heart Association (NYHA) class II. The preparation is based on extracts from *Crataegus* and *Spigelia anthelmia* (wormbush). Cralonin is registered in Germany as homeopathic preparation (Registration No. 9054.00.00) and has a long and well-documented history of use for mild cardiac insufficiency [12,13].

Abbreviations: AE, adverse event; BP, blood pressure; DBP, diastolic blood pressure; DP, double product; HR, heart rate; NYHA, New York Heart Association; PS, propensity score; QOL, quality of life; SBP, systolic blood pressure; SD, standard deviation; SEM, standard error of mean; tid, *tris in diem*, thrice daily

*Corresponding author. Postfach 100349, 76484 Baden-Baden, Germany. Tel.: +49-7221-501-291; fax: +49-7221-501-660.
E-mail address: weiser.michael@heel.de (M. Weiser).

Preparation and administration of Cralonin follow the rules of homeopathy.

The study was designed to disprove inferiority of a Cralonin preparation to ACE inhibitor/diuretics therapy. Focus was on clinical symptoms as observed by the practising physician and the patients themselves, not on underlying cardiac parameters.

In the case of Cralonin, there is a real risk that the subset of patients, who are willing to be randomised to treatments as widely different as an established mainstream therapy and a homeopathic medication, exhibit important differences from the target population [14]. Also, homeopathic remedies are prescribed to a very wide range of patients and treatment is highly individualised, with the possibility of altering medication during the treatment regimen. For these reasons, the study used a non-randomised approach and applied the established methodology of propensity score (PS) analysis to construct matched strata that balance observed co-variates [15–18]. This allowed us to include a broad range of populations in both the Cralonin and control groups. A multivariate analysis was not carried out as this method is not applicable to the demonstration of non-inferiority using one-sided confidence intervals.

2. Methods

This was a multicentre, non-randomised cohort study assessing the non-inferiority of Cralonin to ACE/diuretics therapy. The study was carried out in 27 centres in Germany between July 1 and December 31, 2000. A total of 216 patients were enrolled. All patients were informed about the background and purpose of the study, which was conducted in full compliance with the principles of the Declaration of Helsinki (Br. Med. J. ii (1964) 177) and in accordance with the German 'Recommendations for the planning, performance, and evaluation of postmarketing clinical studies' (Bundesanzeiger Federal Gazette) No. 229 of December 12, 1998.

2.1. Inclusion criteria

Men or women aged 50–75 years, with diagnosed mild cardiac insufficiency NYHA class II, necessitating therapy, but not currently undergoing treatment with either Cralonin drops or ACE inhibitor/diuretics. Patients were outpatients, with or without (stable) hypertension (systolic blood pressure, SBP > 140 mmHg, diastolic blood pressure, DBP > 90 mmHg).

2.2. Exclusion criteria

Unstable coronary heart disease, concomitant cardiac therapy different from study medication, and intolerance toward any of the study treatments. Patients currently on either treatment therapy were also excluded. How-

ever, earlier therapy with either study drug was not a criterion for exclusion.

2.3. Study design

As only patients currently not receiving therapy were included, there was no washout period. Patients received either Cralonin drops (Biologische Heilmittel Heel GmbH, Baden-Baden, Germany) thrice-daily (tid) or ACE inhibitor/diuretics treatment. The dosage for each patient was at the administering practitioner's discretion. The Cralonin preparation consists of pro 100 ml: Crataegus Ø (mother tincture), 70 ml; Spigelia anthelmia D2, 1 ml; Kalium carbonicum D3, 1 ml; ethanol 45% (v/v).

Each patient was followed-up for 8 weeks, with data collected at baseline, at week 4 and at end of study. Treatment efficacy was evaluated on heart rate (HR), blood pressure (BP), double product (DP; evaluated on a bicycle ergometric test and defined as $HR \times BP/100$ where HR is heart rate in bpm and BP blood pressure in mmHg), symptoms (fatigue, listlessness, dyspnoea under strain, pretibial oedema, rapid exhaustion), frequency of nocturnal urinations and exercise tolerance (distance walked and number of stairs ascended without fatigue).

2.4. Measurements

DP was measured at rest and after a 2-min exercise at 50 W. Fatigue, listlessness, performance reduction, dyspnoea under strain and pretibial oedema were evaluated on a scale from 0 to 3, where 0—no difficulties and 3—major difficulties. The walking test assessed the distance the patient was able to walk on level ground without fatigue on a scale from 1 to 6, where 1—<100 m; 2—100–300 m; 3—300–500 m; 4—500–900 m; 5—1000 m (in ~15 min); 6—further than 1000 m (in >15 min). The staircase test evaluated the number of stairs the patient was able to walk without fatigue on a scale from 1 to 7, where 1—<5 steps; 2—5–10 steps; 3—11–15 steps; 4—16–20 steps; 5—21–25 steps; 6—26–30 and 7—>30 steps. Global treatment results were assessed by the practitioner on a scale ranging from very good, good, moderate, no effects to negative development. Tolerability was assessed by recording adverse events (AEs) and by the practitioner's assessment of global tolerability (very good, good, moderate or low). Compliance was assessed by the practitioner as very good, good, moderate or low.

2.5. Statistical methods

As this was a non-randomised cohort study, the principal investigator had no control over the treatment assignment and there might have been large differences

Table 1
Stratification of subjects according to PS

Group	Cralonin (mean PS = 0.66)		ACE inhibitor (mean PS = 0.37)		Total	
	N	%	N	%	N	%
PS < 0.3	11	10.00	52	50.98	63	29.72
0.3 < PS < 0.55	15	13.64	24	23.53	39	18.40
0.55 < PS < 0.7	28	25.45	10	9.80	38	17.92
0.7 < PS	56	50.91	16	15.69	72	33.96
Total	110	100.00	102	100.00	212	100.00

in observed co-variates between the treatment groups. Hence, the direct comparison of treatment effects might be confounded by a number of baseline characteristics. A method of adjusting for treatment differences between co-variates and to reduce bias, is by using a PS, as described by Rosenbaum [16]. PS is a description of the conditional probability of receiving treatment given the observed co-variates. As shown by Rosenbaum and Rubin, PS is a balancing score and is applicable to observational studies to reduce bias, allowing for the application of standard statistical methods [15]. Patients with approximately the same PS value are similar in observed co-variates independent of whether they are treated with test treatment or control treatment and treatment effects can be expected to be largely unbiased by confounding parameters. It has been calculated that, as PS balances all co-variates that are used to calculate PS, division into five strata will eliminate approximately 90% of the bias of each of the co-variates [15,19].

PS was estimated for each patient using logistic regression (i.e. the logarithm of the odds for the probability of receiving Cralonin, $\log(p/(1-p))$, will be seen as linear function of observed co-variates) and patients were divided into four strata according to PS scores. A breakdown of the groups is shown in Table 1. After calculation of treatment effects within each PS stratum, overall treatment effect was calculated by weighted means of the stratum effects as described by Fleiss [17]. For details of the mathematics applied, see Appendix A.

All observed variables were used as underlying co-variates: weight, age, fatigue, listlessness, performance on walking test and staircase test, HR, duration of illness, dyspnoea under strain, DP, SBP and DBP, pre-tibial oedema and reduced overall performance.

Treatment groups were compared after adjustment for PS using a two-way ANOVA model for co-variates based on interval data and Cochran–Mantel–Haenszel test for co-variates with dichotomous values. Prior to stratification, treatment groups differed significantly on five co-variates; however, there were no statistically significant differences after adjustment for PS.

To compare treatment groups for non-inferiority of Cralonin vs. ACE inhibitors/diuretics, the adjusted differences (reduction Cralonin-reduction ACE inhibitors/ diuretics) between treatments were calculated with 97.5% one-sided confidence intervals. Except for the walk test and staircase test, negative treatment differences indicate superiority of Cralonin. The upper limits of the confidence intervals can be interpreted as boundaries for assessing non-inferiority and were compared with two commonly used 'benchmarks' for inter-group differences: small between-treatment difference (0.2 × standard deviation, S.D.) and medium difference (0.5 × S.D.) [20].

3. Results

3.1. Patients

A total of 216 outpatients were enrolled in the study. Four patients were excluded as they were already receiving one of the study medications, and the final analysis was carried out on 212 patients, 110 in the Cralonin group and 102 in the ACE inhibitor/diuretics group. As shown in Table 2, the main reasons for cardiac insufficiency were coronary heart disease, cardiomyopathy, vitium cordis and hypertension.

Of the study population, 110 received Cralonin drops tid and 102 received standard therapy for mild cardiac insufficiency, consisting of ACE inhibitor/diuretics. Most patients in the Cralonin group (80.0%) received the standard dosage of 20 drops tid; 15.4% received 10 drops tid. The control medication was given as monotherapy or combination therapy, at the discretion of the prescribing practitioner. Of the patients in the control group, 52.0% received ACE inhibitors (benazepril, captopril, cilazapril, fosinopril, lisinopril, perindopril or ramipril), 6.9% diuretics (hydrochlorothiazide, furosemide, torasemide, indapamide or triamteren) and 41.2% a combination of both. ACE inhibitors/diuretics were given at doses commonly used in clinical practice; however, doses varied between individuals. Mean treatment period in the Cralonin group was 66.5 days, ranging from 33 to 132 days. The control group was treated for a mean of 65.2 days (32–157 days).

Unadjusted baseline demographic data were comparable for both groups for age and weight, but there was a difference in sex distribution between groups (Table 3). After adjusting for PS, however, differences were no longer statistically significant (Table 3).

Baseline values for efficacy variables were similar between groups (Table 2), with a few exceptions: more patients in the control group were hypertensive (defined as SBP > 140 mmHg, DBP > 90 mmHg) at baseline (72.5 vs. 54.5% in the Cralonin group) and earlier therapy was more common in the control group (64.7 vs. 26.4% in the Cralonin group). These unadjusted

Table 2
Baseline criteria with significance levels before and after PS adjustment

Variable	Cralonin mean±S.D. (n)	Control mean±S.D. (n)	Unadjusted (test result)	Adjusted (test result)
Pretreated (%)	26.4 (110)	64.7 (102)	**	ns
Coronary heart disease (%)	48.2 (110)	49.0 (102)	ns	ns
Vitium cordis (%)	1.8 (110)	1.0 (102)	ns	ns
Nocturnal urinations (%)	81.8 (110)	87.3 (102)	ns	ns
Cardiac myopathy (%)	10.9 (110)	5.9 (102)	ns	ns
Hypertension (%)	54.5 (110)	72.5 (102)	**	ns
Risk factors present (e.g. obesity, smoking, diabetes mellitus) (%)	84.5 (110)	85.3 (102)	ns	ns
HR (bpm)	78.84±8.888 (105)	78.06±10.71 (100)	ns	ns
Fatigue[a]	1.81±0.669 (104)	1.82±0.737 (102)	ns	ns
Listlessness[a]	1.59±0.719 (104)	1.61±0.810 (102)	ns	ns
Duration of illness (yrs)	3.02±1.521 (109)	3.14±1.485 (99)	ns	ns
SBP (mmHg)	146.9±16.24 (109)	150.2±16.04 (100)	ns	ns
DBP (mmHg)	86.43±8.985 (109)	88.04±10.36 (100)	ns	ns
Pretibial oedema[a]	0.82±0.769 (105)	1.01±0.850 (102)	ns	ns
Walk test[b]	3.95±1.245 (106)	3.81±1.376 (101)	ns	ns
Staircase test[c]	3.95±1.298 (110)	3.90±1.432 (102)	ns	ns
Dyspnoea under strain[a]	1.74±0.724 (104)	1.83±0.772 (102)	ns	ns
Reduced overall performance[a]	1.68±0.624 (108)	1.88±0.708 (102)	*	ns
Increase in DP[d]	70.01±33.57 (104)	74.87±39.97 (100)	ns	ns
Nocturnal urinations (n/night)	2.0±0.86 (88)	2.0±0.77 (89)	ns	ns

*$0.05 > P > 0.01$; **$P < 0.01$; ns$= P > 0.05$.

[a] Measured on a scale of 0–3 where 0—no difficulties and 3—major difficulties.

[b] Distance the patient is able to walk on level ground without fatigue; 1—<100 m, 2—100–300 m, 3—300–500 m, 4—500–900 m, 5—1000 m in approximately 15 min, 6—further than 1000 m (in >15 min).

[c] Number of stairs the patient is able to walk without fatigue; 1—<5 steps, 2—5–10 steps, 3—11–15, 4—16–20, 5—21–25, 6—26–30, 7—>30 steps.

[d] min^{-1} mmHg/100.

differences were significant on χ^2-test. However, after adjusting for PS the differences were not shown to be significant (Cochran–Mantel–Haenszel test controlling for stratum). The most common earlier therapies in the Cralonin group were nitrates (10.0%) calcium-channel blockers (7.3%) and diuretics (6.4%). In the control group, most common earlier therapies were ACE inhibitors (37.3%) diuretics (24.5%) and calcium-channel blockers (8.8%).

Baseline BP, HR and performance test scores did not differ significantly between treatments (Table 2), but overall performance was more reduced in the control group, which also tended to have a higher rate of pretibial oedema than the Cralonin group.

3.2. Treatment effects

Both treatments had beneficial effects on most variables studied. Changes in BP, HR and DP are shown in Fig. 1. Marked improvements with both treatments were seen in DP after exercise. Cralonin reduced average scores by 15.4% (from 183.4±39.37 min^{-1} mmHg/100 before treatment to 155.2±37.6 min^{-1} mmHg/100) after 8 weeks, compared with a reduction of 16.0% (from 194.6±43.25 to 163.4±36.92) in the control group.

Benefits from treatment were also seen in both groups on most other criteria. On walk tests and staircase tests, there was a trend towards better scores in the Cralonin

Table 3
Baseline demographics with test results before and after PS adjustment

Variable	Cralonin mean±S.D. (n)	Control mean±S.D. (n)	Unadjusted (test result)	Adjusted (test result)
Age (yrs)	68.5±7.85 (110)	65.6±9.06 (101)	ns	ns
Weight (kg)	76.3±11.91 (109)	76.5±12.74 (101)	ns	ns
Sex (%)				
Male	29.1	47.1	*	ns
Female	70.9	52.9	*	ns

*$P < 0.01$; ns$= P > 0.05$.

group than in the control population (walk test, Cralonin mean improvement 0.8, control 0.6; staircase test, Cralonin mean improvement 1.3, control 1.0). The average number of nocturnal urinations likewise was reduced to a similar extent in both groups, from 2.0 to 1.2. Both treatments reduced fatigue, listlessness and dyspnoea under strain. Score reductions for these criteria were 0.3–1.0 points in both groups, from baseline values in the mild-to-moderate range (1–2). Pretibial oedema (baseline scores 0.8 and 1.0, i.e. 'mild') were reduced by a mean of 0.6 points by both treatments.

3.3. Between-treatment differences at end of study

Fig. 2 summarises adjusted differences in outcomes between the Cralonin and control groups for the 15 criteria evaluated. The non-equivalence hypothesis for a variable was considered disproved if the upper limits of confidence intervals for treatment differences fell within one of two limits: a stringent limit of $0.2 \times$ S.D. and a medium limit of $0.5 \times$ S.D. Using the stringent limit, non-inferiority was demonstrated on 7 out of 15 variables. If the medium difference interval of $0.5 \times$ S.D. was used, non-inferiority was inferred on 13 of 15 variables. Intervals crossed the $0.5 \times$ S.D. boundary only for the criteria SBP during exercise and DBP at rest. However, the differences between treatments were not significant in these cases.

Global assessments of treatment results were somewhat more favourable to Cralonin, with 28.2% judging the results as 'very good' (15.7% in the control group) and with similar percentages judging the results as 'good' (58.2% for Cralonin, 52.0% for ACE inhibitors/ diuretics; $P = 0.002$ for the overall comparison between treatments).

Both treatments were very well tolerated, but the percentage of patients with tolerability evaluated as 'very good' was significantly higher for Cralonin than for the control medication (82.7 vs. 46.1%, $P < 0.0001$).
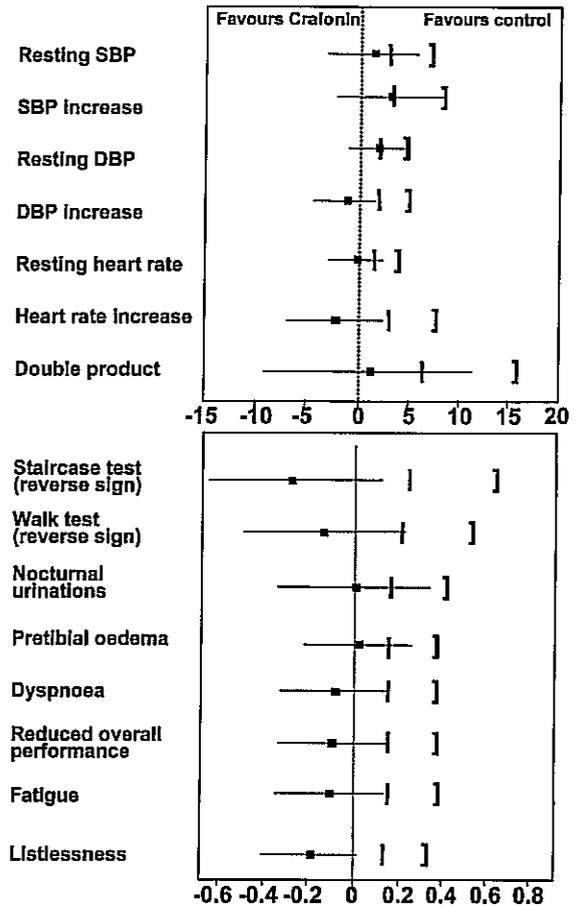
Fig. 2. Adjusted differences between treatments (Cralonin vs. ACE inhibitor/diuretic therapy) with 97.5% one-sided confidence intervals. Vertical bars, non-inferiority limit $0.2 \times$ pooled S.D.; brackets, non-inferiority limit $0.5 \times$ S.D. Note that for uniformity the values for walk and staircase tests are transformed by the factor-1 (reversed sign), as in the original tests, positive values favour Cralonin.

AEs occurred in 1 patient in each treatment group. With Cralonin there was one case of pressure in the heart region and with ACE inhibitor one case of dry cough
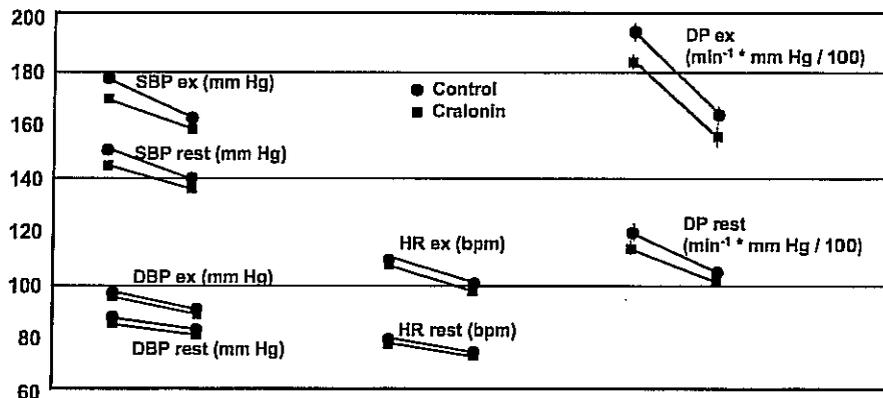
Fig. 1. Changes in BP, HR and DP values ($\pm$S.E.M.) at rest and after exercise (ex) from baseline to end of study for Cralonin (squares) and control groups (circles). S.E.M. values are greater than 2 only for the DP scores.

needing medical attention. Both AEs were considered possibly treatment-related, but none led to discontinuation of the study.

Compliance with treatment was good in both groups. Patients receiving Cralonin demonstrated a greater degree of compliance than the control group. Compliance with Cralonin was judged by practitioners as 'very good' in 57.3% of patients (37.3% in the control group, $P=0.007$ for the differences between groups) and 'good' in 40% (control group 55.9%).

4. Discussion

This study assesses the efficacy and tolerability of the homeopathic preparation Cralonin in patients with mild cardiac insufficiency NYHA class II. Cralonin treatment was shown to be non-inferior to standard ACE inhibitor/diuretics therapy on 13 out of 15 variables, the exceptions being SBP during exercise and DBP at rest. For staircase test and HR increase under exercise, the treatment effects tended towards superiority of Cralonin. As assessments were made at three 4-week intervals, it seems highly unlikely that the differences between the beginning and end of the study were due to a training effect.

In contrast to earlier reports on the efficacy of Cralonin [12,13], the current study is a direct comparison with standard treatment with ACE inhibitors/diuretics on effects on symptoms relevant to the patients' overall status.

The results may be considered controversial, as this is a trial of a homeopathic combination preparation. However, as reviews of clinical trials in homeopathy have concluded, homeopathy can and should be evaluated using the same standards as with allopathic treatments [10,11]. The present study fulfils criteria identified by Benson and Hartz [21] for observational studies able to yield valuable data, that studies shall assess differences between two treatments or between a treatment and no treatment, treatments shall be implemented by physicians and the study must include a control group.

The current study attempts to capture the actual practice by leaving the individualisation of treatment regimens to the respective practitioners. The makeup of populations willing to be randomised to homeopathic or standard treatments can be expected to differ from the general population. Additionally, randomised studies often exclude a significant proportion, between 9 and 51% of screened patients [22]. For these reasons, we decided to forgo the randomised trial in favour of a non-randomised cohort study.

In non-randomised studies co-variates must be balanced by statistical methods, if treated and control groups are to be comparable in the sense of having similar distribution of co-variables. We used PS adjustments [15,23] to construct matched strata that balance observed co-variates. Before adjustment, the baseline variables hypertension, female sex and history of previous treatment differed between treatment groups. However after PS adjustment, these differences were no longer statistically significant.

Recent surveys have challenged the perception that non-randomised studies tend to report greater effects from treatments than randomised trials. Benson and Hartz compared observational studies with randomised clinical trials in 136 cases and 19 treatment areas and found a very good agreement between results. Specifically, cardiological studies showed agreement between randomised and observational results in 6 out of 7 cases [21]. Similarly, the UK Health Technology Assessment Group [24] evaluated studies of 18 treatments, surgical, pharmaceutical and organisational, and concluded that there was no systematic bias in observational studies. Concato et al. came to similar conclusions in an analysis of five clinical topics and 99 reports, 44 of which were related to hypertension and coronary heart disease [25].

As has been pointed out [18], PS adjustment adequately balances observed co-variates but, unlike random assignment of treatment, it cannot balance co-variates that were not observed. However, surveys by Britton et al. and Benson and Hartz [21,24] indicate that this risk is not significantly higher in observational studies than in standard randomised clinical trials. Given the large number of co-variates included in our analysis, it appears unlikely that the risk of bias is larger than the risk of unintentional bias (e.g. non-random allocation of treatment) frequently present even in randomised trials [26].

One consequence of our study design was that the composition of the control medication was not homogenous. Half of the control population, (52.0%) received ACE inhibitor as monotherapy and 41.2% received a combination of ACE inhibitors/diuretics. This reflects the fact that the individual therapy was decided by the prescribing practitioner. This could be seen as a weakness, as outcomes in the control group might have been slightly different with standardised treatment. However, the composition of the control group reflects the treatment situation for cardiac insufficiency in general practice and the results in the control group arguably reflect the outcome of individually optimised treatments.

Another possible weakness is that the data were collected by the attending physician, which may allow for observational bias. This would be expected to be a greater problem with endpoints such as fatigue and listlessness, where evaluations are subjective to a degree. However, endpoints such as DP and HR, which are less susceptible to subjective influence, were very similar to the other endpoints in showing no significant differences between Cralonin and the control group (Fig. 2), which supports the limited conclusions drawn.

A decrease in DP indicates improved oxygen transport and lesser risk of cardiac complications. This is expected to translate into improved performance [27], as was indeed seen in our study in, e.g. staircase tests and walk tests. Similar improved performance scores have been reported from other studies with *Crataegus*-based therapy [28]. As moderate exercise is recognised as being beneficial in heart failure [29], increasing performance would be expected to lead to increased exercise and a positive feedback loop.

There are advantages with Cralonin that speak for the preparation as an alternative to ACE inhibitors/diuretics in mild cardiac insufficiency. Whereas Cralonin has an excellent tolerability profile, documented through long use and in an observational study in 2178 patients [12], ACE inhibitors/diuretics are associated with unwanted effects: cough in the case of ACE inhibitors [30] and reduced quality of life (QOL) with many diuretics [31,32]. Subjective reports on Cralonin from patients show favourable effects on QOL and effects such as reduced nocturnal urination would improve a patient's perceived QOL.

A good tolerability profile is particularly relevant in the case of cardiac insufficiency. Patients with only mild symptoms are unlikely to adhere to a regimen with noticeable side effects, whereas more severely afflicted patients are usually prescribed multiple drug regimens, where compatibility can be an issue. The compatibility of Cralonin with currently recommended medications indicates that the preparation can be safely added to existing drug regimens.

It would be extremely difficult to prove the superiority of a homeopathic preparation in an indication such as heart failure in the current treatment milieu, as it would be unethical to withhold effective treatment from patients in randomised clinical trials. A large, controlled study on *Crataegus* in patients with heart failure class NYHA II–III has recently been announced [33]. However, as this trial does not use a homeopathic preparation, the results may not be applicable to our study. Based on our indications of non-inferiority, and the well-established safety and tolerability record of Cralonin, a controlled trial where Cralonin is added to patients' usual therapy would seem both desirable and ethically defensible.

Appendix A: Propensity score analysis

The PS was estimated for each patient using the following logistic regression procedure in SAS:proc logistic data = base; model trtment = covariate1 covariate2 covariate3.../selection = forward; output out = PROP pred = prob;run;where base is the dataset containing all baseline variables including the co-variates in the model statement and prob is the predicted

PS. Four strata of PS groups were formed to get strata with at least 10 patients for each treatment group.

After stratification the best estimate of the common underlying difference between the 2 treatment means was calculated using a method published by Fleiss [17]:

$$d = \left(\sum w_i \times d_i \right) / \left(\sum w_i \right)$$

where $d_i = x_{i1} - x_{i2}$ denotes the difference between the two means within stratum i, and $w_i = n_{i1} \times n_{i2}/(n_{i1} + n_{i2})$ its weighting factor.

The one-sided 97.5% confidence limits for the underlying overall mean difference are calculated by

$$d + t_{n-8,0.025} \times \sqrt{\left(s^2 / \sum w_i \right)} \text{ (right) resp.}$$

$$d - t_{n-8,0.025} \times \sqrt{\left(s^2 / \sum w_i \right)} \text{ (left),}$$

where $s^2 = \left(\sum [(n_{i1} - 1) \times s_{i1}^2 + (n_{i2} - 1) \times s_{i2}^2] \right) / (n - 10)$ denote the pooled variance within strata and treatment.

References

[1] Eisenberg DM, Davis RB, Ettner SL, et al. Trends in alternative medicine use in the United States, 1990–1997: results of a follow-up national survey. J Am Med Assoc 1998;280:1569–75.

[2] MacLennan AH, Wilson DH, Taylor AW. Prevalence and cost of alternative medicine in Australia. Lancet 1996;347:569–73.

[3] Knipschild P, Kleijnen J, ter Riet G. Belief in the efficacy of alternative medicine among general practitioners in The Netherlands. Soc Sci Med 1990;31:625–6.

[4] Ernst E, Resch KL, White AR. Complementary medicine. What physicians think of it: a meta-analysis. Arch Intern Med 1995;155:2405–8.

[5] Vickers A, Zollman C. ABC of complementary medicine: homeopathy. Br Med J 1999;319:1115–8.

[6] Visser GJ, Peters L, Rasker JJ. Rheumatologists and their patients who seek alternative care: an agreement to disagree. Br J Rheumatol 1992;31:485–90.

[7] Chandola A, Young Y, McAlister J, Axford JS. Use of complementary therapies by patients attending musculoskeletal clinics. J R Soc Med 1999;92:13–6.

[8] Astin JA. Why patients use alternative medicine: results of a national study. J Am Med Assoc 1998;279:1548–53.

[9] Kleijnen J, Knipschild P, ter Riet G. Clinical trials of homeopathy. Br Med J 1991;302:316–23.

[10] Linde K, Clausius N, Ramirez G, et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. Lancet 1997;350:834–43.

[11] Cucherat M, Haugh MC, Gooch M, Boissel JP. Evidence of clinical efficacy of homeopathy. A meta-analysis of clinical trials. HMRAG. Homeopathic Medicines Research Advisory Group. Eur J Clin Pharmacol 2000;56:27–33.

[12] Zenner S, Borho B, Metelmann H. Biological cardiac therapy with cralonin: results of a study on multicentric application with 2178 patients. Ärztezeitschrift für Naturheilverfahren 1990;August:595–606.

[13] Weiser M, Gegenheimer L. Cralonin as adjuvant therapy in heart disease. Biol Med 2000;29:72–9.