Taylor, R. S. et al. (2018) Impact of exercise-based cardiac rehabilitation in patients with heart failure (ExTraMATCH II) on mortality and hospitalisation: an individual patient data meta-analysis of randomised trials. *European Journal of Heart Failure*, 20(12), pp. 1735-1743.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

This is the peer reviewed version of the following article Taylor, R. S. et al. (2018) Impact of exercise-based cardiac rehabilitation in patients with heart failure (ExTraMATCH II) on mortality and hospitalisation: an individual patient data meta-analysis of randomised trials. *European Journal of Heart Failure*, 20(12), pp. 1735-1743, which has been published in final form at http://dx.doi.org/10.1002/ejhf.1311. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

http://eprints.gla.ac.uk/180331/

Deposited on: 22 November 2018

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

Impact of exercise-based rehabilitation in patients with heart failure (ExTraMATCH II) on mortality and hospitalisation: an individual-patient data meta-analysis of randomised trials

Rod S Taylor[1], Sarah Walker[1], Neil A Smart[3], Massimo F Piepoli[4], Fiona C Warren[1], Oriana Ciani[1,2], [5]Christopher O'Connor, [6]David Whellan, [7]Steven J Keteyian, [8]Andrew Coats, [9]Constantinos H Davos, Hasnain M Dalal[1], [10]Kathleen Dracup, [11]Lorraine Evangelista, [12]Kate Jolly, [13]Jonathan Myers, [14]Robert S McKelvie; [15]Birgitta B Nilsson, [16]Claudio Passino, [17]Miles D Witham, [18]Gloria Y Yeh, [19]Ann-Dorthe O Zwisler, on behalf of the ExTraMATCH II Collaboration.

[1]Institute of Health Research, University of Exeter Medical School, EX1 2LU, UK: Rod S Taylor, chair of health services research; Sarah Walker, research fellow; Fiona C Warren, senior lecturer; Oriana Ciani, senior research fellow; Hayes Dalal, associate professor; [2]Centre for Research on Health and Social Care Management, Bocconi University, Milan, 20100, Italy: Oriana Ciani, senior research fellow; [3]University of New England, Armidale, NSW 2351, Australia: Neil Smart professor; [4]Cardiology Unit, Guglielmo da Saliceto Hospital, 29121 Piacenza, Italy: Massimo Piepoli, professor; [5]Duke Clinical Research Institute, North Carolina, NC 27705, USA: Christopher O'Connor, professor; [6]Department of Medicine, Sidney Kimmel Medical College, Philadelphia, PA 19107, USA: David Whellan, professor; [7]Department of Medicine, Henry Ford Hospital, Detroit, MI 48202, USA: Steven J Keteyian, professor; [8]IRCCS, San Raffaele, Pisana, Italy: Andrew Coats, professor; [9]Cardiovascular Research Laboratory, Biomedical Research Foundation, Academy of Athens, Athina 115 27, Greece: Costantinos H Davos, associate professor; [10]University of California San Francisco, School of Nursing, CA 94143, USA: Kathleen Dracup, professor; [11]University of California Irvine, CA 92697, USA: Lorraine Evangelista, professor; [12]Institute of Applied Health Research, University of Birmingham, B15 2TT, UK: Kate Jolly, professor; Western University,[13] St. Joseph's Health Care Centre, London, ON N6A 3K7, Canada: Robert S McKelvie, professor; [14]Palo Alto Health Care System/Stanford University, CA 94305, USA: Jonathan Myers, professor; [15]Division of Medicine, Oslo University Hospital and Faculty

of Health Sciences, Oslo Metropolitan University, N-0130, Norway: Birgitta B Nilsson, associate professor; [16]Fondazione G. Monasterio and Scuola Superiore Sant'Anna, 56127 Pisa, Italy: Claudio Passino, professor; [17]School of Medicine, University of Dundee, Dundee, B15 2TT, UK: Miles D Witham, clinical reader; [18]Beth Israel Deaconess Medical Center, Department of Medicine, Boston, MA 02215, USA: Gloria Y Yeh professor; [19]National Knowledge Centre for Rehabilitation and Palliative Care, University Hospital Odense and University of Southern Denmark, 5800 Nyborg, Denmark: Ann-Dorthe O Zwisler, professor.

Correspondence to: Professor Rod Taylor, Institute of Health Research, University of Exeter Medical School, College House, St Lukes Campus, Heavitree Road, Exeter, EX1 2LU, England, United Kingdom. Email: r.taylor@exeter.ac.uk; Tel: +44 (0)7968 152537

Keywords: cardiac rehabilitation, exercise training, meta-analysis, systematic review

Total word count: 3584

Abstract

Aims To undertake an individual patient data (IPD) meta-analysis to assess the impact of exercise-based cardiac rehabilitation (ExCR) in patients with heart failure (HF) on mortality and hospitalisation, and differential effects of ExCR according to patient characteristics: age, sex, ethnicity, New York Heart Association (NYHA) functional class, ischaemic aetiology, ejection fraction, and exercise capacity.

Methods Randomised trials of exercise training for at least 3 weeks compared with no exercise control with 6-months' follow up or longer, providing IPD time to event on mortality or hospitalisation (all-cause or HF-specific). IPD were combined into a single dataset. We used Cox proportional hazards models to investigate the effect of ExCR and the interactions between ExCR and participant characteristics. We used both two-stage random effects, and 1-stage fixed effect, models.

Results IPD was obtained from 18 trials including 3912 patients with reduced ejection fraction HF. Compared to control, there was no statistically significant difference in pooled time to event estimates in favour of ExCR although confidence intervals were wide: all-cause mortality: hazard ratio (HR): 0.83 (95% confidence interval (CI): 0.67 to 1.04), HF-related mortality: HR 0.84 (95% CI: 0.49 to 1.46), all-cause hospitalisation: HR 0.90 (95% CI: 0.76 to 1.06), and HF-related hospitalisation: HR 0.98 (95% CI: 0.72 to 1.35). No strong evidence was found of differential intervention effects across patient characteristics.

Conclusion ExCR did not have a statistically significant effect on the risk of mortality and hospitalisation in reduced ejection fraction HF. However, uncertainty around effect estimates precludes drawing definitive conclusions.

Introduction

With increasing numbers of people living longer with symptomatic heart failure (HF), the effectiveness and accessibility of health services for HF patients have never been more important. Exercise-based cardiac rehabilitation (ExCR) is recognised as integral to the comprehensive care of HF patients.[1,2] Cardiac rehabilitation is a process by which patients, in partnership with health professionals, are encouraged and supported to achieve and maintain optimal physical health.[2] Exercise training is at the centre of rehabilitation provision for HF. In addition, it is now accepted that programmes should be comprehensive in nature and also include education and psychological care, as well as focus on health and life-style behaviour change and psychosocial wellbeing.[2,3]

Systematic reviews have shown ExCR offers important health benefits for HF patients.[4-7] The 2014 Cochrane review, based on aggregate trial data up to 12-months follow up, reported a reduction in the risk of overall hospitalisation (relative risk (RR) 0.75; 95% CI 0.62 to 0.92), HF-specific hospitalisation (RR 0.61; 95% CI 0.46 to 0.80) compared with no exercise control.[7] However, there is uncertainty as to whether there are differential effects of ExCR across HF patient subgroups. In 2004, the Exercise Training Meta-Analysis of Trials for Chronic Heart Failure (ExTraMATCH) Collaborative Group published an individual patient data (IPD) meta-analysis based on 9 randomised trials in 801 HF patients.[8] ExTraMATCH reported a reduction in all-cause mortality (hazard ratio (HR): 0.65, 95% CI 0.46 to 0.92) and in the composite of mortality and hospital admission (HR 0.72, 95% CI 0.56 to 0.93) with ExCR compared to no exercise control. There were no statistically significant treatment effects across subgroups. Given the small number of trials, patients, and events, those subgroup analyses are likely to be underpowered. As the ExTraMATCH analysis did not take into account the cluster (or trial-level) nature of the data, it is likely to have underestimated the precision of the treatment effect. Since the ExTraMATCH analysis, there have been publications of trials of ExCR in HF, including HF-ACTION, a large US National Institute of Health funded trial with 2331 HF patients recruited from 82 centres.[9]

The ExTraMATCH II collaboration brings together the most comprehensive IPD meta-analysis of randomised trial data for ExCR in HF to date. Using contemporary IPD meta-analysis statistical methods, this study aimed to assess the impact of ExCR on the time to event outcomes (all-cause mortality, HF-specific mortality, all-cause hospital admission, and HF-specific hospital admission), and to identify subgroups of patients with HF that may respond differently to ExCR.

Methods

This study was conducted and reported in accordance with current IPD guidance and Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data (PRISMA IPD) statement.[10,11] The study was registered with the PROSPERO (CRD42014007170) and our full study protocol has been published elsewhere.[12,13]

Search strategy and selection criteria

Trials for inclusion were identified from the original ExTraMATCH IPD meta-analysis and the current Cochrane systematic review of ExCR for HF.[7,8] The Cochrane review searched the following electronic databases: Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library, EMBASE, MEDLINE, CINAHL, PsycINFO, and the NHS Centre for Reviews and Dissemination. Conference Proceedings were searched on Web of Science. Trial registers (Controlled-trials.com and Clinicaltrials.gov) and reference lists of all eligible trials and identified systematic reviews were also searched. No language limitations were imposed. Details of the search strategy used are reported in the study protocol.[12]

Trials which met the following criteria were eligible for inclusion in this analysis: (i) randomised trials of adult patients (aged 18 years and older) with a diagnosis of HF with reduced ejection fraction (HFrEF) or HF with preserved ejection fraction (HFpEF) based on objective assessment of left ventricular ejection fraction and on clinical findings; (ii) trial intervention (ExCR) that included an aerobic exercise training component performed by the lower limbs, lasting a minimum of 3 weeks,[7] either alone or as part of a comprehensive cardiac rehabilitation programme which may also include health education and/or a psychological intervention; (iii) a control arm which did not prescribe an exercise intervention; (iv) a minimum follow-up of 6 months; and (v) a sample size of at least 50 (to ensure that the logistical effort in obtaining, cleaning and organising the data was commensurate with the contribution of the dataset to the analysis).[14,15]

Data management

The principal investigators of included trials were invited by email to participate in this IPD meta-analysis and share their anonymised trial data. Included datasets had ethical approval and consent from their sponsors. The complete list of all requested variables and details regarding collaboration with principal investigators are reported in the study protocol.[12] Each dataset was saved in its original format and then converted and combined into one overall master dataset with standardised variables. All files are stored on a secure password protected computer server managed in accordance with the data management standard operating procedures of the nationally registered Exeter Clinical Trials Unit. Data from each trial were checked on range, extreme values, internal consistency, missing values, and consistency with published reports. Data discrepancies or missing information were discussed with trial investigators. Access to data at all stages of cleaning and analysis was restricted to core members of the research team (OC, RST, FCW, and SW).

Specification of outcomes, subgroups, and risk of bias assessment

We sought patient level time to event data from investigators for the following outcomes: time to all-cause and HF-specific mortality, and time to first all-cause and HF-specific hospitalisation. We also sought IPD on the following pre-defined patient characteristics: age, gender, ejection fraction, New York Heart Association (NYHA) functional class, HF aetiology (ischaemic or. non-ischaemic), ethnicity (white or other), and baseline (pre-randomisation) exercise capacity (e.g. peak oxygen uptake (VO_2)). Study quality/risk of bias was assessed using the TESTEX quality assessment tool.[16]

Statistical analysis

A detailed statistical analysis plan was prepared (available from authors). All analyses were carried out according to the principle of intention to treat (i.e. patients included according to their randomised trial arm) and included only patients with observed baseline data (where required) and outcome data at follow-up. Where missing data was noted within an individual trial, contact with

the author was attempted and data added if available. Given the relatively small levels of missing outcome and covariate data within trials, we did not undertake data imputation. We checked for potential small study bias by assessing funnel plot asymmetry and using the Egger test.[22]

In the primary analysis, a two-stage approach was taken, with each trial first analysed using a Cox proportional hazards regression model and then trial-specific estimates of treatment effect (hazard ratio (HR)) or treatment–covariate interactions (HR of the interaction effect) were combined across trials using a random effects model. A random effects model was preferred due to the high degree of clinical heterogeneity across the individual trials, which included different patient populations, types of ExCR intervention, and comparators.[17] The overall estimate of the effect of ExCR for each outcome, both by trial and as a pooled estimate, was presented as a HR and 95% confidence interval (CI). Additionally, the I^2 and τ^2 statistics were reported alongside the associated p-value for the results of the main analyses.

Secondary analyses were based on a one-stage meta-analysis approach. Due to failure of convergence of one-stage random effects models, which was likely to be due to the low level of statistical heterogeneity between trials (indicated by the τ^2 statistic), a fixed effect approach was used: Cox regression models, stratified by trial. Stratification allowed the baseline hazard to vary between trials, rather than forcing the baseline hazards in individual trials to be proportionate to each other.[19] To investigate subgroup effects, specifically interactions between ExCR effect and patient characteristics, we used the approach recommended by Riley et al.[20] Continuous covariates were centred around the mean value within each trial; binary covariates were centred around the proportion within each trial. To present the results graphically, we performed individual subgroup one-stage fixed-effect IPD meta-analyses.

To test the robustness of primary and secondary analyses, we undertook a number of pre-specified sensitivity analyses: (i) exclusion of the largest trial (HF-ACTION[9]); (ii) truncation of outcomes at 1-, 2- and 5-year follow-up. Small study effects were assessed for each outcome by funnel plot asymmetry

and using the Egger test.[22] Results are reported as estimated HRs with 95% CIs. All analyses were undertaken using Stata 14.2 StataCorp LP, College Station, Texas, USA.

Results

Study selection

A total of 23 trials were deemed eligible for this IPD meta-analysis. Data from six trials have been analysed previously and were available from the ExTraMATCH database.[23-28] We were unable to include data from three trials (355 patients); for two trials data were no longer available [29,30] and the investigators of the other trial could not be contacted.[31] Of the remaining 17 trials, 14 investigators responded positively and shared their trial data. After obtaining IPD, one trial[32] was excluded as it was determined that it included patient data that overlapped with another trial.[33] A further trial was not included as insufficient data was provided to allow calculation of survival time or time to hospitalisation.[34] This resulted in the inclusion of 18 trials comprising 3912 patients (1948 ExCR, 1964 control) with a median follow-up of 19 months for mortality outcomes and 11 months for hospitalisation outcomes. Figure 1 summarises the study selection process. Full citations of included trials are included in eAppendix.[e1-e18]

Characteristics of included trials and participants

Patient baseline characteristics were well balanced between ExCR and control patients (see Table 1). The majority of patients were male (75%), with a mean age of 61 years (standard deviation (SD) 13). The mean baseline left-ventricular ejection fraction was 26.7% (SD 8.1%); no included trials recruited patients with preserved ejection fraction heart failure (ejection fraction >45%), and most patients were in NYHA functional class II (59%) or III (37%). Trials from Europe and North America were published between 1990 and 2012 (see Table 2). Sample size ranged from 50 to 2130 patients. All trials evaluated an aerobic exercise intervention; six also included resistance training.[e3,e4,e8,e9,e14,e15] Exercise training was most commonly delivered in either an exclusively centre-based setting or a centre-based setting in combination with some home exercise sessions. Three trials were conducted in an exclusively home-based setting.[e4,e9,e13] The dose of exercise training ranged widely across trials. ExCR was delivered over a period of 12 to 90 weeks, with between 2 and 7 sessions per week;

median session duration was between 15 and 120 minutes (including warm-up and cool-down). The intensity of exercise ranged between 50 to 85% peak VO_2.

Quality of included trials

The overall quality of included trials was judged to be moderate to good, with a median TESTEX score of 11 (range 9 to 14) out of a maximum score of 15 (eTable1). The criteria of allocation concealment and physical activity monitoring in the control groups were met in only three trials;[e4,e7,e17] the other TESTEX criteria were met in 50% or more of trials.

Effects of intervention on event outcomes

Compared with control, ExCR did not have a statistically significant effect on mortality or hospitalisation. However, all time to event pooled treatment effects from random effects two-stage IPD meta-analysis were in favour of ExCR but with wide confidence intervals; all-cause mortality: HR 0.83 (95% CI 0.67 to 1.04), p=0.107, 17 trials, 3782 patients, I^2 =26%, τ^2=0.04; HF-specific mortality: HR 0.84 (95% CI 0.49 to 1.46), p=0.527, 9 trials, 915 patients, I^2 =0%, τ^2=0.00; all-cause hospitalisation: HR 0.90 (95% CI 0.76 to 1.06), p=0.210, 11 trials, 3190 patients, I^2 =12.4%, τ^2=0.01; and HF-specific hospitalisation: HR 0.98 (95% CI 0.72 to 1.35), p=0.902, 13 trials, 3494 patients, I^2 =45%, τ^2=0.10 (Figure 2a-d). These primary analysis results were broadly consistent across secondary and sensitivity analyses (eTables 2, 3, 4, and 5). Inferences did not change following the addition of trial level data from trials that met our study inclusion criteria but were not able to contribute IPD (data not shown here, available from authors). There was no evidence of significant small study bias for the four outcomes (eFigure 1).

Differential treatment effects **across** ***patient characteristics*** **(*subgroups)***Interaction analyses for the two-stage model revealed no consistent interactions between the effect of ExCR and any of the predefined subgroups (age, gender, ejection fraction, NYHA class, HF aetiology, ethnicity and

baseline peak VO_2) for all-cause mortality, HF-related mortality, all-cause hospitalisation, or HF-related hospitalisation. For comparison of mortality and hospitalisation rates within each subgroup, the HR and associated 95% CI from individual subgroup one-stage IPD meta-analyses are shown in Figures 3a-d, along with the p-value from the interaction test in the two-stage IPD meta-analyses. Some evidence of an interaction effect between ExCR and a patient characteristic ($p<0.05$) was seen in four sensitivity analyses (eTables 2, 3, 4, and 5): (i) ExCR was associated with a larger reduction in all-cause mortality in older patients ($p=0.034$) in the two-stage model with 2-year truncation; (ii) ExCR was associated with a larger reduction in HF-mortality older patients ($p=0.017$) in the two-stage model with 2-year truncation; (iii) ExCR was associated with a larger reduction in HF-mortality in ischemic patients ($p=0.047$) in the one-stage model without truncation; and (iv) ExCR was associated with a larger reduction in all-cause hospitalisation in patients with lower baseline peak VO_2 ($p=0.027$) in the two-stage model with 1-year truncation.

Discussion

Compared with no exercise control, ExCR did not have a statistically significant effect on mortality or hospitalisation in patients with reduced ejection fraction HF. Although we pooled IPD from 18 trials including 3912 patients, treatment effect estimates were imprecise. We found no strong evidence for a differential effect of ExCR according to patient characteristics.

Unlike the previous IPD meta-analysis, ExTraMATCH[8], our analyses did not show a definitive benefit of ExCR in terms of either time to all-cause death or all-cause hospitalisation. The confidence intervals around the effect of ExCR were wide and failed to reach statistical significance. The wide confidence intervals may be due to several factors, including: (i) variation in the ExCR intervention across trials; (ii) variation in adherence to ExCR within trials; (iii) variation in treatment effect of ExCR among adherent patients; and (iv) variation in the composition/effectiveness of usual care within and across trials. A potential explanation for this reduction in strength of effect of ExCR on clinical events could be due to improvements in rates of mortality and hospitalisation with time as a result of the inclusion of more recent trials in this updated IPD analysis. More recent trials are more likely to have utilised prognostic innovations in usual care treatments for HF, including devices (resynchronisation and defibrillator therapy) and disease modifying drugs (beta-blockers, ACE inhibitors, angiotensin II receptor antagonists, and mineralocorticoid receptor antagonists). However, the Cochrane systematic review of ExCR for HF showed that this may not be case. Meta-regression showed no statistical association between trial publication date and the magnitude of ExCR effect on mortality or hospitalisation.[11,35] This Cochrane analysis also showed no association between the magnitude of ExCR effect and trial setting (single or multicentre), type of CR (comprehensive versus exercise only), or ExCR dose.

Our finding of a lack of consistent evidence of a beneficial effect of ExCR for any HF patient subgroup agrees with both the previous ExTraMATCH and Cochrane 2014 analyses..[7,8] However, these two previous studies had major limitations that are likely to have limited their ability to detect subgroup

effects. ExTraMATCH included data on only 801 HF patients and observed 88 deaths and 300 patients with a composite outcome of death or hospitalisation, and therefore lacked statistical power. Using meta-regression analysis, the 2014 Cochrane review found no association between trial level patient characteristics (age, gender, ejection fraction) and ExCR. However, meta-regression analysis is highly prone to study level confounding (ecological fallacy) and should be interpreted with great caution.[36] Nevertheless, our findings are also consistent with the IPD subgroup interaction analyses from the multicentre HF-ACTION trial. The HF-ACTION investigators reported no significant interaction effect on their composite primary outcome (all-cause mortality or hospitalisation) between exercise training intervention and the subgroups of age (≤70 vs. > 70 years), gender, race (white vs. non-white), HF aetiology (ischaemic vs. non ischaemic), ejection fraction (≤25% vs. >25%), and NHYA class (II vs. III/IV).[9,37]

Strengths and limitations

Our ExTraMATCH II study has a number of strengths. We believe this to be the first IPD meta-analysis including sufficiently large numbers of HF patients (3912) and events (701 all-cause deaths, 1642 first all-cause hospitalisations) to be able to identify differential effects of ExCR in patients with different characteristics. We were able to standardise the handling and analysis of time to event outcomes across trials. Our findings were broadly consistent across analytic approaches that included one-- and two-stage IPD meta-analysis models and a range of sensitivity analyses. Finally, we found no evidence of publication bias. Whilst systematic reviews and meta-analyses of IPD from randomised trials are recognised as the gold standard for assessing intervention effects,[39] our study has a number of limitations. First and foremost, there was a lack consistency in how included trials with IPD in our analyses defined and collected clinical event outcome data. As noted in recent commentaries on clinical events in HF trials, with the exception of all-cause mortality, the collection and reporting the other outcomes including cause-specific mortality and hospitalisation can be prone to confounding and bias.[39] We made considerable efforts to contact study authors in order to

clarify issues around the definition of hospitalisations and HF-specific deaths. Although we were able to resolve data issues in many cases, we recognise that a lack of consistency in outcome definition across included trials may exist, weakening the strength of our conclusions for these outcomes. Second, overall, IPD was available to ExTraMATCH II for 3912/4267 patients in 18/23 trials identified as eligible, equating to an omission of only 8% of all participants across all eligible trials. However, not all included trials collected IPD for the time to event outcomes or patient characteristics assessed in this study. The large multicentre HF-ACTION trial did not collect HF-specific hospitalisation, thus reducing our statistical power for this outcome. Although, across the trials that provided outcome data, the proportion of patients with missing clinical event or baseline covariate data was low, this may have introduced bias in our results. Finally, we did not have patient level data on 'ExCR dose', i.e. adherence to duration, frequency and intensity of ExCR undertaken by an individual patient. Using IPD from HF-ACTION, Keteyian et al found exercise volume (defined as metabolic equivalent of task (MET)-hour per week) to be a predictor for the composite outcome of all-cause mortality or hospitalisation (p=0.03).[40] Whilst this analysis indicates that patient level ExCR 'dose' is a key potential explanatory variable, this data was not available across the trials in our analyses.

Implications for practice and further research

In spite of the comprehensiveness of this IPD meta-analysis, findings of this study demonstrate that further evidence is still required to definitively assess the impact of ExCR on mortality and hospitalisation in patients with reduced ejection fraction HF; in particular, to increase the power to examine whether the effect of ExCR varies according to patient characteristics. To more reliably quantify the impact of ExCR on clinical outcomes and examine how these effects may vary across HF patients, there is an urgent need for trial investigators to more consistently collect, report, and share patient-level data in the future. Two central aspects of future data collection include a consensus on the definition, collection, and reporting of clinical event data, especially hospitalisation, plus the

capture of data on patient-level adherence to the amount of exercise training during the ExCR intervention period. Our forthcoming IPD meta-analysis will examine the impact of ExCR exercise capacity on health-related quality of life and explore how this treatment effect may vary across according to HF patient characteristics.[12,13]

Conclusion

ExCR did not significantly reduce the risk of mortality and hospitalisation in patients with reduced ejection fraction HF. Although we pooled the IPD from a number of randomised trials, treatment effect estimates remain imprecise which precludes drawing definitive conclusions. To allow definitive assessment of the effect of ExCR for patients with HF, and to investigate differential treatment effects across specific patient characteristics, a consensus in trial methodology needs to be reached that will allow more detailed and consistently recorded IPD to be routinely collected from clinical trials in ExCR.

Funding: This work is supported by UK National Institute for Health Research funding (HTA 15/80/30). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Acknowledgments: RST, NS, MP, FCW, OC designed the study and obtained the study funding. SW analysed the data. All authors contributed to writing and editing of the manuscript, with the lead taken by RST and SW. CO'C, KD, RH, KJ, RMc, JM, BBN, CP, MDW, GY ADZ contributed individual participant data for this study. All authors commented on the manuscript and agreed the final version.

We acknowledge the contribution of patient data from Dr Rainer Hambrecht, Dr Emanuella Zanelli, and the late Dr Romualdo Belardinelli and late Dr Pantaleo Giannuzzi. We thank Mr Tim Eames, Exeter Clinical Trials Unit, for his advice and support on data management for this study, and Professor Richard Riley, Professor of Biostatistics, Keele University, for his comments on a draft of this paper.

Conflict of Interest: RST and HMD are currently co-chief investigators and KJ a co-investigator on a National Institute for Health Research (NIHR) funded programme grant designing and evaluating the clinical and cost-effectiveness of a home-based cardiac rehabilitation intervention for heart failure patients (RP-PG-1210-12004). All other authors declare no conflicts.

References

1. Bjarnason-Wehrens B, McGee H, Zwisler A-D, et al. Cardiac rehabilitation in Europe: results from the European Cardiac Rehabilitation Inventory Survey. *Eur J Cardiovasc Prev Rehabil.* 2010;17:410-8.
2. British Association for Cardiac Prevention and Rehabilitation (BACPR). *Standards and core components for cardiovascular disease prevention and rehabilitation 2017*. 3rd edition. www.bacpr.com/resources/6A7_BACR_Standards_and_Core_Components_2017.pdf
3. Dalal HM, Doherty P, Taylor RS. Cardiac rehabilitation. *BMJ.* 2015;351:h5000.
4. Lewinter C, Doherty P, Gale, CP, Crouch S, Stirk L, Lewin RJ, LeWinter MM, Ades P A, Kober L, Bland JM. Exercise-based cardiac rehabilitation in patients with heart failure: a meta-analysis of randomised controlled trials between 1999 and 2013. *Eur J Prev Cardiol.* 2015;22:1504-1512.
5. Haykowsky MJ, Timmons MP, Kruger C, McNeely M, Taylor DA, Clark AM. Meta-analysis of aerobic interval training on exercise capacity and systolic function in patients with heart failure and reduced ejection fractions. *Am J Cardiol.* 2013;111;1466-1469.
6. Smart N, Marwick TH. Exercise training for patients with heart failure: a systematic review of factors that improve mortality and morbidity. *Am J Med.* 2004; 116:693-706.
7. Taylor RS, Sagar VA, Davies EJ, Briscoe S, Coats AJ, Dalal H, Lough F, Rees K, Singh S. Exercise-based rehabilitation for heart failure. *Cochrane Database Syst Rev.* 2014;(4):CD003331.
8. Piepoli MF, Davos C, Francis DP, Coats AJ. Exercise training meta-analysis of trials in patients with chronic heart failure (ExTraMATCH). *Brit Med J.* 2004;328:189-93.
9. O'Connor CM, Whellan DJ, Lee KL, et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA*. 2009;301:1439.

10. Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, Rovers M. Individual Participant Data (IPD) Meta-analyses of randomised controlled trials: Guidance on their use. *PLoS Med.* 2015;12:e1001855.
11. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF; PRISMA-IPD Development Group. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA.* 2015;313:1657-65.
12. Taylor RS, Piepoli MF, Smart N, Coats AJ, Ellis S, Dalal H, O'Connor CM, Warren FC, Whellan D, Ciani O; ExTraMATCH II Collaborators. Exercise training for chronic heart failure (ExTraMATCH II): protocol for an individual participant data meta-analysis. *Int J Cardiol.* 2014;174:683-7.
13. Taylor RS, Piepoli MF, Smart N, Coats AJ, Ellis S, Dalal H, O'Connor CM, Warren FC, Whellan D, Ciani O; ExTraMATCH II Collaborators. Exercise training for chronic heart failure (ExTraMATCH II): protocol for an individual participant data meta-analysis PROSPERO register. 2014. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014007170 [last accessed 18th January 2018]
14. Riley R, Lambert P, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *Brit Med J.* 2010;340:c221
15. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *Brit Med J.* 2012;344:d7762.
16. Smart NA WM, Ismail H, Giallauria F, Vigorito C, Cornelissen V, Dieberg G. Validation of a new tool for the assessment of study quality and reporting in exercise training trials: TESTEX. *Int J Evid Based Healthc.* 2015;13:9-18.

17. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21:1539-58.
18. Higgins J PT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557-560.
19. Cleves M, Gould W, Marchenko Y. *An Introduction to Survival Analysis Using Stata*. Third ed: Stata Press; 2016.
20. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stats Med.* 2008;27:1870-93.
21. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ.* 2012;344:d7762
22. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315: 629-34.
23. Belardinelli R, Georgiou D, Cianci G, Purcaro A. Randomized, controlled trial of long-term moderate exercise training in chronic heart failure: effects on functional capacity, quality of life, and clinical outcome. *Circulation* 1999;99:1173-82.
24. Hambrecht R, Gielen S, Linke A, Fiehn E, Yu J, Walther C et al. Effects of exercise training on left ventricular function and peripheral resistance in patients with chronic heart failure: A randomized trial. *JAMA.* 2000;283:3095-101.
25. McKelvie RS, Teo KK, Roberts R, McCartney N, Humen D, Montague T et al. Effects of exercise training in patients with heart failure: the Exercise Rehabilitation Trial (EXERT). *Am Heart J.* 2002;144:23-30.
26. Willenheimer R, Rydberg E, Cline C, Broms K, Hillberger B, Oberg L, Erhardt L. Effects on quality of life, symptoms and daily activity 6 months after termination of an exercise training programme in heart failure patients. *Int J Cardiol.* 2001;77:25-31.

27. Zanelli E, Volterrani M, Scalvini S, Musmeci G, Campana M, Zappa C, et al. Multidisciplinary non-pharmacological intervention prevents hospitalization, improves morbidity rates and functional status in patients with congestive heart failure. *Eur Heart J.*1997;18 (abstract suppl):647.
28. Wielenga RP, Huisveld IA, Bol E, Dunselman PH, Erdman RA, Baselier MR, et al. Safety and effects of physical training in chronic heart failure. Results of the chronic heart failure and graded exercise study (CHANGE) *Eur Heart J.* 1999;20:872-9.
29. Davidson PM, Cockburn J, Newton PJ, et al. Can a heart failure-specific cardiac rehabilitation program decrease hospitalizations and improve outcomes in high-risk patients? *Eur J Cardiovasc Prev Rehabil.* 2010;17:393–402.
30. Austin J, Williams R, Ross L, et al. Randomised controlled trial of cardiac rehabilitation in elderly patients with heart failure. *Eur J Heart Fail.* 2005;7:411–17.
31. Klecha A, Kawecka-Jaszcz K, Bacior B, et al. Physical training in patients with chronic heart failure of ischaemic origin: effect on exercise capacity and left ventricular remodeling. *Eur J Cardiovasc Prev Rehabil.* 2007;14:85-91.
32. Dubach P, Myers J, Dziekan G, Goebbels U, Reinhart W, Muller P, et al. Effect of high intensity exercise training on central hemodynamic responses to exercise in men with reduced left ventricular function. *J Am Coll Cardiol.* 1997;29:1591-8.
33. Mueller L, Myers J, Kottman W, Oswald U, Boesch C, Arbrol N, Dubach P. Exercise capacity, physical activity patterns and outcomes six years after cardiac rehabilitation in patients with heart failure. *Clin Rehab.* 2007;21:923-31.
34. Gary RA, Dunbar SB, Higgins MK, Musselman DL, Smith AL. Combined exercise and cognitive behavioral therapy improves outcomes in patients with heart failure. *J Psychosom Res.* 2010;69(2):119-31.
35.

36. Sagar VA, Davies EJ, Briscoe S, Coats AJ, Dalal HM, Lough F, Rees K, Singh S, Taylor RS. Exercise-based rehabilitation for heart failure: systematic review and meta-analysis. Open Heart. 2015;2(1):e000163Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med.* 2002;21:1559-73.

37. Mentz RJ, Bittner V, Schulte PJ, Fleg JL, Piña IL, Keteyian SJ, Moe G, Nigam A, Swank AM, Onwuanyi AE, Fitz-Gerald M, Kao A, Ellis SJ, Kraus WE, Whellan DJ, O'Connor CM. Race, exercise training, and outcomes in chronic heart failure: findings from Heart Failure - a Controlled Trial Investigating Outcomes in Exercise TraiNing (HF-ACTION). *Am Heart J.* 2013;166:488-95

38. Chalmers I; Cochrane Collaboration. The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci.* 1993;703:156-163.

39. Zannad F, Garcia AA, Anker SD, Armstrong PW, Calvo G, Cleland JG, Cohn JN, Dickstein K, Domanski MJ, Ekman I, Filippatos GS, Gheorghiade M, Hernandez AF, Jaarsma T, Koglin J, Konstam M, Kupfer S, Maggioni AP, Mebazaa A, Metra M, Nowack C, Pieske B, Piña IL, Pocock SJ, Ponikowski P, Rosano G, Ruilope LM, Ruschitzka F, Severin T, Solomon S, Stein K, Stockbridge NL, Stough WG, Swedberg K, Tavazzi L, Voors AA, Wasserman SM, Woehrle H, Zalewski A, McMurray JJ. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. Eur J *Heart Fail.* 2013;15:1082-94.

40. Keteyian SJ, Leifer ES, Houston-Miller N, Kraus WE, Brawner CA, O'Connor CM, Whellan DJ, Cooper LS, Fleg JL, Kitzman DW, Cohen-Solal A, Blumenthal JA, Rendall DS, Piña IL; HF-ACTION Investigators. Relation between volume of exercise and clinical outcomes in patients with heart failure. *J Am Coll Cardiol.* 2012;60:1899-905.

Figure legends

Figure 1. PRISMA flow diagram

Figure 2. Effect of ExCR on mortality and hospitalisation across patient subgroups

Figure 3. Effect of ExCR on mortality and hospitalisation across patient subgroups

Figure 1. PRISMA flow diagram

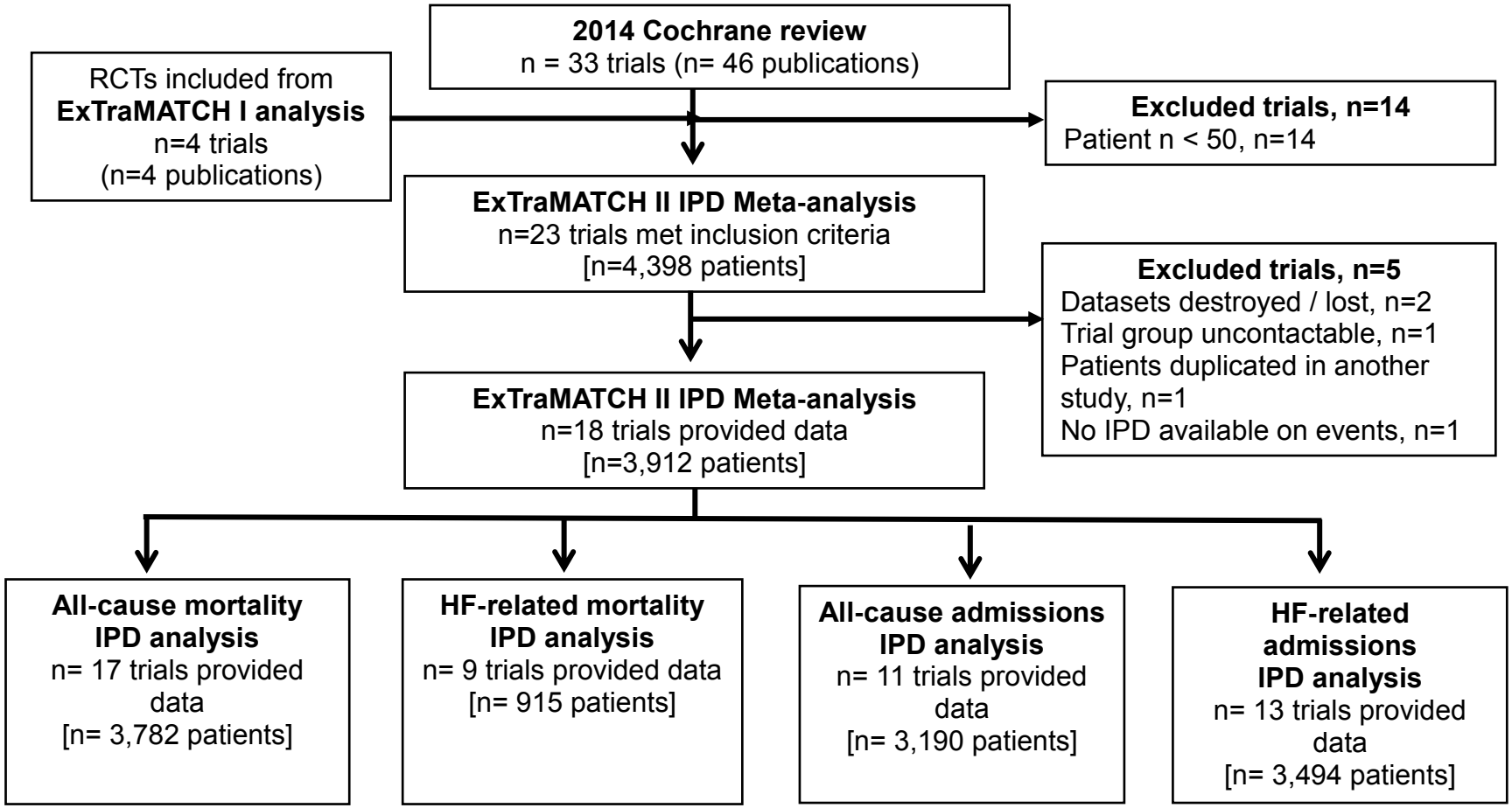


Table 1. Baseline characteristics of patients

Characteristic	ExCR (n=1,948)	Control (n=1,964)	All (n=3,912)
Age (years), mean (SD)	61.3 (12.7)	61.4 (13.2)	61.3 (13.0)
Gender			
Male	1442 (74)	1489 (76)	2,931 (75)
Female	506 (26)	475 (24)	981 (25)
Baseline ejection fraction (%); mean (SD)	26.8 (8.2)	26.7 (8.1)	26.7 (8.1)
NYHA status			
Class I	25 (1)	28 (1)	53 (1)
Class II	1107 (59)	1130 (60)	2237 (59)
Class III	700 (37)	708 (37)	1408 (37)
Class IV	47 (3)	26 (1)	73 (2)
Aetiology			
Ischaemic	1094 (57)	1080 (56)	2174 (57)
Non-ischemic	809 (43)	838 (44)	1647 (43)
Ethnicity			
White	1100 (70)	1140 (72)	2240 (71)
Non-white	472 (30)	445 (28)	917 (29)
VO_2 peak (ml/kg/min); mean (SD)	14.9 (4.4)	15.0 (4.6)	14.9 (4.5)

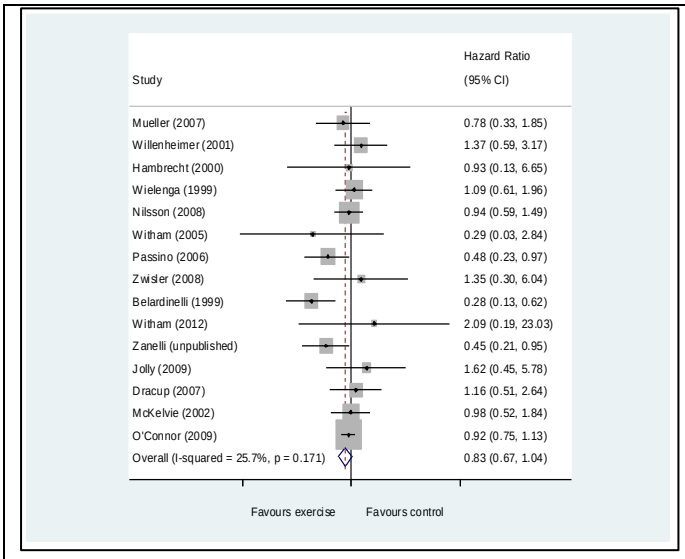
Data are n (%) unless otherwise indicated; percentages may not sum to 100 because of rounding; NHYA: New York Heart Association; SD: standard deviation; VO_2 peak: peak oxygen uptake.

Table 2. Characteristics of included studies

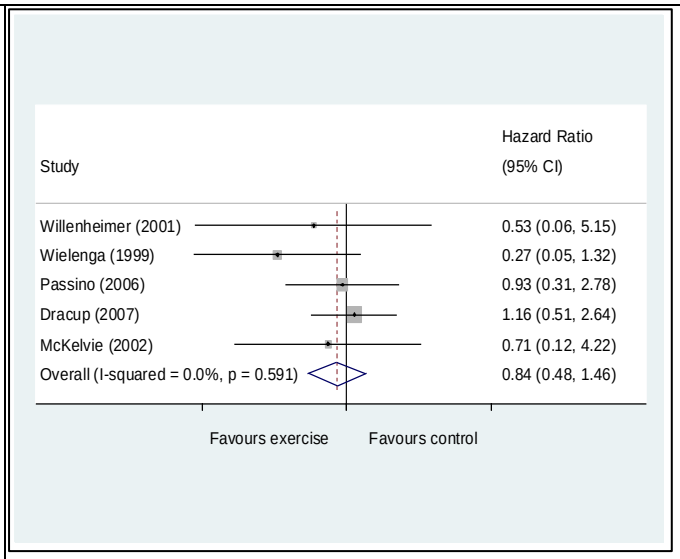
Study characteristics	
Publication year	
1990 to 1999	2 (11)
2000 to 2009	12 (67)
2010 to 2012	3 (17)
Unpublished	1 (6)
Main study location	
Europe	14 (78)
North America*	4 (22)
Single study centre	
Single	12 (67)
Multiple	5 (28)
Not reported	1 (6)
Sample size	
0 to 99	10 (56)
100 to 999	7 (39)
1000 and over	1 (6)
Duration of follow-up in dataset (months), median (range)	
Mortality	18.6 (11.8 to 419)
Hospitalisation	11.2 (2.6 to 98)
Intervention characteristics	
Intervention type	
Exercise only programs	5 (28)
Comprehensive programs	12 (67)
Not reported	1 (6)
Type of exercise	
Aerobic exercise only	12 (67)
Aerobic plus resistance training	6 (33)
Dose of intervention	
Duration of intervention (weeks), median (range)	30 (15 to 90)
Frequency (sessions per week), median (range)	2.8 (2 to 7)
Length of exercise session (mins), median (range)	24 (15 to 120)
Exercise intensity, range	40-80% maximum heart rate 50-85% peak VO_2 12-18 Borg rating
Setting	
Centre-based only	6 (33)
Home-based only	3 (17)
Centre- and home-based	8 (44)
Not reported	1 (6)

Data are n (%) unless otherwise indicated; percentages may not sum to 100 because of rounding; *****O'Connor study was categorised as North America but was also delivered in to a small number of patients in France

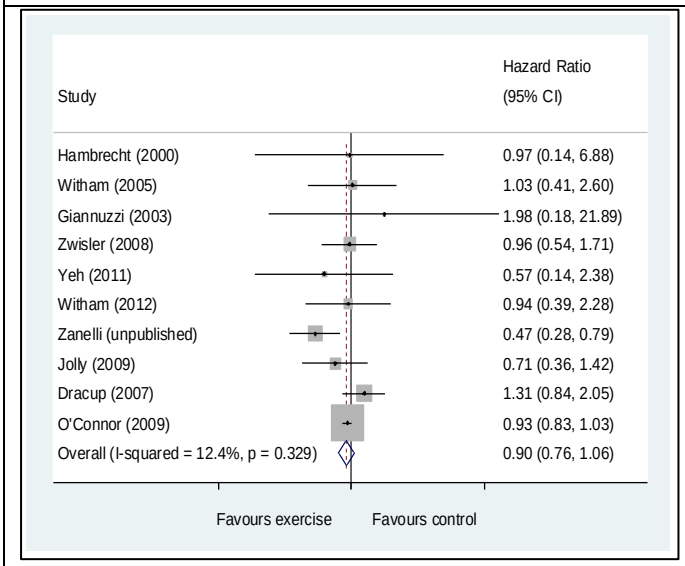
Figure 2. Effect of ExCR on mortality and hospitalisation: two-stage IPD meta-analysis



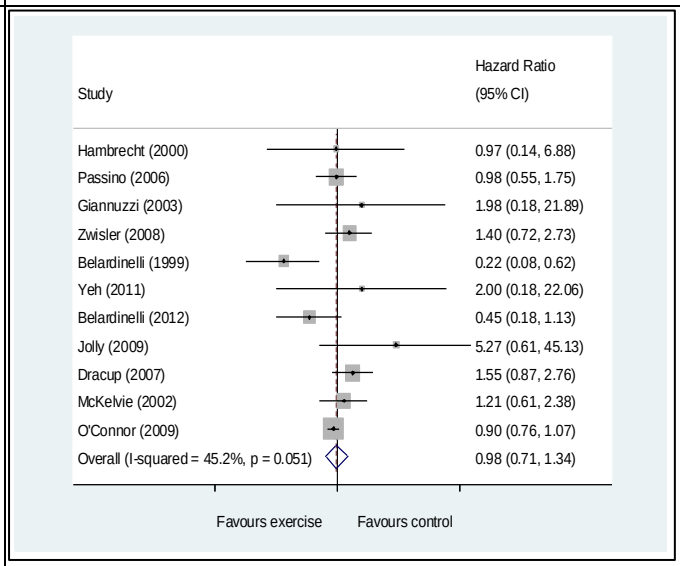
a. All-cause mortality



b. HF-specific mortality

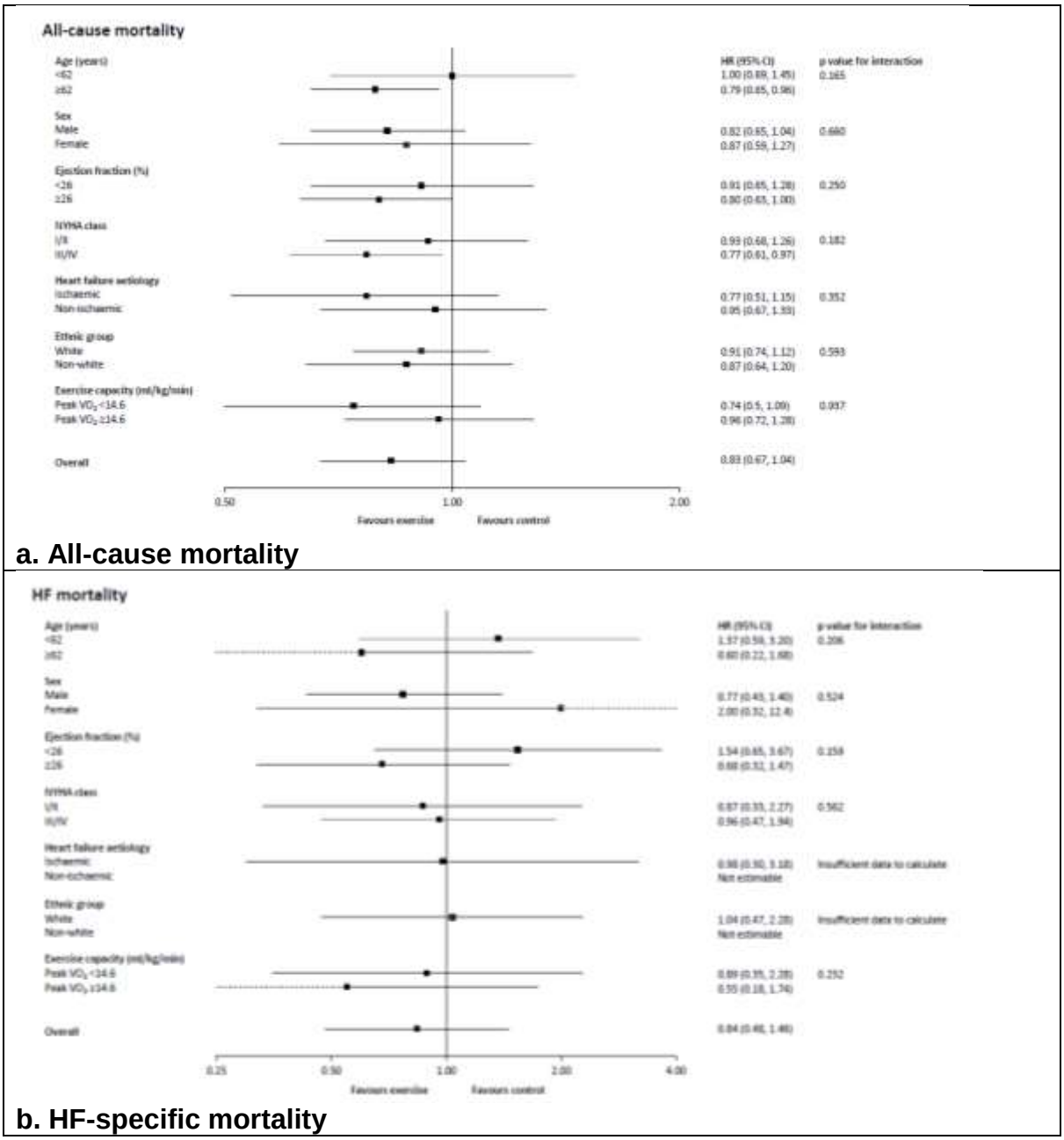


c. All-cause hospitalisation



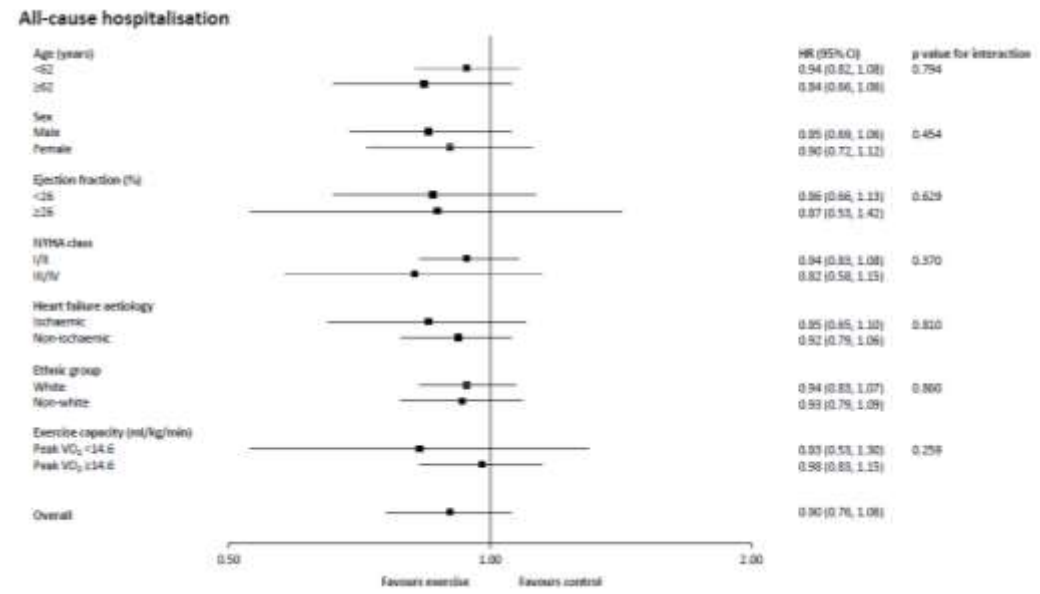
d. HF-specific hospitalisation

Figure 3. Effect of ExCR on mortality and hospitalisation across patient subgroups

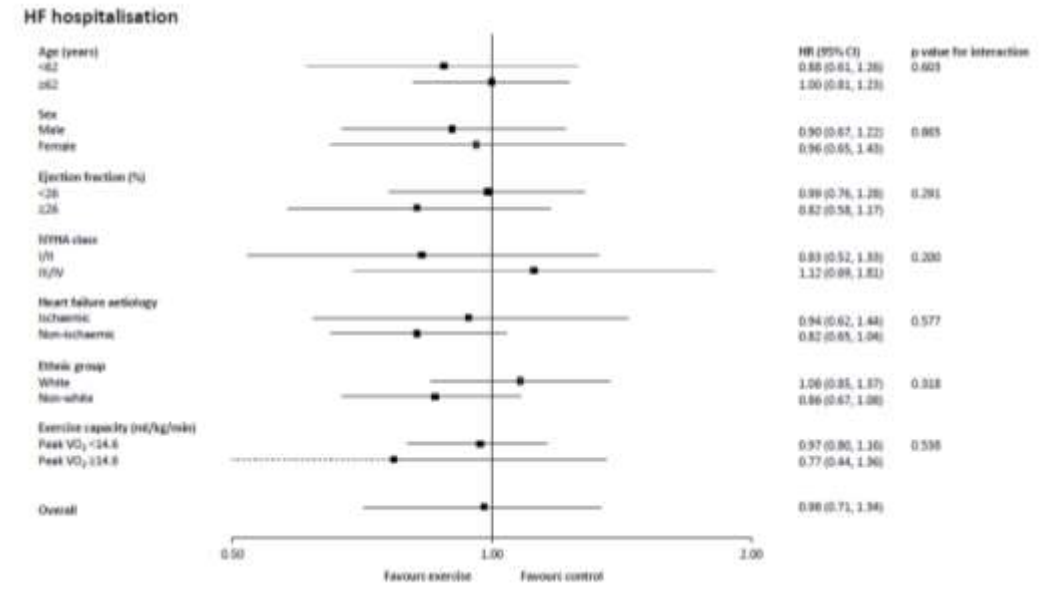


a. All-cause mortality

b. HF-specific mortality



c. All-cause hospitalisation



d. HF-specific hospitalisation

e-Appendix

References of included studies

e1. Belardinelli 1999

- Belardinelli R, Georgiou D, Cianci G, Purcaro A. Randomized, controlled trial of long-term moderate exercise training in chronic heart failure: effects on functional capacity, quality of life, and clinical outcome. *Circulation* 1999;99:1173-82.

e2. Belardinelli 2012

- Belardinelli R Georgiou D, Cianci G, Purcaro A. 10-year exercise training in chronic heart failure: a randomized controlled trial. Journal of the *Am J Cardiol.*2012;60:1521-8.

e3. DANREHAB 2008

- Zwisler AD, Schou L, Soja AM, Brønnum-Hansen H, Gluud C, Iversen L, Sigurd B, Madsen M, Fischer-Hansen J; DANREHAB group. A randomized clinical trial of hospital-based, comprehensive cardiac rehabilitation versus usual care for patients with congestive heart failure, ischemic heart disease, or high risk of ischemic heart disease (the DANREHAB trial) - design, intervention, and population. *Am Heart J.* 2005;899:e16.
- Zwisler AD, Soja AM, Rasmussen S, Frederiksen M, Abedini S, Appel J, Rasmussen H, Gluud C, Iversen L, Sigurd B, Madsen M, Fischer-Hansen J; DANREHAB Group.. Hospital-based comprehensive cardiac rehabilitation versus usual care among patients with congestive heart failure, ischemic heart disease, or high risk of ischemic heart disease: 12-month results of a randomized clinical trial. *Am Heart J*. 2008;155:1106-13.

e4. Dracup 2007

- Dracup K, Evangelista LS, Hamilton MA, Erickson V, Hage A, Moriguchi J et al. Effects of a home-based exercise program on clinical outcomes in heart failure. *Am Heart J.* 2007;154(5):877-83.
- Evangelista LS, Doering LV, Lennie T, Moser DK, Hamilton MA, Fonarow GC, et al. Usefulness of a home-based exercise program for overweight and obese patients with advanced heart failure. *Am J Cardiol.*2006;97:886-90.
- Evangelista LS, Hamilton MA, Fonarow GC, Dracup K. Is exercise adherence associated with clinical outcomes in patients with advanced heart failure? *Phys Sportsmed.* 2010;38:28-36.

e5. Giannuzzi 2003

- Giannuzzi P, Temporelli PL, Corra U, Tavazzi L. Antiremodeling effect of long-term exercise training in patients with stable chronic heart failure: results of the Exercise in Left Ventricular Dysfunction and Chronic Heart Failure (ELVD-CHF) Trial*. Circulation* 2003;108:554-9.

e6. Hambrecht 2000

- Hambrecht R, Gielen S, Linke A, Fiehn E, Yu J, Walther C et al. Effects of exercise training on left ventricular function and peripheral resistance in patients with chronic heart failure: A randomized trial. *JAMA.* 2000;283:3095-101.

e7. HF ACTION 2009

- O'Connor CM, Whellan DJ, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA* 2009;301:1439-50.

- Whellan DJ, O'Connor CM, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ et al. Heart failure and a controlled trial investigating outcomes of exercise training (HF-ACTION): design and rationale. *Am Heart J.* 2007;153:201-11.

e8. Jolly 2009

- Jolly K, Taylor RS, Lip GY, Davies M, Davis R, Mant J, et al. A randomized trial of the addition of home-based exercise to specialist heart failure nurse care: the Birmingham Rehabilitation Uptake Maximisation study for patients with Congestive Heart Failure (BRUM-CHF) study. *Eur J Heart Fail.* 2009;11:205-13.

e9. McKelvie 2002

- McKelvie RS, Teo KK, Roberts R, McCartney N, Humen D, Montague T et al. Effects of exercise training in patients with heart failure: the Exercise Rehabilitation Trial (EXERT). *Am Heart J*. 2002;144:23-30.

e10. Mueller 2007

- Mueller L, Myers J, Kottman W, Oswald U, Boesch C, Arbrol N, Dubach P. Exercise capacity, physical activity patterns and outcomes six years after cardiac rehabilitation in patients with heart failure. *Clin Rehab.* 2007;21:923-31.
- Myers J, Goebbels U, Dzeikan G, Froelicher V, Bremerich J, Mueller P, et al. Exercise training and myocardial remodeling in patients with reduced ventricular function: one-year follow-up with magnetic resonance imaging. *Am Heart J.* 2000;139:252-61.

e11. Nilsson 2008

- Nilsson BB Westheim A, Risberg MA. Long-term effects of a group-based high-intensity aerobic interval-training program in patients with chronic heart failure. *Am J Cardiol.* 2008;102:1220-4.

e12. Passino 2006

- Passino C, Severino S, Poletti R, Piepoli MF, Mammini C, Clerico A et al. Aerobic training decreases B-type natriuretic peptide expression and adrenergic activation in patients with heart failure. *J Am Coll Cardiol.* 2006;47:1835-9.

e13. Willenheimer 2001

- Willenheimer R, Rydberg E, Cline C, Broms K, Hillberger B, Oberg L, Erhardt L. Effects on quality of life, symptoms and daily activity 6 months after termination of an exercise training programme in heart failure patients. *Int J Cardiol.* 2001;77:25-31.

e14. Witham 2005

- Witham MD, Gray JM, Argo IS, Johnston DW, Struthers AD, McMurdo ME. Effect of a seated exercise program to improve physical function and health status in frail patients > or = 70 years of age with heart failure. *Am J Cardiol.*2005;95:1120-4.

e15. Witham 2012

- Witham MD, Fulton RL, Greig CA, Johnston DW, Lang CC, van der Pol M, et al. Efficacy and cost of an exercise program for functionally impaired older patients with heart failure: a randomized controlled trial. *Circulation.* 2012;5:209-16.

e16. Wielenga 1999

- Wielenga RP, Huisveld IA, Bol E, Dunselman PH, Erdman RA, Baselier MR, et al. Safety and effects of physical training in chronic heart failure. Results of

the chronic heart failure and graded exercise study (CHANGE) *Eur Heart J.* 1999;20:872-9.

e17. Yeh 2011

- Yeh GY, McCarthy EP, Wayne PM, Stevenson LW, Wood MJ, Davis RB, et al. Tai chi exercise in patients with chronic heart failure: A randomized clinical trial. *Arch Int Med.* 2011;171:750-7.

e18. Zanelli 1997

- Zanelli E, Volterrani M, Scalvini S, Musmeci G, Campana M, Zappa C, et al. Multidisciplinary non-pharmacological intervention prevents hospitalization, improves morbidity rates and functional status in patients with congestive heart failure. *Eur Heart J.*1997;18(abstract suppl):647

eTable 1. Assessment of quality using TESTEX scale of included studies in IPD meta-analysis

Study (publication year)	Eligibility Criteria Specified	Randomisation Specified	Allocation Concealed	Groups Similar at baseline	Blinding of Assessors	Outcome measures in >85% participants*	Intention to treat analysis**	Between-group statistical comparisons reported†	Point measures & measures of variability reported	Activity monitoring in control group	Relative Exercise intensity reviewed	Exercise Volume and Energy Expended	Overall TESTEX score (maximum score 15)
Belardinelli (1999)	1	0	0	1	0	3	1	1	1	0	0	1	9
Belardinelli (2012)	1	0	0	1	0	3	1	1	1	0	0	1	9
Dracup (2007)	1	0	0	1	0	3	1	2	1	1	1	1	10
Giannuzzi (2003)	1	0	0	1	0	2	1	2	1	0	1	1	10
Hambrecht (2000)	1	1	0	1	0	3	0	2	1	0	1	1	11
Jolly (2009)	1	1	1	1	0	2	1	2	1	0	1	1	12
McKelvie (2002)	1	1	1	1	1	2	1	1	1	0	1	1	12
Mueller (2007)	1	0	0	1	0	2	1	2	1	0	1	1	10
Nilsson (2008)	1	1	0	1	1	2	1	2	1	0	0	1	11
O'Connor (2008)	1	1	1	1	1	3	1	2	1	1	0	1	14
Passino (2006)	1	0	0	1	0	2	1	2	1	0	1	1	10
Wielenga (1999)	1	0	0	1	0	2	1	2	1	0	0	1	9
Willenheimer (2001)	1	0	0	1	1	2	1	2	1	0	0	1	9
Witham (2005)	1	1	0	1	1	3	1	2	1	0	1	0	12
Witham (2012)	1	1	0	1	1	3	1	2	1	0	1	0	12
Yeh (2011)	1	1	0	1	1	3	1	2	1	1	0	0	12
Zanelli (unpublished)	††	††	††	††	††	††	††	††	††	††	††	††	††
Zwisler (2008)	1	1	0	1	1	3/3	1	2/2	1	0	0	0	11

*Three points possible; **If ITT was not specifically mentioned, but it was noted that no participants withdrew and all analysed 1 point was awarded; †Two points possible; ††Zanelli – not scored as no full publication

eTable 2. All-cause mortality - overall treatment effect and subgroups effects

	Primary analysis	**Secondary analysis**	**Sensitivity analyses**				
	Two-stage model, random effects HR (95% CI)p-value	**One-stage Cox model, stratified by study with fixed treatment effect** HR (95% CI) p-value	**Two-stage model, random effects excluding HF-Action** HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed effect excluding HF-Action** HR (95% CI) p-value	**Two-stage model, random effects 1 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 2 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 5 year truncation** HR (95% CI) p-value
Overall effect	0.84 (0.67, 1.04) p=0.107	0.85 (0.73, 0.99) p=0.034	0.81 (0.61, 1.06) p=0.129	0.79 (0.64, 0.97) p=0.027	0.87 (0.58, 1.31) p=0.507	0.86 (0.67, 1.10) p=0.217	0.84 (0.66, 1.06) p=0.140
Interaction terms							
Age (years)*	0.99 (0.98, 1.00) p=0.165	0.99 (0.98, 1.01) p=0.254	0.98 (0.96, 1.01) p=0.144	0.99 (0.96, 1.01) p=0.228	0.98 (0.95, 1.00) p=0.077	0.98 (0.96, 1.00) p=0.034	0.99 (0.97, 1.00) p=0.097
Gender (male vs female)	1.10 (0.73, 1.66) p=0.660	1.06 (0.70, 1.60) p=0.783	0.71 (0.35, 1.43) p=0.341	0.70 (0.36, 1.36) p=0.300	0.76 (0.34, 1.68) p=0.490	0.96 (0.55, 1.67) p=0.872	1.17 (0.75, 1.82) p=0.481
Ejection fraction (%)*	0.99 (0.97, 1.01) p=0.250	0.99 (0.97, 1.01) p=0.332	0.98 (0.95, 1.01) p=0.124	0.98 (0.96, 1.01) p=0.201	1.04 (1.00, 1.08) p=0.055	0.99 (0.97, 1.02) p=0.688	0.99 (0.97, 1.01) p=0.506
NYHA class (NYHA I/II vs NYHA III/IV)	0.80 (0.58, 1.11) p=0.182	0.79 (0.57, 1.08) p=0.134	0.82 (0.49, 1.38) p=0.459	0.75 (0.46, 1.22) p=0.244	0.50 (0.23, 1.07) p=0.073	0.84 (0.54, 1.30) p=0.431	0.83 (0.59. 1.18) p=0.297
HF aetiology (ischaemic vs	0.73 (0.38, 1.39)	1.19 (0.86, 1.64)	0.69 (0.36, 1.31)	0.87 (0.54, 1.41)	0.69 (0.19, 2.54)	0.79 (0.38, 1.67)	0.70 (0.33, 1.47)

non-ischaemic)	p=0.335	p=0.297	p=0.255	p=0.575	p=0.574	p=0.542	p=0.345
Ethnic group (white vs non-white)	1.12 (0.74, 1.69) p=0.593	1.11 (0.74, 1.68) p=0.604	†	1.05 (0.25, 4.31) p=0.949	0.72 (0.34, 1.53) p=0.396	0.83 (0.50, 1.38) p=0.468	1.12 (0.74, 1.69) p=0.593
Exercise capacity							
Peak VO_2 directly measured*	1.00 (0.95, 1.05) p=0.937	0.99 (0.95, 1.04) p=0.783	0.98 (0.90, 1.08) p=0.712	0.99 (0.91, 1.07) p=0.777	0.97 (0.88, 1.06) p=0.456	0.99 (0.93, 1.05) p=0.780	0.98 (0.91, 1.06) p=0.630
Peak VO_2, directly measured and predicted*	1.00 (0.95, 1.06) p=0.903	1.00 (0.96, 1.04) p=0.954	1.00 (0.91, 1.08) p=0.923	1.00 (0.93, 1.07) p=0.984	0.99 (0.90, 1.08) p=0.734	1.00 (0.94, 1.06) p=0.961	1.00 (0.93, 1.07) p=0.924
Standardised scores using peak VO_2, 6MWT, ISWT units, and watts*	1.03 (0.83, 1.27) p=0.802	1.02 (0.85, 1.22) p=0.851	0.99 (0.71, 1.39) p=0.955	1.01 (0.75, 1.35) p=0.967	0.97 (0.66, 1.41) p=0.858	1.00 (0.78, 1.30) p=0.972	1.01 (0.76, 1.35) p=0.938

*centered; †Study estimate not available as too few studies provide data; peak VO_2: peak oxygen uptake; ISWT: incremental shuttle walk test; 6MWT: 6-minute walk test.

eTable 3. HF-specific mortality - overall treatment effect and subgroups effects^

	Primary analysis	**Secondary analysis**	**Sensitivity analyses**		
	Two-stage model, random effects HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed treatment effect** HR (95% CI) p-value	**Two-stage model, random effects 1 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 2 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 5 year truncation** HR (95% CI) p-value
Overall effect	0.84 (0.48, 1.46) p=0.527	0.75 (0.44, 1.28) p=0.294	†	1.30 (0.59, 2.87) p=0.515	0.84 (0.49, 1.53) p=0.575
Interaction terms					
Age (years)*	0.96 (0.91, 1.02) p=0.206	0.96 (0.92, 1.01) p=0.162	†	0.91 (0.84, 0.98) p=0.017	0.95 (0.90, 1.00) p=0.066
Gender (male vs female)	0.53 (0.08, 3.73) p=0.524	0.61 (0.11, 3.49) p=0.583	†	†	†
Ejection fraction (%)*	0.95 (0.89, 1.02) p=0.159	0.96 (0.90, 1.02) p=0.179	†	1.01 (0.82, 1.24) p=0.912	0.96 (0.89, 1.04) p=0.309
NYHA class (NYHA I/II vs NYHA III/IV)	0.54 (0.07, 4.28) p=0.562	0.78 (0.23, 26.65) p=0.691	†	†	0.54 (0.07, 4.28) p=0.562
HF aetiology (ischaemic vs non-ischaemic)	Data only available for one study	3.30 (1.~~01~~02, 10.7) p=0.047	†	†	†
Ethnic group (white vs non-white)	†	†	†	†	†
Exercise capacity					
Peak VO_2, directly measured*	0.90 (0.76, 1.07) p=0.232	0.93 (0.78, 1.09) p=0.362	†	0.98 (0.73, 1.31) p=0.893	0.86 (0.69, 1.06) p=0.146
Peak VO_2, directly measured and	0.91 (0.77, 1.07) p=0.263	0.94 (0.80, 1.10) p=0.423	†	0.98 (0.76, 1.26) p=0.854	0.88 (0.72, 1.06) p=0.184

predicted*					
Standardised scores using peak VO_2, 6MWT, ISWT units and watts score	0.69 (0.35, 1.35) p=0.276	0.82 (0.43, 1.56) p=0.545	†	0.86 (0.31, 2.37) p=0.773	0.61 (0.28, 1.32) p=0.210

^HF-Action did not provide HF-mortality so sensitivity analysis of omission not undertaken; *centered; † Study estimate not available as too few studies provide data; peak VO_2: peak oxygen uptake; ISWT: incremental shuttle walk test; 6MWT: 6-minute walk test.

eTable 4. All-cause hospitalisation - overall treatment effect and subgroups effects

	Primary analysis	**Secondary analysis**	**Sensitivity analyses**				
	Two-stage model, random effects HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed treatment effect** HR (95% CI) p-value	**Two-stage model, random effects excluding HF-Action** HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed treatment effect excluding HF-Action** HR (95% CI) p-value	**Two-stage model, random effects 1 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 2 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 5 year truncation** HR (95% CI) p-value
Overall effect	0.90 (0.76, 1.06) p=0.210	0.91 (0.83, 1.01) p=0.072	0.86 (0.64, 1.14) p=0.293	0.85 (0.68, 1.09) p=0.210	0.94 (0.75, 1.18) p=0.583	0.91 (0.74, 1.11) p=0.330	0.90 (0.76, 1.06) p=0.210
Interaction terms							
Age (years)*	1.00 (0.99, 1.01) p=0.794	1.00 (0.99, 1.01) p=0.854	1.00 (0.98, 1.03) p=0.808	1.00 (0.98, 1.02) p=-0.969	1.00 (0.99, 1.01) p=0.636	1.00 (0.99, 1.01) p=0.798	1.00 (0.99, 1.01) p=0.794
Gender (male vs female)	1.09 (0.87, 1.36) p=0.454	1.09 (0.88, 1.36) p=0.424	0.66 (0.38, 1.14) p=0.136	0.68 (0.39, 1.16) p=0.158	1.05 (0.80, 1.37) p=0.745	1.15 (0.91, 1.46) p=0.239	1.09 (0.87, 1.35) p=0.454
Ejection fraction (%)*	1.00 (0.98, 1.01) p=0.629	1.00 (0.98, 1.01) p=0.646	1.00 (0.96, 1.04) p=0.857	1.00 (0.96, 1.05) p=0.831	1.00 (0.98, 1.01) p=0.632	0.99 (0.98, 1.01) p=0.343	1.00 (0.98, 1.01) p=0.629
NYHA class (NYHA I/II vs NYHA III/IV)	0.91 (0.74, 1.12) p=0.370	0.90 (0.73, 1.10) p=0.308	0.89 (0.43, 1.87) p=0.763	0.79 (0.39, 1.60) p=0.508	0.81 (0.63, 1.05) p=0.110	0.87 (0.70, 1.09) p=0.235	0.91 (0.74, 1.12) p=0.355

HF aetiology (ischaemic vs non-ischaemic)	0.96 (0.71, 1.31) p=0.810	1.00 (0.82, 1.22) p=0.988	0.73 (0.39, 1.39) p=0.340	0.73 (0.40, 1.31) p=0.284	1.08 (0.84, 1.38) p=0.562	1.04 (0.84, 1.29) p=0.723	1.01 (0.83, 1.24) p=0.910
Ethnic group (white vs non-white)	1.02 (0.83, 1.26) p=0.860	1.02 (0.83, 1.26) p=0.852	1.02 (0.47, 2.21) p=0.959	1.06 (0.49, 2.32) p=0.879	1.14 (0.88, 1.48) p=0.322	1.06 (0.85, 1.33) p=0.607	1.02 (0.83, 1.26) p=0.860
Exercise capacity							
Peak VO_2, directly measured*	1.01 (0.99, 1.04) p=0.259	1.02 (0.99, 1.04) p=0.234	1.05 (0.95, 1.16) p=0.352	1.06 (0.96, 1.17) p=0.262	1.03 (0.99, 1.06) p=0.124	1.02 (0.99, 1.04) p=0.243	1.01 (0.99, 1.04) p=0.259
Peak VO_2, directly measured and predicted*	1.02 (0.99, 1.04) p=0.153	1.02 (0.99, 1.04) p=0.134	1.07 (0.98, 1.17) p=0.125	1.08 (0.99, 1.17) p=0.078	1.03 (1.00, 1.06) p=0.057	1.02 (0.99, 1.05) p=0.129	1.02 (0.99, 1.04) p=0.153
Standardised scores using peak VO_2, 6MWT, ISWT units and watts*	1.09 (0.98, 1.22) p=0.095	1.10 (0.99, 1.22) p=0.088	1.30 (0.93, 1.83) p=0.120	1.32 (0.95, 1.82) p=0.097	1.16 (1.02, 1.33) p=0.027	1.11 (0.99, 1.24) p=0.077	1.09 (0.98, 1.22) p=0.095

*centered; peak VO_2: peak oxygen uptake; ISWT: incremental shuttle walk test; 6MWT: 6-minute walk test.

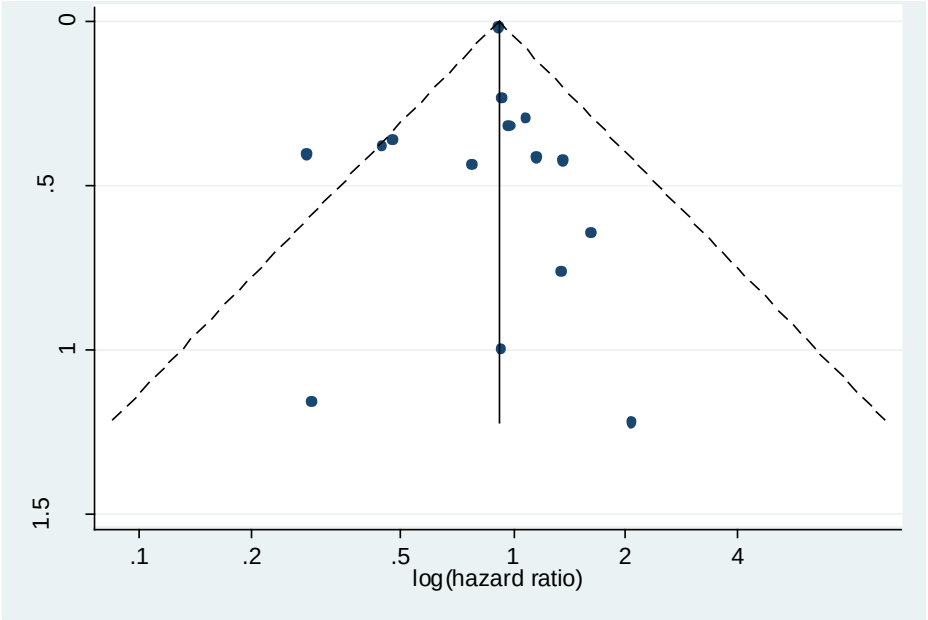
eTable 5. HF-specific hospitalisation - overall treatment effect and subgroups effects in studies included in IPD meta-analysis

	Primary analysis	**Secondary analysis**	**Sensitivity analyses**				
	Two-stage model, random effects HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed treatment effect** HR (95% CI) p-value	**Two-stage model, random effects excluding HF-Action** HR (95% CI) p-value	**One-stage Cox model, stratified by study with fixed treatment effect excluding HF-Action** HR (95% CI) p-value	**Two-stage model, random effects 1 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 2 year truncation** HR (95% CI) p-value	**Two-stage model, random effects 5 year truncation** HR (95% CI) p-value
Overall effect	0.98 (0.71, 1.34) p=0.902	0.94 (0.81, 1.08) p=0.368	1.00 (0.65, 1.54) p=0.999	1.03 (0.79, 1.35) p=0.829	1.08 (0.88, 1.33) p=0.470	1.06 (0.83, 1.34) p=0.658	0.97 (0.70, 1.34) p=0.855
Interaction terms							
Age (years)*	1.00 (0.99, 1.02) p=0.603	1.00 (0.99, 1.02) p=0.632	1.00 (0.98, 1.03) p=0.958	1.00 (0.97, 1.02) p=0.906	1.00 (0.99, 1.02) p=0.640	1.00 (0.99, 1.02) p=0.611	1.00 (0.99, 1.02) p=0.580
Gender (male vs female)	1.03 (0.73, 1.46) p=0.865	0.99 (0.71, 1.39) p=0.949	0.70 (0.32, 1.53) p=0.372	0.65 (0.33, 1.29) p=0.215	0.76 (0.46, 1.24) p=0.274	1.06 (0.68, 1.66) p=0.803	0.93 (0.49, 1.75) p=0.815

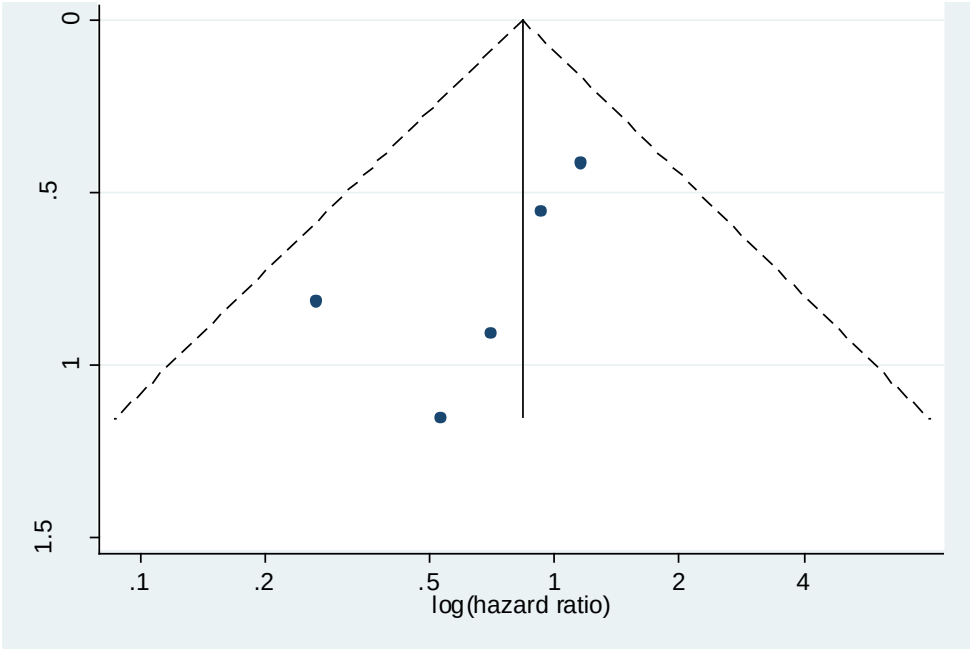
Ejection fraction (%)*	0.51 (0.14, 1.79) p=0.291	0.99 (0.97, 1.01) p=0.325	0.99 (0.96, 1.03) p=0.540	0.99 (0.97, 1.01) p=0.350	0.99 (0.97, 1.02) p=0.569	0.99 (0.97, 1.01) p=0.350	0.99 (0.97, 1.02) p=0.569
NYHA class (NYHA I/II vs NYHA III/IV)	1.55 (0.79, 3.02) p=0.200	1.14 (0.84, 1.54) p=0.399	2.05 (0.86, 4.92) p=0.107	1.74 (0.92, 3.29) p=0.089	0.81 (0.51, 1.29) p=0.375	1.17 (0.68, 2.03) p=0.573	1.21 (0.72, 2.04) p=0.475
HF aetiology (ischaemic vs non-ischaemic)	1.20 (0.64, 2.25) p=0.577	1.28 (0.94, 1.74) p=0.111	0.95 (0.31, 2.95) p=0.928	1.10 (0.57, 2.16) p=0.771	1.47 (0.94, 2.29) p=0.128	1.28 (0.90, 1.84) p=0.172	1.29 (0.79, 2.12) p=0.309
Ethnic group (white vs non-white)	1.18 (0.85, 1.65) p=0.318	1.19 (0.86, 1.66) p=0.291	†	1.79 (0.60, 5.37) p=0.301	1.25 (0.79, 1.98) p=0.334	1.20 (0.83, 1.74) p=0.327	1.18 (0.85, 1.65) p=0.318
Exercise capacity							
PeakVO_2, directly measured*	0.97 (0.88, 1.07) p=0.538	0.97 (0.93, 1.01) p=0.149	0.94 (0.80, 1.11) p=0.467	0.98 (0.90, 1.07) p=0.658	0.99 (0.90, 1.10) p=0.882	0.99 (0.93, 1.06) p=0.769	0.98 (0.89, 1.08) p=0.685
Peak VO_2, directly measured and predicted*	0.97 (0.89, 1.07) p=0.539	0.97 (0.93, 1.01) p=0.116	0.95 (0.82, 1.10) p=0.483	0.97 (0.89, 1.05) p=0.424	0.98 (0.92, 1.05) p=0.610	0.99 (0.94, 1.03) p=0.535	0.98 (0.90, 1.07) p=0.670
Standardised scores using peak VO_2, 6MWT, ISWT units and watts*	0.88 (0.62, 1.26) p=0.483	0.86 (0.72, 1.03) p=0.093	0.83 (0.46, 1.49) p=0.527	0.81 (0.56, 1.16) p=0.246	0.92 (0.69, 1.23) p=0.576	0.93 (0.75, 1.16) p=0.517	0.91 (0.69, 1.20) p=0.505

*centered; †Study estimate not available as too few studies provide data; peak VO_2: peak oxygen uptake; ISWT: incremental shuttle walk test; 6MWT: 6-minute walk test.

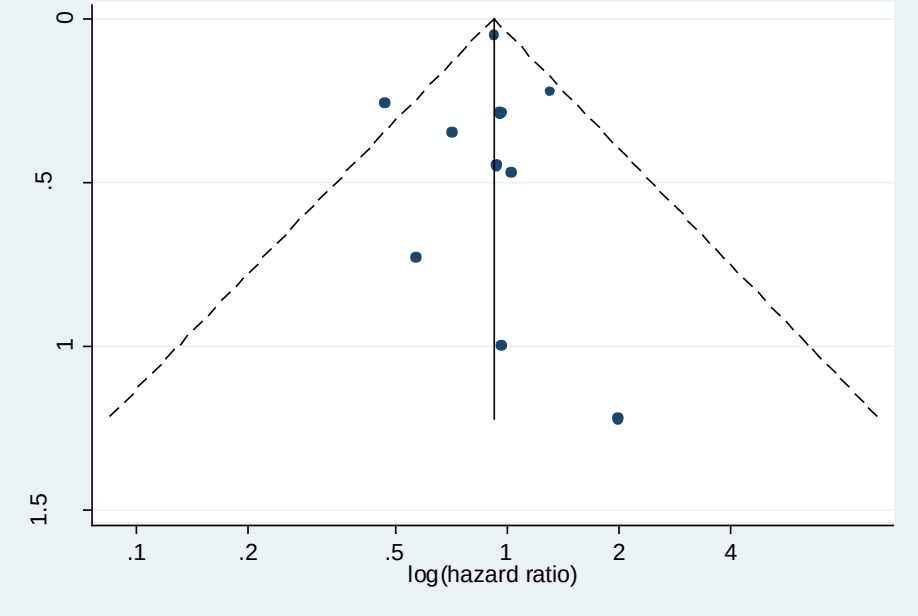
eFigure 1. Funnel plots



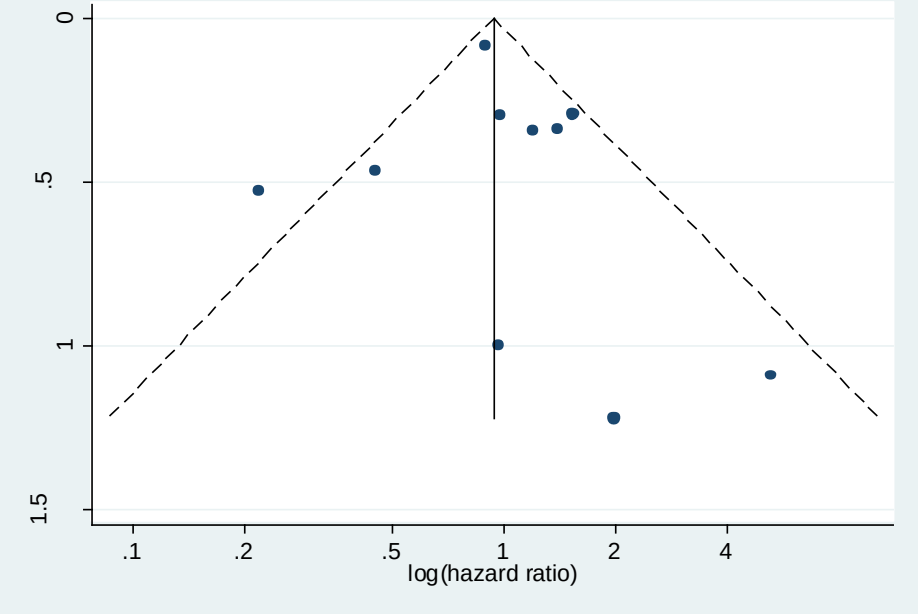
a. All-cause mortality
Egger test -0.26, p=0.458



b. HF-specific mortality
Egger test -1.60, p=0.147



c. All-cause hospitalisation
Egger test 0.16, p=0.739



d. HF-specific hospitalisation
Egger test 0.32, p=0.610