REMOTE ISCHAEMIC CONDITIONING REDUCES INFARCT SIZE IN ANIMAL IN VIVO MODELS OF ISCHAEMIA-REPERFUSION INJURY: A SYSTEMATIC REVIEW AND META-ANALYSIS

Daniel I Bromage*, Jack MJ Pickard*, Xavier Rossello, Oliver J Ziff, Niall Burke, Derek M Yellon, Sean M Davidson

*These authors contributed equally to this work

The Hatter Cardiovascular Institute, 67 Chenies Mews, London, WC1E 6HX, United Kingdom

Correspondence to:

Prof Derek M. Yellon

The Hatter Cardiovascular Institute, 67 Chenies Mews, London, WC1E 6HX, United Kingdom
d.yellon@ucl.ac.uk

Tel: +44 203 447 9591

Fax: +44 203 447 9818

Word count (including abstract, manuscript text, references, and figure legends):

7295

© The Author 2016. Published by Oxford University Press on behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
ABSTRACT

Aims

The potential of remote ischaemic conditioning (RIC) to ameliorate myocardial ischaemia-reperfusion injury (IRI) remains controversial. We aimed to analyse the pre-clinical evidence base to ascertain the overall effect and variability of RIC in animal in vivo models of myocardial IRI. Furthermore, we aimed to investigate the impact of different study protocols on the protective utility of RIC in animal models and identify gaps in our understanding of this promising therapeutic strategy.

Methods and Results

Our primary outcome measure was the difference in mean infarct size between RIC and control groups in in vivo models of myocardial IRI. A systematic review returned 31 reports, from which we made 22 controlled comparisons of remote ischaemic preconditioning (RIPreC) and 21 of remote ischaemic perconditioning and postconditioning (RIPerC/RIPostC) in a pooled random-effects meta-analysis. In total, our analysis includes data from 280 control animals and 373 animals subject to RIC. Overall, RIPreC reduced infarct size as a percentage of area at risk by 22.8% (95% CI 18.8-26.9%), when compared to untreated controls (P<0.001). Similarly, RIPerC/RIPostC reduced infarct size by 22.2% (95% CI 17.1-25.3%; P<0.001). Interestingly, we observed significant heterogeneity in effect size (T2=92.9% and I2=99.4%; P<0.001) that could not be explained by any of the experimental variables analysed by meta-regression. However, few reports have systematically characterised RIC protocols, and few of the included in vivo studies satisfactorily met study quality requirements, particularly with respect to blinding and randomisation.

Conclusions

RIC significantly reduces infarct size in in vivo models of myocardial IRI. Heterogeneity between studies could not be explained by the experimental variables tested, but studies are limited in number and lack consistency in quality and study design. There is therefore a clear need for more well-performed in vivo studies with particular emphasis on detailed characterisation of RIC protocols and investigating the potential impact of gender. Finally,
more studies investigating the potential benefit of RIC in larger species are required before translation to humans.

INTRODUCTION

Myocardial ischaemia-reperfusion injury describes the deleterious consequences of several pathological processes and cardiac interventions. Most commonly, it is caused by thrombotic occlusion of the coronary artery in ST-segment elevation myocardial infarction (STEMI) and subsequent reperfusion by primary percutaneous coronary intervention (PPCI), but it may result from a range of elective and emergent causes of myocardial ischaemia, including cardiopulmonary bypass and spontaneous reperfusion of STEMI. Despite constantly improving medical and surgical practice, myocardial IRI remains associated with significant morbidity and mortality. For example, in STEMI and despite PPCI, 30-day, 1-year, and 5-year cardiac mortality remains 7.3 %, 8.4 %, and 13.8 %, respectively.\(^1\)

It has been demonstrated that myocardium can be protected from lethal IRI by the application of multiple brief cycles of ischaemia and reperfusion to an organ or tissue remote from the heart, either before, during, or after the index ischemia (preconditioning, perconditioning or postconditioning, respectively). Limb remote ischaemic conditioning (RIC) is a cheap, non-invasive intervention that, since its inception in 1997,\(^2\) has been successfully demonstrated in several pre-clinical studies of myocardial IRI. Subsequently, several phase II, proof-of-concept clinical studies have translated these findings to variety of clinical settings, albeit frequently using cardiac enzymes as surrogate markers of cellular injury, including coronary artery bypass surgery (CABG),\(^3\)\(^-\)\(^6\) elective abdominal aortic aneurysm repair,\(^7\)\(^,\)\(^\)\(^8\) elective cervical decompression surgery,\(^9\) elective PCI,\(^10\) and in PPCI for STEMI.\(^11\)\(^-\)\(^13\)

However, more recent, large clinical-endpoint studies of RIC in cardiac surgery have been neutral.\(^14\),\(^15\) Although cardiac surgery may be an inappropriate setting for RIC, given the small peri-operative injury and lack of injurious warm ischemia-reperfusion,\(^16\) these findings
have prompted an interrogation of the pre-clinical evidence base for RIC and a perceived lack of systematic pre-clinical characterisation of the optimal RIC stimulus.\textsuperscript{16} This is in contrast to direct ischaemic conditioning that, despite being limited by the necessity to intervene before the index ischaemia, has been thoroughly characterised.\textsuperscript{17}

Well-designed animal studies can provide useful information on relevant factors that may influence outcome. This comprehensive systematic review and meta-analysis scrutinizes basic studies of RIC in \textit{in vivo} animal models of myocardial IRI. Our aim was to ascertain the overall effect and variability of RIC in this context, compared to control (sham procedure or no treatment). We further aimed to investigate determinants of efficacy, including variables such as RIC protocol and use of supplementary oxygen. Our hypothesis was that study quality and publication bias would result in over-estimation of the effect size associated with RIC.\textsuperscript{18,19}

**METHODS**

**Systematic review**

The systematic review was performed in accordance with Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines.\textsuperscript{20} A literature search was conducted on 21\textsuperscript{st} August, 2015 by JP. Keywords and MeSH terms were used to search Medline and Embase (via OVID) between 1997-present, and further studies were identified by consultation with experts in the field. Details of the search strategy are available in the Supplementary material, \textit{Search strategy}.

Study eligibility criteria were defined using the PICOS approach.\textsuperscript{21} \textit{In vivo} animal studies were included and were eligible if they investigated the effect of limb RIC (pre-, per- or post-) versus a control (sham procedure or no treatment) on myocardial infarct size, as measured using tetrazolium chloride (TTC),\textsuperscript{22} in any mammalian species, regardless of study design. Transient infra-renal aortic occlusion was considered as bilateral hind limb ischaemia.

Studies were excluded if they did not include or report absolute myocardial infarct size (IS) as a percentage of area at risk (AAR, defined as the myocardial tissue within the vascular
territory that is distal to occluded artery and, if not reperfused, is at risk of irreversible ischaemic death). The AAR varies depending on the exact position of the LAD suture and variable LAD anatomy. Infarct size has a strong positive correlation with AAR and therefore, without correction, a small AAR could create false-positive results for cardioprotection, and vice versa. Furthermore, studies were excluded if they specifically investigated only the ‘second window’ of cardioprotection (RIC to infarction interval > 1 h), if RIPostC was initiated more than 10 min after reperfusion (after which it is generally believed unlikely to be effective), if the animals had co-morbidities, if infarct size was only measured using a method other than TTC, or if they investigated the impact of RIC in the context of heart transplant. Groups in which RIC was administered in combination with another conditioning protocol (local conditioning, for example), or with pharmacological treatments known to have cardioprotective effects, were excluded. Finally, studies investigating neonatal animals were excluded to ensure clinical relevance to IRI.

Reports were excluded if they were not available in English and a publication date restriction of 1997-present was imposed in view of the first publication of the efficacy of RIC in the limb. Review articles, abstract articles, unpublished material and ongoing studies were excluded.

Retrieved records were screened for eligibility using the title and abstract, followed by the full text. Eligibility assessment was performed independently in an un-blinded, standardized manner by DB and JP. Disagreements were resolved by examining the full text of the article or by consensus between reviewers in all cases.

Data was independently extracted by 2 authors (DB and JP) using predefined data fields. Disagreements were resolved by consensus in all cases. We attempted to acquire key missing information by contacting the report authors by e-mail.

Variables for which data were sought were developed using the PICOS approach. Data items were chosen according to experimental variables with evidence for an effect on myocardial IRI as these were considered likely to impact on the efficacy or RIC, and variables
that we considered of potential importance. Full list of data items and assumptions are available in the Supplementary material, Data items.

Meta-analysis

We defined the primary outcome as the weighted (unstandardized) mean difference (WMD) between infarct sizes in the RIC and control groups. WMD was used as all data were presented in the same units and it gives a biologically relevant value. In each publication, we identified all independent comparisons of IS/AAR% in RIC versus control groups. Where a study made multiple comparisons to the same control group, the size of the control group was corrected for the number of comparisons made (n/number of comparisons). The secondary outcome was the effect of 5 pre-defined experimental variables, which we considered most likely to impact on the efficacy of RIC, on WMD.

Comparisons were grouped according to their use of either remote ischaemic preconditioning (RIPreC) or remote ischemic perconditioning and postconditioning (RIPerC/RIPostC). This is due to temporal differences in their application that, despite not necessarily occurring via different mechanistic pathways, have different clinical utility. Specifically, RIPerC/RIPostC are clinically applicable to STEMI, whereas RIPreC is not. Subsequent analysis was performed for each group separately.

For each independent comparison, we calculated the effect size as a raw difference in IS/AAR% means (the mean of the control groups minus the mean of the experimental group) and the corresponding 95% confidence interval (CI). To account for anticipated heterogeneity, we pooled effect sizes using random-effects meta-analysis, which considers the within-study and between-study variability and weights each study accordingly. Heterogeneity was quantified using I2 and T2 statistics. Studies with missing data on any of the pre-defined experimental variables were excluded from the meta-analysis.

Subgroup analyses were performed using univariate meta-regressions to explore which experimental factors and quality indicators contribute to heterogeneity. The percentage of between-study variance explained by variables of interest was assessed using the T2 and adjusted R2 statistics. The significance level was adjusted according to the number of
comparisons using the Holm-Bonferroni method and results were considered significant when \( P<0.01 \).\textsuperscript{32} Classification of subgroups is given in the Supplementary material, \textit{Summary measures}.

Sensitivity analysis was performed to assess the robustness of our findings by performing an additional analysis for both the primary and the secondary endpoints using the standardized mean difference (SMD; the mean of the control group minus the mean of the RIC group, divided by the pooled SD of the two groups). We performed a stratified meta-analysis by subgroup to validate the results obtained by meta-regression.

All analyses were pre-specified and performed using STATA/SE, version 13.1 (StataCorp, College Station, TX, USA); GraphPad Prism version 5.00 for Windows (CA, USA) was used in the production of figures.

\textbf{Risk of bias}

We used a component approach to assess study quality, based on the study report, using the ‘Animal Research: Reporting of \textit{In Vivo} Experiments’ (ARRIVE) guidelines and a 12-item quality score.\textsuperscript{33-35} The derivation of the 12-item quality score is given in Supplementary materials, \textit{Table S3} and the methodological features evaluated for each study are given in Supplementary materials, \textit{Tables S4-5}. Study quality was assessed independently from data extraction and between assessors in an un-blinded, standardized manner by 2 reviewers (OZ and NB). Disagreements were resolved by consensus in all cases. We assessed the relationship between study quality and the overall effect of RIC on IS/AAR\% using meta-regression, as described above.

Potential publication bias was assessed by visual inspection of a funnel plot for asymmetry, and Egger’s regression analysis for small study effects. No protocols were available with which to examine for selective reporting; however, the methods and results sections of all included studies were carefully compared for inconsistencies.

\textbf{RESULTS}
Study selection

Our search returned 539 records, including 169 duplicate reports (consisting of reports returned by both Medline and Embase). 370 reports underwent title and abstract screening, which resulted in 256 exclusions. The remaining 114 reports were retrieved for detailed full text evaluation. 83 articles were excluded, 66 due to failing to meet the inclusion criteria, 12 were abstracts and 2 were not retrievable. Of the remaining 34 reports (studies), 3 were missing data on one or more important experimental variables that we were unable to retrieve by contacting the study authors, and were consequently excluded. These studies are referenced in the Supplementary material, References. The remaining 31 studies were included in the quantitative synthesis (Fig 1). All included studies and their main characteristics are reported in the Supplementary material, Data extraction table.

Meta-analysis

From the 31 included reports, we extracted data on 43 controlled comparisons of RIC in models of myocardial IRI. These were split into 22 comparisons investigating RIPreC, and 21 comparisons investigating RIPerC/RIPostC. In total, our analysis includes data from 280 control animals and 373 animals undergoing RIC. In the RIPreC group, conditioning reduced IS/AAR% by 22.8% (95% CI 18.8-26.9%) when compared to untreated controls (P<0.001; n=22 comparisons, Fig 2A). Significant heterogeneity was observed (T2=89.2 and I2=99.1%; P<0.001). In the RIPreC/RIPostC group, conditioning reduced IS/AAR% by 22.2% (95% CI 17.1-25.3%) when compared to untreated controls (P<0.001; n=21 comparisons, Fig 2B). Again, significant heterogeneity was observed (T2=90.9 and I2=99.5%; P<0.001).

We investigated potential experimental sources of the observed heterogeneity using meta-regression analysis with IS/AAR% as the dependent variable, and did not find any significant associations with efficacy of RIC (Fig 3A-B).

Risk of bias

Reports achieved a median ARRIVE guidelines score of 14 (interquartile range 12-14) out of 20 and a median 12-item quality score of 7 (interquartile range 6-8; Fig 4). A full breakdown of the scores is given in the Supplementary material, Tables S7-8. Meta-regression indicated
that study quality according to either the ARRIVE guidelines score or to a 12-item quality score was not associated with the overall effect (P=0.317 and P=0.846, respectively).

Overall, studies performed particularly poorly in several important areas of experimental design. For example, with respect to the ARRIVE guidelines only 16% reported sample size calculation, 23% reported randomization of animals to experimental groups, 23% defined the primary experimental outcome and 29% reported detail of adverse events during experiments (Fig 4A). Regarding the 12-item quality score, only 32% of studies reported measurement of blood oxygen saturation, 13% reported measurement of blood pressure during the in vivo protocol and 10% performed blinded application of the conditioning protocol (Fig 4B). Notably, only 38% of studies performed a blinded assessment of outcome.

The impact of publication bias on the overall effect was assessed by visual analysis of the funnel plot, which suggested that small and negative studies might be under-represented (Fig 5). However, in Egger’s regression test the null-hypothesis of no small-study effect was not rejected at P=0.216 (estimated bias coefficient 3.75±2.98 SE).

**Sensitivity analysis**

When re-running our analysis using the SMD, all results were similar to those found using the WMD. We found a highly significant (P<0.001) overall effect of RIPreC (SMD of 11.06; 95% CI 8.52-13.60), as well as a similar level of heterogeneity (I²=91.1%). None of the experimental variables were significant after correction by multiple comparison. For RIPerC/RIPostC, all results using the SMD were likewise similar to those found using the WMD. We found a highly significant (P<0.001) overall effect (SMD of 13.14; 95% CI 10.51-15.77) and similar heterogeneity (I²=93.3%). None of the experimental variables were significant after correction by multiple comparison.

**DISCUSSION**

Our major finding is that both RIPreC and RIPerC/RIPostC have a large beneficial effect of both on infarct size. This finding is based on a comprehensive systematic review, including
over 650 animals. However, there were relatively few pre-clinical studies investigating RIC and even fewer systematically characterising the protocol, which limited our analysis.

Furthermore, we found inconsistency in the design of pre-clinical RIC studies with very few that were randomized effectively, included both male and female animals, and were double blinded with respect to procedure and outcome, which highlights the need for further well-designed, pre-clinical studies of RIC.

These are important findings in the context of recent pessimism regarding RIC as a genuine cardioprotective phenomenon. Clinical trials investigating the efficacy of RIC in myocardial IRI have had mixed results, which has been attributed to clinical variables, including propofol administration. However, each of the factors described above may influence outcome through lack of generalizability or unconscious bias, and therefore indirectly impact upon attempts to translate RIC to humans in clinical trials.

**Determinants of RIC**

Interestingly, we found high levels of heterogeneity between studies. To investigate whether study protocol could account for the observed heterogeneity, and to elucidate the determinants of efficacy of RIC, we assessed the impact of experimental variables on effect size using a meta-regression analysis. This approach has been successfully applied to several promising pre-clinical interventions to date. For example, Lim et al. were able to demonstrate that ciclosporin was not effective at limiting myocardial infarct size in pig models of IRI, compared to small animal models. This finding might be important in the context of a subsequent neutral clinical study of ciclosporin before reperfusion in patients with STEMI. Therefore, finding the parameters responsible for heterogeneity can guide pre-clinical and clinical study design.

However, none of the tested experimental variables, which included species, cycle duration, number of cycles, number of conditioned limbs and the use of supplementary oxygen, were associated with effect size. In a recent analysis of pre-clinical studies of local ischaemic preconditioning, which included limited analysis of RIC, species (rodent vs. non-rodents) accounted for a substantial amount of the observed heterogeneity. Although we found no
significant association between animal size (rodents vs. non-rodents), there was a pattern of increased efficacy of RIPreC in large animals.

Wever et al. also describe no association between effect size and the number, timing and duration of cycles, which is reflected in the present study.\textsuperscript{17} Specifically, we found 1, 3 and 4 cycles to be equally effective. Conversely, in one of the few neutral comparisons in our analysis, Lu et al. utilised a protocol consisting of 1 cycle of 5 min RIPreC.\textsuperscript{40} Interestingly, this amounts to the lowest total ischaemic ‘dose’ (a function of cycle number and duration) of any study of RIPreC. Amongst the RIPerC/RIPostC studies, only Li et al. applied 1 cycle of 5 min limb ischaemia and reported a relatively modest, albeit significant, effect (WMD 14.35\%, 95\% CI 13.03-15.67\%).\textsuperscript{41} Others, including Mastitskaya et al., have applied a single cycle of longer duration and achieved greater protection.\textsuperscript{42} It is therefore plausible that ischaemic ‘dose’ or burden, rather than cycle number or duration alone, is the dependent variable, but very few studies have examined this systematically \textit{in vivo}. An exploratory analysis of the present data demonstrated no such association (data not shown), however this was limited by a narrow distribution of total ischaemic times in our dataset. Interestingly, a dose-dependent effect of RIPreC has been observed, with greater cardioprotection after 10 min or 15 min compared to 5 min continuous infra-renal aortic occlusion.\textsuperscript{43} In a study comparing bilateral and unilateral RIPreC in protection against renal ischaemia-reperfusion injury, bilateral was found to be more effective.\textsuperscript{44} Taken together, this may suggest that total ischemic dose is important, but this needs testing in specifically designed experiments.

Furthermore, there may be an upper limit to what is an effective ischaemic dose. A recent characterisation of the RIC protocol \textit{ex vivo} reported that 4 and 6, but not 8, cycles of RIC were protective in mice.\textsuperscript{45} Similarly, although 5 min or 10 min limb ischaemia protected against subsequent liver injury, 30 min and 60 min actually increased injury.\textsuperscript{46} This issue is worthy of further attention, particularly using a more realistic setting of aged animals with co-morbidities. This may suggest a therapeutic window for RIC but there is very limited pre-clinical evidence, in contrast to direct ischaemic conditioning where the number of cycles are demonstrably important,\textsuperscript{17} and it is clear that detailed characterisation of study protocol \textit{in vivo} is urgently necessary in order to answer these questions.
After a period of limb ischaemia, it has been assumed that reperfusion is necessary, either to wash out the putative humoral factor, or because reperfusion-induced ROS may be necessary to activate signalling pathways. Additionally, some studies have suggested that reactive hyperaemia is an important factor in the response. However it is worth bearing in mind that the original description in rabbits, where RIPreC was achieved by 55-65% stenosis of the femoral artery in combination with rapid electrical stimulation of the gastrocnemius muscle for 30 min, did not actually involve reperfusion. This appears to support the hypothesis that multiple mechanisms are involved.

Similarly, our analysis supports the finding of Wever et al. of no association with the number of limbs conditioned. Johnsen et al. further reported that 2 and 5 min, in contrast to 10 min, cycles were beneficial in mice. This remains consistent with the concept of a therapeutic window of RIC but there is a paucity of in vivo experimental data relating to the precise RIC protocol. Our analysis demonstrated 10 min cycles to be equally effective as 5 min cycles, albeit only in larger species as no studies in our meta-regression used 10 min cycles in mice. We also noticed a pattern of reduced efficacy in RIPreC/RIPostC studies using 2 limbs instead of one, which may exceed the therapeutic window. However, this finding was not statistically significant and should only be considered hypothesis-generating.

The use of supplementary oxygen in acute myocardial infarction is controversial, but the role of supplementary oxygen in RIC has not, to our knowledge, been investigated. It could be hypothesised that RIC is driven by cellular hypoxia in the conditioned limb which, at least in part, might be alleviated by ventilation with supplementary oxygen. However, despite finding the conditioning of animals with and without oxygen to be equally effective, there was a pattern of increased efficacy in animals ventilated with supplementary oxygen in the RIPreC comparisons. This has not been specifically investigated but our results would suggest further investigation is warranted.

Furthermore, there is evidence that the degree of cardioprotection conferred by RIC is proportional to the duration of index ischemia. Specifically, Kleinbongard et al. demonstrated longer ischaemic times to be associated with a greater efficacy of RIC, albeit in a study of patients undergoing CABG, which may be related to a greater target for
protection.\textsuperscript{50} In an exploratory analysis, we found no statistical effect of index ischaemic time (Supplementary material, Table S5-6); however, there was a pattern of reduced efficacy of RIC at longer ischaemic durations in RIPreC but not RIPerC/RIPostC. This might suggest an important role for the timing of intervention, but should be interpreted with caution in view of the limited number of comparisons available after stratification according to study protocol.

There were several experimental variables that were of considerable interest but that lacked sufficient power for statistical analysis. For example, it has been suggested that propofol interferes with the development of RIC,\textsuperscript{51, 52} and it has been implicated as a potential reason for the apparent lack of translation of RIC.\textsuperscript{53} Interestingly, we observed studies where propofol (together with either opioid analgesia ± pancuronium) was administered to have effect sizes above the mean.\textsuperscript{54, 55}

Furthermore, studies reporting the use of mixed gender experimental groups reported apparently smaller effect sizes, and one of the few neutral studies was performed only in female animals.\textsuperscript{56} Gender can potentially impact upon ischaemia-reperfusion injury possibly due to the cardioprotection conferred by oestrogen,\textsuperscript{57, 58} and potential temporal variability in cardioprotection as a result of the oestrous cycle of female rats.\textsuperscript{59} However, investigation of the role of gender is conspicuously absent from pre-clinical studies of RIC and these qualitative data require further testing in formal, well-designed studies.

An important result of the present study is the need for more pre-clinical characterisation of RIC, which as well as the aforementioned variables should include the potential role of occlusion technique (arterial clamping vs. external compression), which may be important due to the possible involvement of limb collaterals,\textsuperscript{60} the putative effect of shear stress in RIC,\textsuperscript{47} and the role of pre-treatment with heparin, which has been reported to be cardioprotective in the context of IRI.\textsuperscript{61-65} Likewise, it has been hypothesised that the interval between the conditioning stimulus and the onset of index ischaemia might be an important determinant of the efficacy of remote preconditioning.\textsuperscript{16}
A further consideration is the impact of co-morbidities and co-medications on the efficacy of RIC. The animal models used in pre-clinical studies frequently do not reflect the complex risk factor, co-morbidity and pharmaceutical profile of humans with cardiovascular disease, which may impede the development of RIC. For example, Jensen et al. treated isolated perfused rabbit hearts subjected to IRI with plasma dialysate from diabetic patients treated with RIC. They were unable to confer cardioprotection using dialysate from patients with peripheral neuropathy compared to non-diabetic patients and diabetic patients without neuropathy. Baranyai et al. similarly found that acute hyperglycaemia abrogated the beneficial effect of RIPerC in Wistar rats. However, despite considerable investigation of these features in local ischaemic preconditioning (reviewed by Ferdinandy et al.), there is little experimental data relating to these factors in the context of RIC. Indeed, animals with co-morbidities were specifically excluded from the present analysis due to the small number of studies available. Nonetheless, co-morbidities should be considered when investigating the optimum protocol for the delivery of RIC.

**Risk of bias**

Poor methodological quality and publication bias can result in over-estimation of effect size. In turn, this can engender enthusiasm about the benefit of a treatment where, in fact, none exists. It is therefore essential to examine the impact of study quality on size of effect, to which end we found no statistical relationship using meta-regression.

However, we made some interesting observations. Aspects of the report relating to the experimental procedure, including control of temperature and recording of the ECG, were generally well reported. However, there was generally poor observation of the ARRIVE guidelines, particularly in relation to reporting of sample size calculations, randomisation, blinding and adverse events, which can result in selective exclusion. For example, in an interesting meta-analysis of systematic reviews, Hirst et al. reported that failure to randomise significantly increased effect size. Furthermore, appropriate monitoring of experimental animals, including recording of blood oxygen saturation and blood pressure, were poorly reported. These facets are clearly essential to ensure good quality research and were a central tenet of a recent position paper on improving the pre-clinical assessment of
novel cardioprotective therapies. In some cases these omissions will represent inadequate reporting but in others it is likely that these crucial elements of study design were not performed. We elected not to analyse each quality criterion independently due to insufficient reporting and to avoid false positive findings due to multiple comparisons. However, poor adherence to certain quality criteria may account for the heterogeneity observed in this meta-analysis.

Finally, our assessment of publication bias by visual analysis of the funnel plot suggested that small, neutral studies may be under-represented; however, this did not statistically impact on the overall effect size, which is reassuring.

Limitations

The validity of this meta-analysis is contingent upon the quality of reporting of the included studies. Unpublished studies and those with missing data could not be included in the meta-analysis, and others did not meet important quality criteria including poor information regarding statistical analysis and blinding. However, the absence of a statistical impact from study quality or publication bias is reassuring in this regard. We were limited by being unable to consider manuscripts not available in English and we acknowledge that we did not carry out a systematic literature review to determine which experimental variables to include in our analyses, which might be subject to selection bias. A relatively small number of studies were included in the meta-analysis, thereby limiting the power of the study, which was further affected by multiple comparisons within individual studies; however, we elected to include all comparisons to avoid selection bias. Meta-regression is inherently limited; however, to ensure this was as robust as possible we performed a stratified meta-analysis by subgroup that yielded similar results, including a highly significant overall effect of RIC, significant heterogeneity, and no effect of any of the experimental variables we included in our model.

RECOMMENDATIONS AND CONCLUSIONS

This systematic review and meta-analysis of pre-clinical in vivo studies of myocardial IRI demonstrates a significant and highly reproducible beneficial effect of RIC. This effect was
highly heterogeneous, a finding that may be due to un-measurable, multifactorial differences between individual experimenters and laboratories. However, importantly, in vivo studies to date suggest the optimal RIC stimulus has not yet been identified. There was a paucity of pre-clinical characterisation of the RIC protocol and poor reporting of quality indicators. This is important not so the protocol can be translated to humans, but in order to understand the important parameters and/or markers that will facilitate optimization of the protocol in humans.

There has been a great deal of debate regarding neutral clinical studies of RIC.\textsuperscript{16,36} However, before we try to understand these failings and design future studies it is essential to fully describe RIC in pre-clinical experiments. At present, studies variably (and apparently randomly) apply the intervention to one or two limbs, for varying periods of time, with a variable number of cycles and with inconsistent timing with respect to the injurious ischemic episode. To this end, we have identified a need for more, well-performed studies with a focus on characterisation rather than detailed elucidation of mechanisms. In particular, these should concentrate on investigating the potential impact of gender and the number, timing and duration of cycles on the efficacy of RIC. These aims would be greatly aided by the identification of a biomarker or a critical physiological parameter(s) that correlates with protection. More studies investigating the potential benefit of RIC in larger species are required before translation to humans and we highlight the potential role of supplementary oxygen as particularly interesting for exploration. Finally, future research should focus on investigating other potential reasons for neutral clinical studies of RIC, including co-morbidities and adjunctive therapies.
FUNDING

This work was supported by funding from the British Heart Foundation (grant number PG/15/52/31598) and the Biomedical Research Centre (grant number BRC233/CM/SD/101320). D.B. is a Medical Research Council Clinical Research Training Fellow. J.P. is a British Heart Foundation PhD Student. X.R. has received support from the Fundacion Alfonso Martin Escudero fellowship grant. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health’s National Institute for Health Research Biomedical Research Centres funding scheme.

The funders played no role in study design, collection, analysis, interpretation of data, writing of the report, or in the decision to submit the paper for publication. They accept no responsibility for the contents.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the authors of all of the studies included in this meta-analysis, with particular thanks to those who supplied information not included in their published manuscript.

CONFLICT OF INTEREST

None declared.
FIGURE LEGENDS

Figure 1. Flow chart of the study selection process
A systematic review yielded 539 reports. After removal of duplicates and the application of inclusion and exclusion criteria, 31 studies were included in the meta-analysis.

Figure 2. Forest plots of meta-analysis of conditioning efficacy in A) RIPreC and B) RIPerC/RIPostC
Forest plots of the effect of A) RIPreC and B) RIPerC/RIPostC on IS/AAR%, pooled using random-effects meta-analysis. 22 and 21, respectively, controlled comparisons were included, amounting to data from 280 control animals and 373 animals undergoing RIC.

Figure 3. Impact of experimental factors on the efficacy of A) RIPreC and B) RIPerC/RIPostC
WMD and the corresponding 95% CI for each variable were obtained by subgroup stratification. However, the reported P-value was obtained by meta-regression to reduce false-positive findings. Studies that used either mice or rats were grouped as ‘small animals’, and those using rabbits or pigs were grouped as ‘large animals’. A P value of <0.01 was considered significant.

Figure 4. Reporting of study quality indicators
Study quality was assessed using the ARRIVE guidelines on reporting in vivo experiments (A) and a 12-item quality score (B). Values are expressed as the percentage of studies reporting each quality indicator.

Figure 5. Assessment of publication bias
A funnel plot comparing treatment effect to a measure of study size (precision of the effect estimate). The vertical line represents the mean effect size. This plot was assessed visually, with further analysis of publication bias performed using Egger’s regression test.
REFERENCES


72. Sena ES, van der Worp HB, Bath PM, Howells DW, Macleod MR. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol* 2010;8:e1000344.


236 articles identified in Medline search

299 articles identified in EMBASE search

4 articles identified from expert consultation

370 articles underwent Level 1 screening

169 duplicates excluded

256 articles excluded following Level 1 screening

114 articles underwent Level 2 screening

83 articles excluded:
66 due to inclusion criteria not being met
12 as duplicate studies published in separate journals
2 due to unavailability of full text
3 due to missing data

31 studies included for meta-analysis
A. Remote preconditioning

<table>
<thead>
<tr>
<th>Study</th>
<th>WMD (95% CI)</th>
<th>%</th>
<th>Control – Sample size</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basalay 2012 (1)</td>
<td>22.00 (20.63, 23.37)</td>
<td>4.78</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Callier 2015</td>
<td>18.30 (15.83, 20.77)</td>
<td>4.75</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Hausenloy 2012 (1)</td>
<td>35.50 (32.10, 38.90)</td>
<td>4.68</td>
<td>2.5</td>
<td>5</td>
</tr>
<tr>
<td>Heinri 2011</td>
<td>22.00 (9.28, 34.72)</td>
<td>3.29</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>HIBart 2013</td>
<td>17.65 (11.78, 13.52)</td>
<td>4.83</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Kalaseh 2013 (1)</td>
<td>16.50 (13.50, 17.70)</td>
<td>4.82</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Krabtch 2013 (2)</td>
<td>17.80 (16.78, 18.82)</td>
<td>4.82</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Krabtch 2014</td>
<td>18.00 (16.92, 19.08)</td>
<td>4.82</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Kharbanda 2002</td>
<td>27.00 (24.23, 29.77)</td>
<td>4.73</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Lim 2010</td>
<td>35.10 (34.38, 35.82)</td>
<td>4.83</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Lu 2012 (1)</td>
<td>26.40 (18.56, 34.24)</td>
<td>4.10</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Lu 2012 (2)</td>
<td>3.60 (−5.41, 12.61)</td>
<td>3.91</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Masahiro 2012</td>
<td>18.00 (16.86, 19.14)</td>
<td>4.82</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Rosaf 2014</td>
<td>20.00 (15.62, 24.38)</td>
<td>4.58</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Shahid 2008</td>
<td>34.00 (32.47, 35.53)</td>
<td>4.80</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Weinbrunner 2002 (1)</td>
<td>10.00 (4.95, 15.07)</td>
<td>4.50</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Weinbrunner 2002 (2)</td>
<td>20.00 (15.06, 24.94)</td>
<td>4.51</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Weinbrunner 2002 (3)</td>
<td>25.00 (19.45, 30.55)</td>
<td>4.44</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Weinbrunner 2002 (4)</td>
<td>44.00 (39.05, 48.95)</td>
<td>4.51</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Wong 2012</td>
<td>23.23 (16.60, 29.86)</td>
<td>4.29</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Zhang 2006</td>
<td>30.70 (25.53, 35.87)</td>
<td>4.49</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Zhu 2013 (1)</td>
<td>20.21 (17.29, 23.13)</td>
<td>4.74</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Overall: $I^2$ = 99.7% p = 0.0006</td>
<td>22.82 (18.75, 26.89)</td>
<td>100.00</td>
<td>26.89</td>
<td>100.00</td>
</tr>
</tbody>
</table>

B. Remote perconditioning and postconditioning

<table>
<thead>
<tr>
<th>Study</th>
<th>WMD (95% CI)</th>
<th>%</th>
<th>Control – Sample size</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andreka 2007</td>
<td>25.40 (20.29, 30.51)</td>
<td>4.34</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Basalay 2012 (2)</td>
<td>29.00 (21.01, 24.97)</td>
<td>4.82</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Basalay 2012 (3)</td>
<td>29.00 (21.00, 35.00)</td>
<td>4.82</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Basalay 2012 (4)</td>
<td>20.00 (16.03, 24.07)</td>
<td>4.82</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Graff 2013</td>
<td>33.50 (30.35, 36.65)</td>
<td>4.60</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Hausenloy 2012 (2)</td>
<td>30.50 (27.04, 33.96)</td>
<td>4.72</td>
<td>2.5</td>
<td>6</td>
</tr>
<tr>
<td>Jeanneault 2012</td>
<td>30.20 (27.74, 32.66)</td>
<td>4.79</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kim 2014 (1)</td>
<td>25.50 (24.76, 36.44)</td>
<td>4.07</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Kim 2014 (2)</td>
<td>22.20 (20.87, 23.53)</td>
<td>4.85</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Li 2009</td>
<td>14.55 (13.03, 16.07)</td>
<td>4.85</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Sachintra 2014</td>
<td>6.50 (4.64, 8.36)</td>
<td>4.67</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Schertl 2006</td>
<td>22.00 (20.61, 23.39)</td>
<td>4.85</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Tamekou 2011</td>
<td>30.24 (27.56, 33.12)</td>
<td>4.76</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Wei 2011</td>
<td>14.67 (12.70, 16.64)</td>
<td>4.66</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Xin 2010</td>
<td>15.40 (12.82, 18.18)</td>
<td>4.40</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Xu 2015</td>
<td>12.00 (9.99, 14.01)</td>
<td>4.82</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Yu 2014</td>
<td>24.50 (20.03, 29.07)</td>
<td>4.87</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Zhang 2013</td>
<td>14.70 (11.98, 17.42)</td>
<td>4.50</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Zhu 2013 (2)</td>
<td>19.47 (16.68, 22.26)</td>
<td>4.76</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Zhu 2013 (3)</td>
<td>14.50 (11.86, 17.12)</td>
<td>4.73</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Overall: $I^2$ = 99.5% p = 0.0005</td>
<td>21.20 (17.07, 25.32)</td>
<td>100.00</td>
<td>25.32</td>
<td>100.00</td>
</tr>
</tbody>
</table>
A  Preconditioning Studies

WMD  95% CI  P-value  Adj R²
---  --------  -------  ------
21.95  (16.75, 26.24)  0.169  5.19%
31.14  (22.86, 39.52)  
22.45  (17.41, 27.48)  0.601  -4.18%
18.51  (6.42, 30.60)  
25.58  (16.87, 34.28)  
23.47  (17.19, 29.75)  0.742  -5.00%
23.13  (13.73, 32.52)  
21.63  (18.54, 24.71)  
22.03  (16.37, 27.69)  0.640  -4.27%
23.98  (18.00, 29.96)  
27.21  (21.89, 32.53)  0.042  17.36%
19.45  (16.05, 22.86)  

B  Per-/Post-conditioning Studies

WMD  95% CI  P-value  Adj R²
---  --------  -------  ------
20.65  (15.55, 25.74)  0.554  -3.81%
22.79  (17.97, 27.62)  
19.14  (12.86, 25.43)  0.254  2.08%
24.11  (14.81, 33.42)  
22.84  (20.75, 29.92)  
22.64  (19.07, 29.20)  0.347  -0.10%
19.76  (13.63, 25.88)  
19.63  (11.29, 27.96)  
24.09  (20.97, 27.21)  0.051  15.17%
18.16  (11.57, 24.75)  
20.19  (15.65, 24.52)  0.653  -4.07%
21.67  (16.16, 27.19)  

Figure 3
Figure 4

A. ARRIVE guidelines

Passed study quality criteria?

B. Study quality score

Passed study quality criteria?