ORIGINAL RESEARCH



# Effect of Study Design on the Reported Effect of Cardiac Resynchronization Therapy (CRT) on Quantitative Physiological Measures: Stratified Meta-Analysis in Narrow-QRS Heart Failure and Implications for Planning Future Studies

Richard J. Jabbour, MRCP; Matthew J. Shun-Shin, MRCP; Judith A. Finegold, MA, MRCP; S. M. Afzal Sohaib, MRCP; Christopher Cook, MRCP; Sukhjinder S. Nijjer, MRCP; Zachary I. Whinnett, PhD, MRCP; Charlotte H. Manisty, PhD, MRCP; Josep Brugada, MD; Darrel P. Francis, MA, MD, FRCP

***Background*—**Biventricular pacing (CRT) shows clear benefits in heart failure with wide QRS, but results in narrow QRS have appeared conflicting. We tested the hypothesis that study design might have influenced findings.

***Method and Results*—**We identified all reports of CRT-P/D therapy in subjects with narrow QRS reporting effects on continuous physiological variables. Twelve studies (2074 patients) met these criteria. Studies were stratified by presence of bias-resistance steps: the presence of a randomized control arm over a single arm, and blinded outcome measurement. Change in each endpoint was quantified using a standardized effect size (Cohen's d). We conducted separate meta-analyses for each variable in turn, stratified by trial quality. In non-randomized, non-blinded studies, the majority of variables (10 of 12, 83%) showed significant improvement, ranging from a standardized mean effect size of +1.57 (95%CI +0.43 to +2.7) for ejection fraction to +2.87 (+1.78 to +3.95) for NYHA class. In the randomized, non-blinded study, only 3 out of 6 variables (50%) showed improvement. For the randomized blinded studies, 0 out of 9 variables (0%) showed benefit, ranging from −0.04 (−0.31 to +0.22) for ejection fraction to −0.1 (−0.73 to +0.53) for 6-minute walk test.

***Conclusions*—**Differences in degrees of resistance to bias, rather than choice of endpoint, explain the variation between studies of CRT in narrow-QRS heart failure addressing physiological variables. When bias-resistance features are implemented, it becomes clear that these patients do not improve in *any* tested physiological variable. Guidance from studies without careful planning to resist bias may be far less useful than commonly perceived. **(*J Am Heart Assoc*. 2015;4:e000896 doi: 10.1161/JAHA.114.000896)**

**Key Words:** cardiac resynchronization therapy • heart failure • narrow QRS



Cardiac resynchronization therapy (CRT) undoubtedly provides both symptomatic and prognostic benefit in patients with heart failure and a wide QRS complex.[1–3] Whether it is effective in patients with narrow QRS complexes has appeared contentious. Studies addressing this have implemented bias-resistance steps (such as the inclusion of a randomized control arm and blinding of endpoint assessment) to varying degrees, and have addressed a variety of endpoints. While these studies have been reviewed in the past, no meta-analysis has focused on trial design as a potential explanatory variable for the differing results.[4]

We formally assessed the effect of CRT in patients with narrow QRS, to identify whether the conflict between different study results was an effect of trial design. To make it possible to compare effects on different endpoints, we calculated for each the standardized effect size (Cohen's d).

## Methods

### Eligibility and Search Strategy

We identified all reports of studies of heart failure patients with narrow QRS (<130 ms) and had either CRT pacemaker

*From the International Centre for Circulatory Health, National Heart and Lung Institute, Imperial College, London (R.J.J., M.J.S.-S., J.A.F., S.M.A.S., C.C., S.S.N., Z.I.W., C.H.M., D.P.F.); Hospital Clinic, University of Barcelona, Barcelona, Spain (J.B.).*

Correspondence to: Richard J. Jabbour, MRCP, International Centre for Circulatory Health, National Heart and Lung Institute, Imperial College London, 59-61 North Wharf Road, London W2 1LA, United Kingdom. E-mail: rjj03@ic.ac.uk

Received July 4, 2014; accepted October 24, 2014.

or CRT defibrillator implantation (CRT-P or CRT-D) inserted. MEDLINE (1946–September 2013), EMBASE (1974–September 2013), the Cochrane central register of controlled trials (Cochrane Library 2011, Issue 4), and www.controlled-trials.com (a meta-registry of randomized controlled clinical trials that includes the ISRCTN register) were searched using appropriate terms in the online appendix. Reference lists of the retrieved articles were hand-searched for additional publications. Conference presentations of the reported trials were used if they provided incremental information.

## Effect Sizes

The primary aim of this meta-analysis was to assess whether bias-resistance elements of study design affect study results. For each study we included all reported measured variables of functional status provided on a continuous scale and common left ventricular function measurements. For randomized controlled trials we defined the effect size as the difference between the change scores in each arm. For single arm studies we defined the effect size as the reported change score in the intervention arm. To allow measurements of different physiological quantities to be compared on a common scale we calculated for each the standardized effect size (Cohen's d) by dividing by the standard deviation of that variable in the patients before CRT implantation.

## Classification of Studies by Presence of Bias-Resistance Features

We stratified the studies into 3 broad groups depending on the number of bias-resistance features:

- 0 bias-resistance features (Neither an equivalent control group nor blinding of measurements);
- 1 bias-resistance feature (Randomization to a control group or intervention, but without blinding of patient and echocardiographic operator);
- 2 bias-resistance features (Randomization with blinding of patient and echocardiographic operator).

The bias-resistance features were only considered valid if the results presented used them. For example, if a study had a randomized control arm but only presented data from the intervention arm then we were obliged to consider the report to be of a single arm study.

We further assessed all studies using the Cochrane "risk of bias" tool to qualitatively identify if there were any additional sources of biases that could have affected the results.

## Data Abstraction

Data was abstracted in duplicate by 2 authors (RJ and CC). Disagreements were resolved by a third author (MJS).

## Data Analysis

We summarized the data and tested for inequality between the groups using a random-effects meta-analysis using the statistical environment "R" with the "metafor" package.[5] We stratified by trial quality along with end-point. Data were graphically presented using the package "ggplot2."[6]

# Results

## Search Results and Classification of Studies by Bias-Resistance of Design

Three hundred eighty-two articles met the initial search criteria, of which 131 were duplicates and 51 were excluded at the abstract stage. From 136 full-text articles screened, 12 studies (2074 patients) met the inclusion criteria (Figure 1).[7–18] Of these 12 studies, 9 enrolled patients with QRS durations <120 ms,[7–10,12,13,15,16,18] and 3 enrolled patients with QRS <130 ms.[11,14,17] The characteristics and classification of studies are presented in Table 1.

Across the 12 studies, various echocardiographic left ventricular functional and size parameters, and functional measured physiological variables were reported including: ejection fraction, end systolic volume, end diastolic volume, end systolic diameter, end diastolic diameter, sub-maximal exercise duration, quality of life score, NYHA class change, 6-minute walk distance, myocardial performance index, peak $VO_2$, and peak $VE/VO_2$ slope [Table 2].

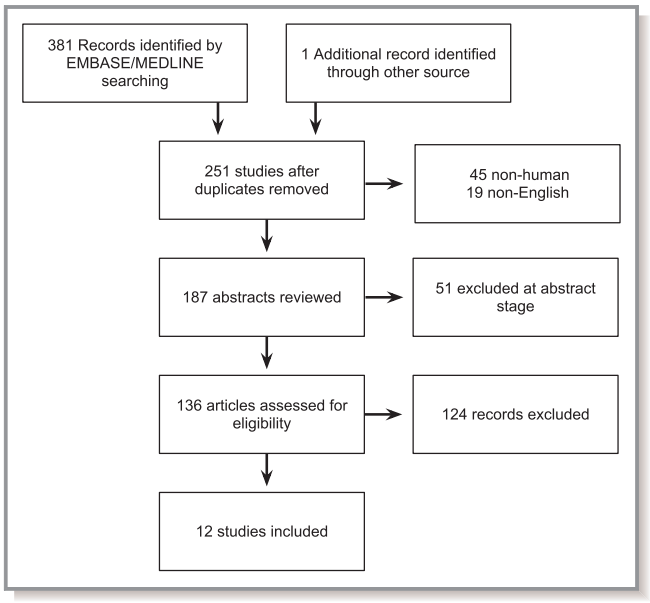


**Figure 1.** PRISMA flow diagram of studies

**Table 1.** Selected Baseline Characteristics of Studies Involving CRT and Narrow QRS

| Study | Blinded | Random | Arms | Inclusion Criteria | Exclusion Criteria | Enrollment Period | Intervention Group | Control/ Comparison Group | Outcomes Reported | Duration of Follow-up | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ECHO-CRT[11] | Yes | Yes | Two | NYHA III/IV<br>Dyssynchrony†<br>EF≤35%<br>QRS<130 ms | AF<br>Acute heart failure | August 2008 to March 2013 | CRT-D n=404 | D-ICD n=405 | Death<br>NYHA<br>QOL | 6 months | Double blind |
| LESSER-EARTH[12] | Yes | Yes | Two | ICD indication<br>EF≤35%<br>QRS<120 ms | Permanent AF<br>ACS<6 weeks<br>Previous CRT device | October 2003 to January 2011 | CRT-D n=44 | D-ICD n=41 | Sub Ex<br>6MWT<br>EF, LVESV<br>NYHA, QOL | 12 months | Double blind |
| RethinQ[17] | Yes | Yes | Two | NYHA III<br>QRS<130 ms<br>EF≤35% | Previous CRT device | August 2005 to January 2007 | CRT-D n=87 | D-ICD n=85 | QOL, 6MWT<br>Peak oxygen consumption | 6 months | Double blind |
| NARROW-CRT[10] | No | Yes | Two | EF≤35%<br>Dyssynchrony†<br>Max. medical therapy<br>QRS<120 ms | Conventional indication for pacing<br>Persistent AF | January 2008 to May 2010 | CRT-D n=60 | D-ICD n=60 | HF clinical composite score<br>HF hospitalization<br>HF death<br>LVEF<br>LVESV/DV | 12 months | No change in echo parameters stated in control arm |
| RESPOND[8] | No | Yes | Two | EF≤35%<br>NYHA III/IV<br>QRS<120 ms<br>Chronic HF<br>Max. medical therapy | Conventional indication for pacing<br>Recent MI | August 2007 to September 2009 | CRT-D n=29 | OPT n=31 | NYHA<br>LVEF<br>QOL<br>6MWT<br>LVESV/DV | 6 months | Maximum medical therapy control arm |
| Achilli et al[16] | No | No | Two | EF≤35%<br>Dyssynchrony†<br>Chronic HF<br>NYHA III/IV | Permanent AF<br>Valvular disease*<br>ACS<3 months<br>Severe COPD | February 2000 to March 2002 | CRT n=14 | CRT n=38<br>QRS >120 ms | NYHA<br>6MWT<br>LVEF, LVESD/DD | 45 months | Patients divided into two groups based on QRS duration |
| Bleeker et al[18] | No | No | Two | EF≤35%<br>Dyssynchrony†<br>Chronic HF<br>NYHA III/IV<br>QRS<120 ms | ACS<3 months<br>Decompensated HF | Insufficient information | CRT-D/P n=33 | CRT-D/P n=33<br>QRS >120 ms | NYHA<br>6MWT<br>QOL<br>LVEF<br>LVESV/DV | 6 months | Broad QRS comparison group |

Continued

DOI: 10.1161/JAHA.114.000896

**Table 1.** Continued

| Study | Blinded | Random | Arms | Inclusion Criteria | Exclusion Criteria | Enrollment Period | Intervention Group | Control/ Comparison Group | Outcomes Reported | Duration of Follow-up | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ESTEEM-CRT[7] | No | No | Single | EF≤35%<br>Dyssynchrony†<br>Chronic HF<br>NYHA III<br>QRS<120 ms | Persistent AF<br>Sustained VT<br>COPD<br>Bradycardia pacing | June 2005 to December 2007 | CRT-D n=68 | NA | NYHA, QOL<br>LVESV/DV<br>Peak V02<br>EF, LVESD/DD | 12 months | Single arm; Multi-centre |
| Gasparini et al[9] | No | No | Two | NYHA III/IV<br>Chronic HF<br>EF≤35%<br>QRS<120 ms | Insufficient information | October 1999 to April 2005 | CRT n=45 | CRT n=331<br>QRS >120 ms | NYHA<br>EF, LVESV<br>6MWT | 28 months (Range: 6–68 months) | Broad QRS control group |
| PROSPECT[14] | No | No | Single | EF≤35%<br>Dyssynchrony†<br>NYHA III/IV<br>QRS<130 ms | Insufficient information | March 2004 to November 2005 | CRT-P/D n=41 | NA | NYHA, QOL<br>6MWT, CCS<br>LVEF, LVESD/SV<br>MPI | 6 months | Single arm sub-study; multi-centre |
| van Bommel[13] | No | No | Single | EF≤35%<br>Dyssynchrony†<br>Chronic HF<br>QRS<120 ms | Insufficient information | Insufficient information | CRT-P/D n=123 | NA | LVEF<br>LVESV/DV<br>NYHA | 6 months | Single arm; multi-centre |
| Yu et al[15] | No | No | Two | NYHA III/IV<br>Dyssynchrony†<br>QRS<120 ms | Insufficient information | Insufficient information | CRT n=51 | CRT n=51<br>QRS >120 ms | NYHA, QOL<br>6MWT, Ex<br>Mitral regurgitation<br>LVEF, MPI<br>LVESV/DV | 3 months | Broad QRS comparison group |

AF indicates atrial fibrillation; CCS, clinical composite score; D-ICD, dual chamber ICD; EDD, left ventricular end diastolic diameter; EDV, left ventricular end diastolic volume; EF, ejection fraction; ESD, left ventricular end systolic diameter; ESV, left ventricular end systolic volume; Ex, exercise capacity, metabolic equivalent; MPI, myocardial performance index; OPT, optimal medical therapy; QOL, quality of life questionnaire; Sub Ex, exercise duration at submaximal level.
*Surgically correctable significant valvular disease.
†Echocardiographic evidence of interventricular/intraventricular asynchrony.

DOI: 10.1161/JAHA.114.000896

**Table 2.** Continuous Variables Analysed in Meta-Analysis

| | Outcome Analyzed | |
|---|---|---|
| | Clinical | Echocardiographic |
| ECHO-CRT[11] | QOL | |
| LESSER-EARTH[12] | 6MWT, Sub Ex | EF, ESV |
| RethinQ[17] | 6MWT, Peak $VO_2$, QOL | EF, EDD, ESD, EDV, ESV |
| NARROW-CRT[10] | | EF, EDD, ESD, EDV, ESV |
| RESPOND[8] | 6MWT, NYHA, QOL | EF, EDV, ESV |
| Achilli et al[16] | 6MWT, NYHA | EF, EDD, ESD |
| Bleeker et al[18] | 6MWT, NYHA, QOL | EF, EDV, ESV |
| ESTEEM-CRT[7] | Peak $VO_2$, $VE/VCO_2$, QOL | EF, EDD, ESD, EDV, ESV |
| Gasparini et al[9] | 6MWT | EF, ESV |
| PROSPECT[14] | 6MWT, NYHA, QOL | EF, EDD, ESD, EDV, ESV, MPI |
| van Bommel[13] | | EF, EDV, ESV |
| Yu et al[15] | 6MWT, NYHA, QOL, Ex | EF, EDD, ESD, EDV, ESV, MPI |

EDD indicates left ventricular end diastolic diameter; EDV, left ventricular end diastolic volume; EF, ejection fraction; ESD, left ventricular end systolic diameter; ESV, left ventricular end systolic volume; Ex, exercise capacity, metabolic equivalent; MPI, myocardial performance index; NYHA, NYHA class change; QOL, quality of life questionnaire; Sub Ex, exercise duration at submaximal level; $VE/VCO_2$, $VE/VCO_2$ slope.

## Classification of Studies by Bias-Resistance Features

Eight studies, including 435 patients, had neither a randomized controlled arm nor blinding.[7,9,10,13–16,18] Three of these were single armed studies.[7,13,14] In four of these studies comparison data was presented, but arose from patients with a QRS above the threshold and so were analysed as single armed.[9,15,16,18] One of these studies was carried out as a randomized trial, but only presented continuous variable data from the intervention arm, and hence was analysed as a single armed study.[10]

One study had a randomized control arm but neither patients nor echocardiographers were blinded to whether CRT was active, enrolling 60 patients.[8]

Three studies had both a randomized control arm and blinding of patients and sonographers totaling 1066 patients. In these trials the patients received a CRT-D device, with the control patients having the CRT function inactive.[11,12,17]

The baseline characteristics and effects of CRT in the intervention arm are available in Table 3 and the Cochrane risk of bias assessment tool in Table 4.

## Is it Bias-Resistance of Study Design or Choice of End-Point That Leads to Unintentionally False-Positive Results?

A series of meta-analyses, one for each end-point, stratified by the presence of bias-resistance features, is shown in Figure 2.

The green diamonds show the meta-analysis summary results of the randomized, blinded studies. None (0/9, 0%) of these showed a significant effect of CRT on its end-point. However, the trials with fewer bias-resistance features showed a different pattern.

The orange diamonds show the results of the randomized, unblinded study. Half (3/6, 50%) of the end-points showed statistically significant improvement.

The red diamonds show the meta-analysis summary results of the studies with neither a randomized controlled arm nor blinding. Most (10/12, 83%) end-points showed statistically significant improvement.

## Danger of Viewing Multiple Positive End-Points in a Trial as Strong Evidence

Commentators sometimes remark on the multiplicity of positive end-points within a single arm study as though their great number might somehow overcome the weakness of the study in lacking blinding, or even lacking a control arm. This is not wise because this counts the same patients on multiple occasions as though they were separate.[19]

Figure 3 illustrates this danger of viewing multiple positive endpoints as strong evidence. It shows 1 point for each endpoint reported in each of the studies. The number of patients (and bias-resistance of the design) is given by the horizontal position, and the apparent standardized effect size is given by the vertical position. The left-hand group of studies, with only a single arm, and therefore the least bias-resistance, give the

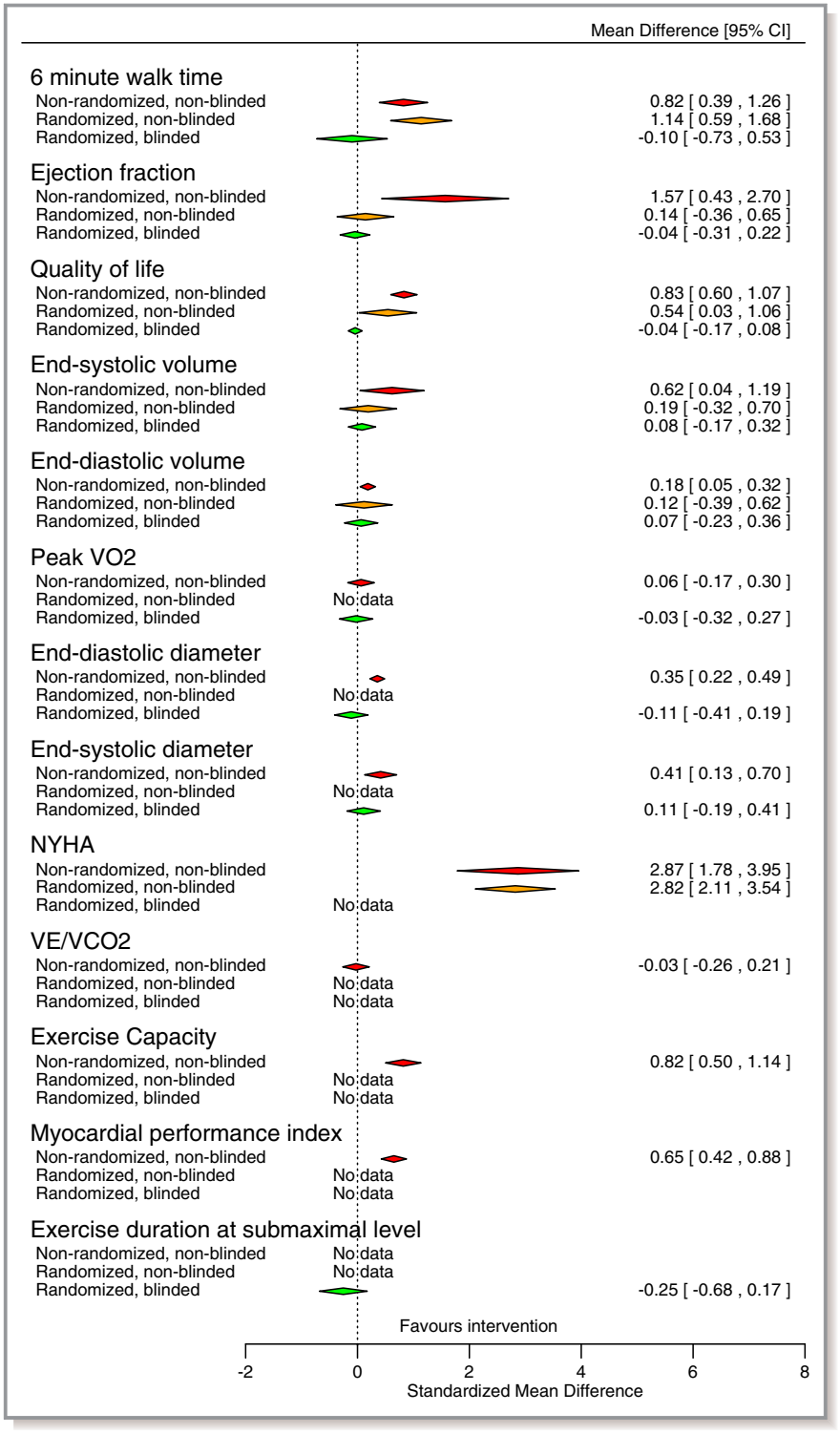


**Figure 2.** Meta-analyses of effects on physiological variables, with studies stratified by number of bias-resistance features in the study design. For each variable a meta-analysis was conducted stratified by the presence of bias-resistance features. The majority (10 out of 12) of variables reported in studies without randomization and blinding (red symbols) favored intervention to a statistically significant degree. In contrast, all 9 of the outcome variables reported by randomized, blinded trials (green symbols) were neutral. The 6 variables reported by studies with randomization but not blinding (orange symbols) were equally divided between suggesting significant response to intervention (3 variables) and not (3 variables).

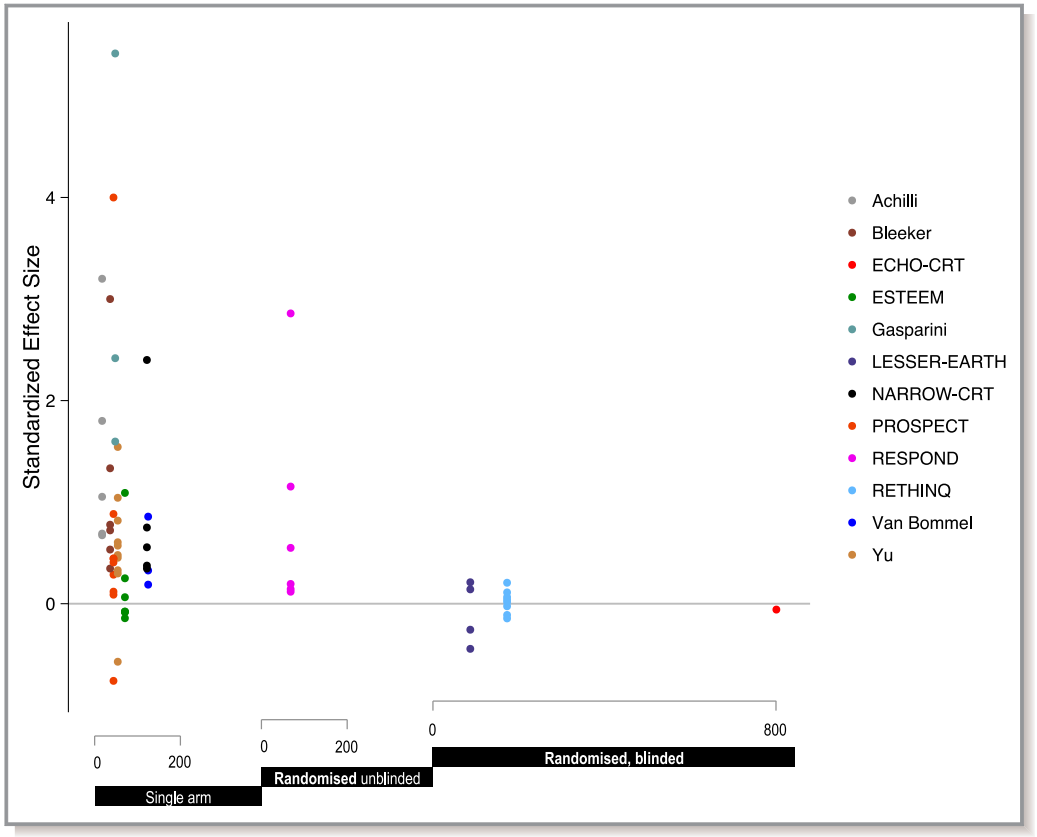


**Figure 3.** Effect on physiological-variable endpoints of bias-resistance features and sample size. It is not statistically valid to "merge" multiple endpoints from the same study as though they are independent. Unfortunately, however, it is common practice when commenting on a study to highlight that multiple endpoints are showing a consistent indication. This plot illustrates why such presentation is invalid. Each study is represented by a series of points, one for each reported endpoint. The horizontal position represents the bias-resistance group and the sample size (and is therefore is common for all end-points for a single study). As can be seen, the less bias resistant the design, the greater the tendency to a positive result.

impression of efficacy. The right-hand group of studies, which have randomization and blinding, do not.

The key factor is the bias-resistance of the study design, rather than the choice or number of endpoints.

## Discussion

The conflict between study reports on the efficacy of biventricular pacing (CRT) on physiological variables in heart failure with narrow QRS, seems to originate not in the choice of physiological endpoint, but in the design of the study. There may have been unintentional bias introduced when study design did not possess bias-resistance features. Measurements made for routine clinical purposes do not have the correct properties for drawing reliable scientific inferences. Readers may not realize that such data are not equivalent to data from a scientific experiment carefully designed to answer a question reliably.

In the 3 studies implementing both a randomized control arm and blinding, the effect of CRT on endpoints is neutral on these physiological variables. Two of these trials stopped early due to futility because of detrimental signals in event rates.[11,12]

### Rationale for CRT Implantation in Relation to QRS Complex Width

Broad QRS complexes were the defining characteristic of the early patients receiving CRT from the very first case reports through to the pivotal physiological studies and landmark trials.[2,20–23] The powerful symptomatic and morbidity/mortality reduction were a strong stimulus for attempting expansion into patients with narrow QRS.[3,24]

One rationale for such expansion has been the umbrella concept of dyssynchrony. It was conjectured that CRT might, even in the absence of electrical dyssynchrony (wide QRS), alleviate isolated mechanical (echocardiographic) dyssynchrony. More recently however, it has emerged that the apparent predictive power of mechanical (echocardiographic) dyssynchrony for benefit from CRT exists only when studies

**Table 3.** Baseline Characteristics and Effects of CRT in Narrow QRS Arm of Each Study

| Baseline Characteristics | ECHO-CRT[11] | LESSER-EARTH[12] | RethinQ[17] | NARROW-CRT[10] | RESPOND[8] | Achilli et al[16] | Bleeker et al[18] | ESTEEM-CRT[7] | Gasparini et al[9] | PROSPECT[14] | van Bommel[13] | Yu et al[15] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Double Blind | | | Unblinded | | | | | | | | |
| | Two Arms | | | | | Single Arm | | | | | | |
| Randomized | Yes | Yes | Yes | Yes | Yes | No | No | No | No | No | No | No |
| Patients (n) | 404 | 44 | 87 | 60 | 29 | 14 | 33 | 68 | 45 | 41 | 123 | 51 |
| Age (years) | 58±13 | 62±10 | 60±12 | 65±9 | 67±8 | 68.3±8 | 63±11 | 58±14 | 68±9 | 64±13 | 61±11 | 63±11 |
| Male (%) | 73 | 64 | 71 | 88 | 86 | 71 | 85 | 68 | 84.4 | 71 | 79 | 78 |
| Ischaemics (%) | 54 | 73 | 54 | NA | 76 | 29 | 70 | 60 | NA | 49 | 61 | 49 |
| Ejection Fraction (%) | 27±6 | 25±6 | 25±5 | 28±5 | 22±8 | 25±5 | 22±6 | 25±7 | 29±4 | 25±6 | 27±7 | 28±7 |
| Intervention | CRT | CRT | CRT | CRT | CRT | CRT | CRT | CRT | CRT | CRT | CRT | CRT |
| Control | D-ICD | D-ICD | D-ICD | D-ICD | OMT | NA | CRT** | NA | NA | NA | NA | NA |
| QRS (ms) | 105±13 | 105±10 | 107±12 | 107±14 | 92±11 | 110±11 | 110±8 | 102±10 | 109±9 | 108±14 | 106 (98–114) | 103±13 |
| Effects of CRT in Narrow QRS arm | | | | | | | | | | | | |
| Reduction in NYHA class | NA | NA | NA | NA | 1.1±1.5 | 1.6±1.8 | 0.9±1.4 | NA | NA | 0.8±1.2 | NA | 0.7±1.5 |
| Reduction in QOL score | 12±21 | NA | 8 (10 to 1)‡ | NA | 12.2±25 | NA | 14±22 | 24±21* | NA | 17±26 | NA | 8±28 |
| Improvement in 6MWT (m) | NA | –11.3 (–31.7to 9.7)† | 26 (0 to 46)‡ | NA | 103±186 | 94±136 | 96±152 | NA | 182±286 | 48±97 | NA | 46±122 |
| Reduction in LVEDV (ml) | NA | NA | 16 (29 to 8)‡ | 18 (47 to 6)† | 15.4 (44.6)§ | NA | 27±43 | 1±35* | NA | 8±64 | 12±40 | 14±29 |
| Reduction in LVESV (ml) | NA | –6.4 (–18.8to 5.9)† | 19 (34 to 12)‡ | 30 (48 to 11)† | 26.2 (66.3)§ | NA | 40±64 | 1±26* | 71.8±113 | 9±37 | 17±56 | 19±39 |
| Reduction in LVEDD (mm) | NA | NA | 0 (2 to 0)‡ | 3 (5 to 0)† | NA | 6.2±12 | NA | 2±5* | NA | 4±9 | NA | 0.3±1 |
| Reduction in LVESD (mm) | NA | NA | 1 (3 to 0)‡ | 5 (7 to 2)† | NA | 5.8±11 | NA | 0±10* | NA | 4±8 | NA | 0.5±1 |
| Improvement in LVEF (%) | NA | 3.3 (0.7 to 6)† | 1.2 (0.4 to 4.4)‡ | 12 (10 to 13)† | 6.7 (18)§ | 9±10 | 8±13 | 0±7* | 23.3±37 | 2±9 | 6±20 | 7±15 |

OMT indicates optimal medical therapy; D-ICD, dual chamber ICD.
*Data shown from European Society of Cardiology Conference presentation (2008).
†Mean (95% confidence interval).
‡Median (95% confidence interval).
§Median (interquartile range).
**Effectively non-comparable control group as broad QRS.

DOI: 10.1161/JAHA.114.000896

**Table 4.** Cochrane Risk of Bias Assessment Tool

| Study | Adequate Sequence Generation | Allocation Concealment | Blinding | Incomplete Outcome Data | Free of Selective Reporting |
|---|---|---|---|---|---|
| ECHO CRT[11] | Computer generated randomization | Yes | Double Blind | Intentional to treat analysis; low discontinuation | Yes |
| LESSER-EARTH[12] | Randomized, no further details | Yes | Double Blind | Intentional to treat analysis; low discontinuation | Yes |
| RethinQ[17] | Computer generated randomization | Yes | Double Blind | Intentional to treat analysis; low discontinuation | Yes |
| NARROW-CRT[10] | Block randomization; Delivered by sealed envelope technique | Insufficient detail | Patients blinded | Insufficient details | Only echocardiographic parameters from CRT arm presented |
| RESPOND[8] | Computer generated randomization; Delivered by sealed envelope technique | Insufficient detail | Unblinded | Intentional to treat analysis; low discontinuation | Yes |
| Achilli et al[16] | Effectively single arm with non-comparable wide QRS group | Open label | Unblinded | Insufficient details | Yes |
| Bleeker et al[18] | Effectively single arm with non-comparable control group | Open label | Unblinded | Insufficient details | Yes |
| ESTEEM-CRT[7] | Single Arm | Not applicable | Unblinded | Insufficient details | Yes |
| PROSPECT[14] | Single Arm | Not applicable | Unblinded | Insufficient details | Yes |
| Gasparini et al[9] | Effectively single arm with non-comparable wide QRS group | Open label | Unblinded | Insufficient details | Yes |
| van Bommel[13] | Single Arm | Not applicable | Unblinded | Insufficient details | Yes |
| Yu et al[15] | Effectively single arm with non-comparable wide QRS group | Open label | Unblinded | Insufficient details | Yes |



do not have formal enrollment and blinding of measurements.[25]

Our findings are concordant with a recent large meta-analysis of RCTs of CRT, which showed through a spline-based regression analysis a direct relationship between QRS width and prognostic benefit from CRT, with no statistically significant benefit from CRT once the QRS duration falls below approximately 130 ms.[1]

## Implications for Research

Highlighting the importance of bias-resistance steps such as randomization and blinding is not novel, having been introduced in 1948 with Hill's randomized trial of streptomycin in tuberculosis.[26] The published reports of observational, and incompletely blinded, studies of therapy produce effect estimates that tend to show exaggerated benefits, and can even be in the opposite direction to thoroughly blinded, randomized controlled trials.[27] Nevertheless, uncontrolled, non-randomized, and unblinded routine clinical data are widely available in every hospital and it is inevitable that such data will enter the literature. Selective enthusiasm to report (and review favorably, and publish) positive data, together called "publication bias," can further distort the literature towards positivity, unhelpfully.

## Study Limitation?

The lesson to learn may be that lack of a suitable control arm with randomization and blinding in a study of treatment outcomes reporting measurements acquired through routine clinical practice should not be considered merely a minor "study limitation." The potential for this bias is so large that a published study in this class cannot be trusted to give even approximate guidance.

Rarely are readers explicitly warned that the measurement (and its associated confidence interval) from non-bias-resistant studies can be so misleading as to get the direction of effect completely backwards. Table 5 lists, for each study lacking bias resistance steps, where the study remarks on the potential for the effect size to be biased and what it says on the subject.

**Table 5.** Analysis of Single Arm Trials and Acknowledgement of Limitations Associated With Them

| | Abstract | Methods | Results | Conclusion |
|---|---|---|---|---|
| Achilli et al[16] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| Bleeker et al[18] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| ESTEEM-CRT[7] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| Gasparini et al[9] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| PROSPECT[14] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| van Bommel[13] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |
| Yu et al[15] | Not mentioned | Not mentioned | Not mentioned | Not mentioned |

The weakness of lack of bias-resistance is far worse than the weakness of small size of RCTs. Small RCTs suffer from noise, which is an equal tendency to over- or under-estimate effect sizes. As sample size grows, or as RCTs are synthesized by meta-analysis, the nature of noise is to progressively subside. Bias, on other hand, is in a consistent direction. Thus a non-bias-resistant study cannot safely be used to give a "rough idea" of what a bias-resistant study would find. The roughness of the idea can dominate any genuine effect, and is easily underestimated by authors and reader alike. Moreover, the larger the biased study, the more likely its confidence interval does *not* contain the true value.

Meta-analysis of the early observational studies used in our present analysis, when taken in aggregate, point us towards physiological benefit.[4] When the question is restudied with bias removed, we see physiological neutrality alongside event-based evidence of harm.[11,12] This phenomenon has also been seen in observational study designs for balloon pump therapy, which have consistently generated a wide range of results which, if read superficially, may be misleading.[28]

Nevertheless it should not be forgotten that, while the impact of such noise shrinks with sample size, the effect of bias does not. In an irony underappreciated by many of us, bias-vulnerable studies are more likely to be falsely statistically significant if large than small.[29]

## Impact on Publication and Interpretation of Unblinded Data

This analysis is not a criticism of the conduct of the studies listed. One had the primary endpoint of mortality and therefore did not require blinding to measure their primary endpoint without bias. Instead, our study puts a spotlight on what might be the consequences of drawing inferences from data that inevitably become available from studies that are unblinded or do not have a control arm. The problem is not unique to any study's team, but is common to all of us, perhaps as the result of our inevitable conditioning by normal clinical practice, where it may be good practice to portray to patients a favorable picture of their response to intervention, in order to maximize the overall symptomatic improvement in that individual (only part of which is directly mediated by the device).

Our challenge is to build a community understanding that when addressing mechanistic questions we should not rely on unblinded clinical data, originally obtained for individual-patient clinical purposes, to be a suitable bias-resistant basis for correct evaluation of physiological benefit. In our analysis, what the non-bias-resistant clinical data gave was not a feeble and uncertain pointer, but multiple clear, consistent, and statistically significant pointers to benefit, but when tested in a bias-resistant manner these turned out to be wrong.

If the magnitude of unintentional bias can be so large, we should think carefully before reviewing such data as evidence in a scientific forum.

## Hypothesis Generating?

Uncontrolled, unblinded, non-randomized data are widely available and often examined for features that might suggest an RCT of therapeutic intervention. However, although the sophisticated authors of such an article might fully understand that an observational study is not a recommendation on which to base therapeutic decisions but a highlighting of an interesting area to trial, many readers cannot resist making this intellectual jump.

The particular hypothesis that CRT would be "beneficial in heart failure with narrow QRS" could have been generated directly from the hypotheses in the positive trials for wide QRS, by simply changing one word. Such a method of generation would have produced the hypothesis:

1. in less time;
2. at less cost;
3. without obscuring the fact that it was a hypothesis, and
4. preserving the clear understanding (from broad-QRS trials) of how to test it.

It is a fallacy that unblinded, non-randomized studies are a useful step before blinded RCTs for assessing treatment efficacy. In the case of CRT:

1. In broad-QRS heart failure, the data that was not designed to resist bias suggested positive benefits. The subsequent blinded, randomized trials indicated *benefit*.
2. In narrow-QRS heart failure, the data that was not designed to resist bias suggested positive benefits. The subsequent blinded, randomized trials indicated *harm*.

Thus, the appearance of positive benefits in studies not designed to resist bias, in 1 group predicted genuine benefit in survival, and in the other group genuine harm to survival.

## Clinical Implications

When practicing clinical medicine we may find it difficult to resist looking at unblinded non-randomized studies especially when they are numerous or large or both. Nevertheless, we should remember the many instances in which they have been seriously misleading.[28, 30-32] We should focus on studies that incorporate vigorous steps to avoid bias, where such studies exist. Where there are no such studies, we should identify this and focus our energies on designing and implementing trials that have these characteristics. They need not be expensive, if we do not load them with compulsory features beyond inexpensive online randomization and a simple incontrovertible endpoint such as all-cause mortality. If the endpoint is a quantitative physiological marker, then expenditure on measuring this without bias is not a waste of resources, but a necessity for preventing the entire trial being a waste of resources.

## Study Limitations

Our search strategy might have missed some studies, although we attempted to be comprehensive. Secondly, we focused on physiological variables rather than event endpoints.[7,10,11,14] This was needed to allow examination of all 3 classes of study, since uncontrolled studies are unable to state effects on event rates. Thirdly, some data from the control arm of 1 randomized trial was missing and could not be obtained from the corresponding author, and therefore this trial had to be interpreted as a single arm study.[10] Finally, we could only use the variables provided by the authors. This is valid as long as the authors did not selectively present variables that showed an improvement; however, this susceptibility to bias exists for any reader of such reports. In addition, if this is an explanation for their discrepancy from the randomized trials then it further underlines the importance of preferentially using the prospectively specified endpoints from blinded randomized controlled trials as guidance for clinical decision-making.

## Conclusion

Patients with heart failure and narrow QRS complexes appear to show a physiological improvement with CRT, but only in unblinded studies or those without a randomized control arm. When blinding and randomization are implemented the CRT effect on physiological markers is neutral.

This experience of CRT in narrow-QRS heart failure is a particularly clear illustration of the need throughout cardiology to take elaborate steps to prevent inadvertent bias. The standard method of resisting bias is randomization and blinding. This requires planning and may appear to increase cost. However, in retrospect it might have been preferable to try harder at the outset to acquire bias-resistant data.

## Sources of Funding

Nijjer is supported by the MRC (G1100443); Francis (FS/10/038), Finegold (FS/14/25/30676), Sohaib (FS/11/92/29122), Shun-Shin (FS/14/27/30752) and Whinnett (FS/13/44/30291) are supported by the British Heart Foundation.

## Acknowledgement

The authors are grateful for infrastructural support from the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London.

## Disclosures

None.

## References

1. Cleland JG, Abraham WT, Linde C, Gold MR, Young JB, Claude DJ, Sherfesee L, Wells GA, Tang AS. An individual patient meta-analysis of five randomized trials assessing the effects of cardiac resynchronization therapy on morbidity and mortality in patients with symptomatic heart failure. *Eur Heart J*. 2013;34:3547–3556.
2. Cleland JG, Daubert JC, Erdmann E, Freemantle N, Gras D, Kappenberger L, Tavazzi L. The effect of cardiac resynchronization on morbidity and mortality in heart failure. *N Engl J Med*. 2005;352:1539–1549.
3. Finegold JA, Raphael C, Levy WC, Whinnett Z, Francis DP. Quantification of survival gain from cardiac resynchronization therapy: non-linear growth with time, and greater gain in low-risk patients, make raw trial data an underestimate of real-world behaviour. *J Am Coll Cardiol*. 2013;62:2406–2413.
4. Jeevanantham V, Zareba W, Navaneethan S, Fitzgerald D, Yu CM, Achilli A, Bax J, Daubert J. Metaanalysis on effects of cardiac resynchronization therapy in heart failure patients with narrow QRS complex. *Cardiol J*. 2008;15:230–236.

5. R Development Core Team. R: a language environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010. [computer program]. Available at: http://www.R-project.org. Accessed January 2014.
6. ggplot2: elegant graphics for data analysis. New York: Springer; 2009. [computer program].
7. Donahue T, Niazi I, Leon A, Stucky M, Herrmann K. Acute and chronic response to CRT in narrow QRS patients. *J Cardiovasc Transl Res*. 2012;5:232–241.
8. Foley PWX, Patel K, Irwin N, Sanderson JE, Frenneaux MP, Smith REA, Stegemann B, Leyva F. Cardiac resynchronisation therapy in patients with heart failure and a normal QRS duration: the RESPOND study. *Heart*. 2011;97:1041–1047.
9. Gasparini M, Regoli F, Galimberti P, Ceriotti C, Bonadies M, Mangiavacchi M, Andreuzzi B, Bragato R, Pini D, Klersy C, Gronda E. Three years of cardiac resynchronization therapy: Could superior benefits be obtained in patients with heart failure and narrow QRS? (vol 30, pg S34, 2007). *Pacing Clin Electrophysiol*. 2007;30:1425.
10. Muto C, Solimene F, Gallo P, Nastasi M, La RC, Calvanese R, Iengo R, Canciello M, Sangiuolo R, Diemberger I, Ciardiello C, Tuccillo B. A randomized study of cardiac resynchronization therapy defibrillator versus dual-chamber implantable cardioverter-defibrillator in ischemic cardiomyopathy with narrow QRS: the NARROW-CRT study. *Circ Arrhythm Electrophysiol*. 2013;6:538–545.
11. Ruschitzka F, Abraham WT, Singh JP, Bax JJ, Borer JS, Brugada J, Dickstein K, Ford I, Gorcsan J III, Gras D, Krum H, Sogaard P, Holzmeister J. Cardiac-Resynchronization Therapy in Heart Failure with a Narrow QRS Complex. *N Engl J Med*. 2013;369:1395–1405.
12. Thibault B, Harel F, Ducharme A, White M, Ellenbogen KA, Frasure-Smith N, Roy D, Philippon F, Dorian P, Talajic M, Dubuc M, Guerra PG, Macle L, Rivard L, Andrade J, Khairy P. Cardiac Resynchronization Therapy in Patients With Heart Failure and a QRS Complex < 120 Milliseconds The Evaluation of Resynchronization Therapy for Heart Failure (LESSER-EARTH) Trial. *Circulation*. 2013;127:873–881.
13. van Bommel RJ, Tanaka H, Delgado V, Bertini M, Borleffs CJW, Marsan NA, Holzmeister J, Ruschitzka F, Schalij MJ, Bax JJ, Gorcsan J. Association of intraventricular mechanical dyssynchrony with response to cardiac resynchronization therapy in heart failure patients with a narrow QRS complex. *Eur Heart J*. 2010;31:3054–3062.
14. van Bommel RJ, Gorcsan J, Chung ES, Abraham WT, Gjestvang FT, Leclercq C, Monaghan MJ, Nihoyannopoulos P, Peraldo C, Yu CM, Demas M, Gerritse B, Bax JJ. Effects of cardiac resynchronisation therapy in patients with heart failure having a narrow QRS Complex enrolled in PROSPECT. *Heart*. 2010;96:1107–1113.
15. Yu CM, Chan YS, Zhang Q, Yip GW, Chan CK, Kum LC, Wu L, Lee AP, Lam YY, Fung JW. Benefits of cardiac resynchronization therapy for heart failure patients with narrow QRS complexes and coexisting systolic asynchrony by echocardiography. *J Am Coll Cardiol*. 2006;48:2251–2257.
16. Achilli A, Sassara M, Ficili S, Pontillo D, Achilli P, Alessi C, De SS, Guerra R, Patruno N, Serra F. Long-term effectiveness of cardiac resynchronization therapy in patients with refractory heart failure and "narrow" QRS. *J Am Coll Cardiol*. 2003;42:2117–2124.
17. Beshai JF, Grimm RA, Nagueh SF, Baker JH, Beau SL, Greenberg SM, Pires LA, Tchou PJ. Cardiac-resynchronization therapy in heart failure with narrow QRS complexes. *N Engl J Med*. 2007;357:2461–2471.
18. Bleeker GB, Holman ER, Steendijk P, Boersma E, van der Wall EE, Schalij MJ, Bax JJ. Cardiac resynchronization therapy in patients with a narrow QRS complex. *J Am Coll Cardiol*. 2006;48:2243–2250.
19. Tanner-Smith EE, Tipton E. Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Res Synth Methods*. 2014;5:13–30.
20. Cazeau S, Leclercq C, Lavergne T, Walker S, Varma C, Linde C, Garrigue S, Kappenberger L, Haywood GA, Santini M, Bailleul C, Daubert JC. Effects of multisite biventricular pacing in patients with heart failure and intraventricular conduction delay. *N Engl J Med*. 2001;344:873–880.
21. Bristow MR, Saxon LA, Boehmer J, Krueger S, Kass DA, De MT, Carson P, DiCarlo L, DeMets D, White BG, DeVries DW, Feldman AM. Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N Engl J Med*. 2004;350:2140–2150.
22. Auricchio A, Stellbrink C, Block M, Sack S, Vogt J, Bakker P, Klein H, Kramer A, Ding J, Salo R, Tockman B, Pochet T, Spinelli J. Effect of pacing chamber and atrioventricular delay on acute systolic function of paced patients with congestive heart failure. The Pacing Therapies for Congestive Heart Failure Study Group. The Guidant Congestive Heart Failure Research Group. *Circulation*. 1999;99:2993–3001.
23. Leclercq C, Cazeau S, Ritter P, Alonso C, Gras D, Mabo P, Lazarus A, Daubert JC. A pilot experience with permanent biventricular pacing to treat advanced heart failure. *Am Heart J*. 2000;140:862–870.
24. Sohaib SMA, Chen ZC, Whinnett Z, Francis DP, Manisty C. Systematic review of genuine symptomatic response to cardiac resynchronization therapy: acknowledging the contribution of spontaneous response. *European Heart Journal* 2012; 33:995 (Abstract).
25. Nijjer SS, Pabari PA, Stegemann B, Palmieri V, Leyva F, Linde C, Freemantle N, Davies JE, Hughes AD, Francis DP. The limit of plausibility for predictors of response: application to biventricular pacing. *JACC Cardiovasc Imaging*. 2012;5:1046–1065.
26. Daniels M, Hill AB. Chemotherapy of pulmonary tuberculosis in young adults; an analysis of the combined results of three Medical Research Council trials. *Br Med J*. 1952;1:1162–1168.
27. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005;294:218–228.
28. Sjauw KD, Engstrom AE, Vis MM, van der Schaaf RJ, Baan J Jr, Koch KT, de Winter RJ, Piek JJ, Tijssen JG, Henriques JP. A systematic review and meta-analysis of intra-aortic balloon pump therapy in ST-elevation myocardial infarction: should we change the guidelines? *Eur Heart J*. 2009;30:459–468.
29. Shun-Shin MJ, Francis DP. Why even more clinical research studies may be false: effect of asymmetrical handling of clinically unexpected values. *PLoS One*. 2013;8:e65323.
30. Rossebo AB, Pedersen TR, Boman K, Brudi P, Chambers JB, Egstrup K, Gerdts E, Gohlke-Barwolf C, Holme I, Kesaniemi YA, Malbecq W, Nienaber CA, Ray S, Skjaerpe T, Wachtell K, Willenheimer R. Intensive lipid lowering with simvastatin and ezetimibe in aortic stenosis. *N Engl J Med*. 2008;359:1343–1356.
31. Rosenhek R, Rader F, Loho N, Gabriel H, Heger M, Klaar U, Schemper M, Binder T, Maurer G, Baumgartner H. Statins but not angiotensin-converting enzyme inhibitors delay progression of aortic stenosis. *Circulation*. 2004;110:1291–1295.
32. Rich JD, Cannon CP, Murphy SA, Qin J, Giugliano RP, Braunwald E. Prior aspirin use and outcomes in acute coronary syndromes. *J Am Coll Cardiol*. 2010;56:1376–1385.

# Effect of Study Design on the Reported Effect of Cardiac Resynchronization Therapy (CRT) on Quantitative Physiological Measures: Stratified Meta –Analysis in Narrow–QRS Heart Failure and Implications for Planning Future Studies

Richard J. Jabbour, Matthew J. Shun-Shin, Judith A. Finegold, S. M. Afzal Sohaib, Christopher Cook, Sukhjinder S. Nijjer, Zachary I. Whinnett, Charlotte H. Manisty, Josep Brugada and Darrel P. Francis

*J Am Heart Assoc.* 2015;4:e000896; originally published January 6, 2015;
doi: 10.1161/JAHA.114.000896
The *Journal of the American Heart Association* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Online ISSN: 2047-9980