

GUEST EDITOR'S PAGE



The Nonischemic Cardiomyopathy Defibrillator Conundrum



Is a Meta-Analysis Enough?

Jagmeet P. Singh, MD, PhD,^{a,b} Jeanne E. Poole, MD,^c Paul Kolm, PhD^d

*It ain't what you don't know that gets you
into trouble. It's what you know for sure
that just ain't so.*

—Mark Twain (1)

Previous studies have shown us that cardiomyopathy patients of a nonischemic etiology (NICM) have a lower absolute risk for cardiac arrhythmias and sudden death as compared to patients with ischemic cardiomyopathy (ICM) (2,3). The guidelines have also reflected this by delegating a class IIa indication for implantable cardioverter-defibrillators (ICD) in NICM patients versus a class I indication in ICM patients (4). Importantly, the guiding principle for the use of ICD in the NICM population was maximally influenced by the SCD-HeFT (Sudden Cardiac Death in Heart Failure Trial), which included patients with an ejection fraction of $\leq 35\%$ and did not upfront discriminate based on the QRS duration. Therefore, SCD-HeFT included a sizeable proportion of patients that may have had coexisting electrical dyssynchrony. In both the SCD-HeFT and DEFINITE (Defibrillators in Nonischemic Cardiomyopathy Treatment Evaluation) trials, 20% of patients had a left bundle branch block. Notably, cardiac resynchronization therapy (CRT) was still investigational at the time the SCD-HeFT, and DEFINITE trials were being conducted, and the disease modifying impact of CRT could not be taken into consideration. The DANISH (Danish Study to Assess the Efficacy of ICDs in Patients with

Nonischemic Systolic Heart Failure on Mortality) (5), however, while investigating the impact of ICD in NICM patients through a randomized study, in accordance with contemporary guidelines, also appropriately used CRT in more than one-half of their patients. Consequently, the absence of benefit of ICD in this group, probably secondary to CRT-induced favorable cardiac remodeling of CRT, has brought into question the overall role of ICD in the NICM population. There are a few considerations worth noting regarding the DANISH trial. First, the patient population of nearly 60% CRT-indicated patients is likely greater than what would be expected in a standard heart failure clinical practice, and second, the remaining narrow complex ICD candidates in the DANISH study may not be entirely reflective of the real-world conventional high-risk ICD patients. Nevertheless, we can ask whether the guidelines for the use of ICD from the yesteryears are applicable in the NICM population with narrow QRS, and whether there is a justifiable role in the wider QRS cohort, who are also indicated for concomitant CRT. In the absence of a large prospective randomized study powered enough with a sample size to answer this, Narayanan et al. (6) sought to answer this question through a systematic review and a meta-analysis. Unlike other meta-analyses attempting to address this question, the investigators compared ICD with medical therapy without the use of CRT and then specifically examined the question of whether a defibrillator therapy should be indicated when

^aDeputy Editor, *JACC: Clinical Electrophysiology*; from the ^bCardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts; ^cUniversity of Washington, Seattle, Washington; and the ^dValue Institute, Christiana Care Health System, Newark, Delaware. Dr. Singh has served as a consultant for Biotronik, Boston Scientific, Medtronic, St. Jude Medical & Liva Nova Group, Respicardia Inc., and Impulse Dynamics. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

TABLE 1 Questions to Ask While Evaluating a Meta-Analysis

Is there a mix of patient populations that would affect the outcome?
What is the I^2 statistic (ignore the p value)?
Do the authors discuss, qualitatively or quantitatively, study quality?
Is there a substantial difference with respect to study dates?
What story do the forest plots tell?

combined with CRT. The big conundrum that ensues is whether a well-conducted meta-analysis such as this one really provides us enough evidence to change our current clinical practice?

Meta-analyses have appeared with increasing frequency in the medical research publications. There were 6 published in *JACC: Clinical Electrophysiology* in 2016, and 5 have already been published through July 2017. Although meta-analysis is viewed by some as the ultimate in evidenced-based method (7) others have a less positive view (8). Arguably, the most important aspect of meta-analysis, especially in medical research, is selecting the studies to be included. Obviously, the studies need to assess the same outcome, but not all studies, including randomized trials, are created equal with respect to patient populations, inclusion/exclusion criteria, treatment/therapy options, data quality, follow-up, etc. Study heterogeneity has long been recognized as a threat to the validity of meta-analysis (9). Meta-analytic methods do not “automatically” overcome between-study differences, including study quality. Fortunately, there is a statistic, which is readily calculated in meta-analytic software, that helps quantify study heterogeneity and can help assess whether the results of the study are valid. The I^2 statistic quantifies the degree of heterogeneity between studies included in the analysis (10,11). Although virtually all meta-analytic papers cite a typical categorization of I^2 (e.g., low: < 25%; moderate: 25% to 50%; high: > 50%) in their methods section and also report I^2 in their results section, very few subsequently interpret the results taking into consideration the implications of this crucial statistic. When there is heterogeneity between studies, the conventional thought is to do a random effects meta-analysis rather than a fixed-effects analysis. In a fixed-effects meta-analysis, the estimated effect is a weighted average of each study's effect where the weights are the precisions (the reciprocal of a study's estimated variance) of each study. A fixed-effects analysis tests the hypothesis that the treatment effects are equal across all studies and that any variation is only a result of sampling (random) error (12). A random-effects analysis incorporates an estimate of between-study heterogeneity into the weights (12,13), and typically results in more conservative conclusions than a fixed-effects analysis, that is, wider confidence intervals and larger p values, though this is not always true (14). It has also been argued that a random effects analysis in the face of study heterogeneity leads to inconclusive results (15). From a computational standpoint, a random effects

meta-analysis redistributes study weights based on sample size or effect size variability and simply moves the analysis back to an estimate based on unweighted means. A fixed-effects analysis in the face of study heterogeneity results in estimates that corresponds to the dominant study, essentially eliminating the others. The implication is that the only meta-analysis that can be interpreted in a straightforward manner is a fixed-effects analysis where the studies are homogeneous (i.e., $I^2 = 0$). There is a statistical test for whether I^2 is significantly different from 0, but significance depends on the number of studies in the analysis. So one study may have an I^2 that is statistically significantly different from 0, but another with the same I^2 that is not significantly different from 0. When studies are heterogeneous, the only meaningful analysis is more qualitative, which is an assessment of how the studies differ from a quality standpoint and how this in turn could affect the results. It has been suggested that the quantitative aspect of meta-analysis can be maintained if a measure of “study quality” is included in the computations (15-17). The criticism here is that “study quality” will be an arbitrary and subjective judgment (18). However, as long as the quality criteria are known, meta-analysis consumers can judge the merits of the criteria and the subsequent results. Otherwise, meta-analysis results where there is study heterogeneity are likely not meaningful and should be viewed with skepticism (Table 1).

The Narayanan et al. (6) study in this issue illustrates a meta-analysis where the results could be meaningfully evaluated. First, the investigators correctly stratify the meta-analyses by CRT status, thus not mixing patient populations. The investigators correctly note the limitations of the small number of studies included as well as the fact that event rates were low (which would also be a problem in the original studies). The I^2 statistic for each analysis was 0, so the fixed-effects analyses were appropriate and the results should be interpretable. But does the statistical absence of heterogeneity truly reflect a homogeneous population, and thereby allow for the generalizability of these results? Even in such a well-conducted analysis as this one, the temporal distribution of the included studies, the consequent mixture of patient types, accompanied by variability in the use of optimal medical therapy can significantly confound the results. To elaborate on this point a little further, the largest of these studies (i.e., SCD-HeFT) was conducted from 1997 to 2001, in which 78% of

the patients used beta-blockers, 86% angiotensin-converting enzyme/angiotensin receptor blocker, and 30% aldosterone antagonists (4). In the DEFINITE trial, 85% received a beta-blocker (3). This is different from the extent of optimal medical therapy used in contemporary clinical trials (e.g., DANISH study) and consequently could lead to overestimating the benefit of ICD., it remains unclear what proportion of patients in each study were optimized to the maximal dose of their medications over the course of each study. At the same time, it is also possible that the DANISH study (5) population could have been reflective of a sicker population than what we typically see in the United States. Especially because elevated N-terminal pro-B-type natriuretic peptide was an enrollment criteria in the DANISH study (5), it is possible that the study selected patients more likely to die from non-sudden cardiac death causes.

Also, even though the heterogeneity of the patient groups within the studies may appear to be low, a meta-analysis cannot account for all covariates and patient characteristics that could affect clinical outcome. Notably, the non-CRT subgroup analysis included studies that did not exclude patients with a wide QRS. Previous large population-based studies have shown that patients with a wider QRS are sicker and have a higher propensity for sudden death. This in turn could further exaggerate the observed beneficial effect of ICD within this subgroup analysis. These studies were conducted before the broader use of CRT and more recent advances in pharmacotherapy, which have been shown to affect the clinical course in these patients and could potentially attenuate the benefit of implantable defibrillators. Also noteworthy is that in the most contemporaneous study (DANISH), more than one-half of the patients with NICM needed a CRT device, which is quite different from what is observed in a real-world heart failure practice (5).

We would like to congratulate Narayanan et al. (6) for a very well-conducted meta-analysis. The investigators extended themselves here by obtaining more specific subgroup data via personal communications. Nevertheless, the amalgamation of numbers and data from a slew of studies, accompanied by a deep yearning to provide an answer does not guarantee that we may actually be able to extricate the right answer from a given set of clinical trials. Beyond recognizing the importance of the I^2 statistic in assessing heterogeneity, we need to read between the lines and pay attention to the “missing” covariates that could still significantly affect the interpretation of the analysis. It seems reasonable to reiterate that the role of ICD in the NICM substrate has been supported by either small randomized clinical trials or subgroup analysis from studies conducted almost 2 decades ago. One could infer that the number needed to treat with ICD to reduce overall mortality or sudden cardiac death in either the wide QRS patient (with concomitant CRT) or in the low-risk narrow QRS patient would be higher than previously presumed, bringing into question the cost-effectiveness of this intervention in its current form. As clinicians, we need to periodically revisit our indications and guidelines, especially with the influx of new therapies. Repeating large randomized studies may be impractical, but prospective, well-thought-out registries with more thorough individualized risk stratification strategies (using biomarkers and imaging) may help us better ferret out the appropriate high-risk subsets that truly benefit from ICD. We may not necessarily need to reinvent the wheel, but we can serve our patients better, by trying to realign it.

ADDRESS FOR CORRESPONDENCE: Dr. Jagmeet P. Singh, Cardiology Division, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, GRB 842, Boston, Massachusetts 02114. E-mail: jsingh@mgh.harvard.edu.

REFERENCES

1. BrainyQuote: Mark Twain quotes. Available at: <https://www.brainyquote.com/quotes/quotes/m/marktwain109624.html>. Accessed August 11, 2017.
2. Kadish A, Dyer A, Daubert JP, et al., for the DEFINITE Investigators. Prophylactic defibrillator implantation in patients with nonischemic dilated cardiomyopathy. *N Engl J Med* 2004; 350:2151-8.
3. Bardy GH, Lee KL, Mark DB, et al., for the SCD-HeFT Investigators. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. *N Engl J Med* 2005;352:225-37.
4. Epstein AE, DiMarco JP, Ellenbogen KA, et al. 2012 ACCF/AHA/HRS focused update incorporated into the ACCF/AHA/HRS 2008 guidelines for device-based therapy of cardiac rhythm abnormalities: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2013;61:e6-75.
5. Køber L, Thune JJ, Nielsen JC, et al., for the DANISH Investigators. Defibrillator implantation in patients with nonischemic systolic heart failure. *N Engl J Med* 2016;375:1221-30.
6. Anantha Narayanan M, Vakil K, Reddy YN, et al. Efficacy of implantable cardioverter-defibrillator therapy in patients with nonischemic cardiomyopathy: a systematic review and meta-analysis of randomized controlled trials. *J Am Coll Cardiol EP* 2017;3:962-70.
7. Haidich AB. Meta-analysis in medical research. *Hippokratia* 2010;14 Suppl 1:29-37.
8. Chalmers TC. Problems induced by meta-analyses. *Stat Med* 1991;10:971-9, discussion 979-80.

9. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997;315:1533-7.
10. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
11. Higgins JP, Thompson SD, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
12. Senn S. Trying to be precise about vagueness. *Stat Med* 2007;26:1417-30.
13. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
14. Poole C, Greenland S. Random-effects meta-analyses are not always conservative. *Am J Epidemiol* 1999;150:469-75.
15. Al Khalaf MM, Thalib L, Doi SA. Combining heterogeneous studies using the random-effects model is a mistake and leads to inconclusive meta-analyses. *J Clin Epidemiol* 2011;64:119-23.
16. Doi SA, Thalib L. A quality effects model for meta-analysis. *Epidemiology* 2008;19:94-100.
17. Doi SA. Modelling methodologic quality into meta-analyses and the pitfalls of not doing this. *Ann Thorac Surg* 2009;87:985-6.
18. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics* 2001;2:463-71.