Treatment Crossovers Did Not Affect Randomized Treatment Comparisons in the Mode Selection Trial (MOST)

Anne S. Hellkamp, MS,* Kerry L. Lee, PhD,* Michael O. Sweeney, MD, FACC,† Mark S. Link, MD, FACC,‡ Gervasio A. Lamas, MD, FACC,§ for the MOST Investigators

Durham, North Carolina; Boston, Massachusetts; and Miami Beach, Florida

OBJECTIVES
We evaluated the impact of treatment crossovers on study results in the Mode Selection Trial (MOST).

BACKGROUND
The MOST study, a 2,010-patient, 6-year trial comparing dual-chamber pacing (DDDR) and ventricular pacing (VVIR) in sinus node dysfunction (SND) (1). In the major intent-to-treat analyses, the MOST study demonstrated no difference between pacing modes in the primary composite endpoint of all-cause death or stroke, although significant reductions in heart failure hospitalization (HFH) and atrial fibrillation (AF) were observed with DDDR pacing. However, a moderate proportion (38%) of VVIR-randomized patients were temporarily or permanently crossed over to DDDR pacing.

METHODS
Intent-to-treat (ITT) analyses compared treatment arms by randomized pacing mode. On-treatment analyses used time-dependent covariates to account for all crossovers. All analyses used Cox proportional hazards models and included covariates prespecified in the study design: age, gender, Charlson index, and prior stroke, heart failure, myocardial infarction, supraventricular tachyarrhythmia, and ventricular tachycardia or fibrillation.

RESULTS
Of 996 VVIR-randomized patients, 375 (38%) were DDDR paced at some time, accounting for 27% of follow-up days among all VVIR-randomized patients. Of 1,014 DDDR-randomized patients, 53 (5%) were VVIR paced at some time, accounting for 1.5% of follow-up days among all DDDR-randomized patients. On-treatment analyses showed slightly lower hazard ratios favoring DDDR versus VVIR compared with ITT: death or stroke 0.88 (on-treatment) versus 0.91 (ITT); death 0.94 versus 0.95; stroke 0.74 versus 0.81; HFH 0.72 versus 0.73; and AF 0.72 versus 0.77. Interpretation of treatment effects was unchanged.

CONCLUSIONS
Although treatment crossovers accounted for >25% of follow-up time in the VVIR-randomized group, this did not affect study results. End point comparisons between randomized modes are accurate reflections of DDDR versus VVIR pacing in this study population. (J Am Coll Cardiol 2006;47:2260–6) © 2006 by the American College of Cardiology Foundation

Heart Rhythm Disorders

The Mode Selection Trial (MOST) was a 2,010-patient, 6-year trial comparing dual-chamber pacing (DDDR) and ventricular pacing (VVIR) in sinus node dysfunction (SND) (1). In the major intent-to-treat analyses, the MOST study demonstrated no difference between pacing modes in the primary composite endpoint of all-cause death or stroke, although significant reductions in heart failure hospitalization (HFH) and atrial fibrillation (AF) were observed with DDDR pacing. However, a moderate proportion (38%) of VVIR-randomized patients were temporarily or permanently crossed over to DDDD pacing. The purpose of this study was to assess the impact of these treatment crossovers on the interpretation of the MOST study results.

Intent-to-treat analyses are the gold standard of clinical trials. Randomization of treatment assignment ensures that the treatment groups are similar in terms of relevant baseline characteristics, and that any difference in outcomes between the groups can reasonably be attributed to the treatment. However, if a large proportion of patients discontinue treatment or change to another treatment arm, it becomes more difficult to determine the existence and magnitude of treatment effect. On-treatment analyses address this problem by comparing patients who received one treatment to patients who received another, but because treatment assignment is no longer completely random, results can be biased by any association between likelihood of treatment change and risk of event (e.g., if factors that make a patient more likely to change treatment also make them more likely to have the event).

Crossovers may have affected the MOST study by making outcomes in the two treatment arms more homogeneous than they were in reality. The VVIR arm as randomized was really a mix of VVIR-paced patients and patients who were DDDR-paced during part of their follow-up. Therefore, where an underlying treatment effect exists, a comparison of randomized...
treatments might show a nonsignificant effect or might show a significant effect but underestimate its size. Where no underlying treatment effect exists, crossovers would have no impact on randomized mode comparisons.

METHODS

The design and conduct of the MOST study have been previously described (1,2). Briefly, 2,010 patients with SND were randomized to DDDR or VVIR pacing at the time of pacemaker implantation. Patients were seen at 1, 3, and 6 months after implant and every 6 months thereafter; median follow-up was 33 months. Because all patients received a dual-chamber device, mode changes (crossovers) required only programming changes. All mode changes were documented by the sites and reported as part of the case report form. Deaths, strokes, and first HFH were adjudicated by a clinical events committee blinded to treatment. First documented episode of AF was adjudicated by an electrocardiographic core laboratory.

Although pacemakers could be programmed in a variety of ways (e.g., DDIR, DDD, VVI), for the purposes of this analysis all dual-chamber or atrial pacing modes were considered DDDR and all ventricular pacing modes were considered VVIR.

Time in each mode for each patient was calculated from mode change dates provided by the sites. For displaying time in each mode graphically, patients were classified for each month of follow-up by the mode in which they spent the majority of that month. For the purpose of characterizing crossover and noncrossover groups at baseline, patients were classified as DDDR (all patients randomized to DDDR), VVIR-crossover (all VVIR-randomized patients who permanently crossed to DDDR or who spent at least one third of follow-up time in DDDR), and VVIR-no crossover (all other VVIR-randomized patients). Continuous baseline variables are summarized by group as median (25th, 75th), and categoric variables are summarized as percent (number). Baseline variables were compared between VVIR-crossover and VVIR-no crossover groups using Wilcoxon rank sum tests for continuous variables and likelihood ratio chi-square tests for categoric variables.

Cox proportional hazards models (3) were used to examine the association of pacing mode with event risk. Intent-to-treat analyses compared treatment arms according to randomized pacing mode. On-treatment analyses used a time-dependent covariate to account for all mode changes during follow-up. Thus, event-free follow-up time for a patient before any crossover would be credited to the treatment arm to which the patient was randomized, whereas at crossover, the patient would shift to the other arm. Some patients shifted between pacing modes multiple times. All Cox model treatment comparisons were adjusted for eight baseline covariates that were prespecified for secondary adjusted comparisons in the study design: age, gender, prior stroke, prior HF, prior myocardial infarction, Charlson comorbidity index (4), prior supraventricular arrhythmia, and prior ventricular tachycardia or fibrillation. Relative risk for each event is presented as a hazard ratio and 95% confidence interval.

Event-free survival curves for groups defined by randomized pacing modes were generated using Kaplan-Meier estimates (5). For on-treatment event-free survival curves, event rate calculations were modified by counting patients in the mode they were in at each event time. Thus the curves represent the cumulative event-free rates of each group over time, although in the on-treatment curves, the actual membership of each group changed over time as a result of pacing mode crossovers.

All analyses were conducted using SAS version 8.2 (SAS Institute, Cary, North Carolina).

RESULTS

Of 996 VVIR-randomized patients, 375 (38%) were DDDR paced for at least part of their follow-up time in the study; 313
(31%) remained in DDDR at the last follow-up visit. Time spent in DDDR pacing accounted for 27% of follow-up days among all VVIR-randomized patients (Fig. 1). Of 1,014 DDDR-randomized patients, only 53 (5%) were VVIR paced at some point, accounting for 1.5% of follow-up days among DDDR-randomized patients.

In several respects, VVIR-randomized patients who crossed over to DDDR pacing tended to be sicker than patients who remained in VVIR: They had more prior HF, angina, hypercholesterolemia, and prior revascularization and were more likely to be on antiarrhythmic therapy at the time of hospital admission (all p < 0.05) (Table 1). There was also a trend

<table>
<thead>
<tr>
<th>Event</th>
<th>ITT HR (CI)</th>
<th>OT HR (CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death/stroke</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td>0.91 (0.75, 1.10)</td>
<td>0.88 (0.73, 1.07)</td>
<td>0.32</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>0.95 (0.78, 1.16)</td>
<td>0.94 (0.76, 1.15)</td>
<td>0.64</td>
</tr>
<tr>
<td>AF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF</td>
<td>0.81 (0.54, 1.23)</td>
<td>0.74 (0.48, 1.12)</td>
<td>0.33</td>
</tr>
<tr>
<td>HFH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFH</td>
<td>0.73 (0.56, 0.95)</td>
<td>0.72 (0.55, 0.94)</td>
<td>0.021</td>
</tr>
<tr>
<td>AF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF</td>
<td>0.77 (0.64, 0.92)</td>
<td>0.72 (0.60, 0.86)</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>0.70 (0.54, 0.92)</td>
<td>0.68 (0.52, 0.93)</td>
<td>0.017</td>
</tr>
</tbody>
</table>

*Age and weight are given as median (quartiles). All other variables are given as % (n). †Crossover patients are VVIR-randomized patients who permanently crossed to DDDR or who spent at least one third of follow-up time in DDDR. ‡Low ejection fraction is defined as ejection fraction <50% or, where ejection fraction was not measured, clinical impression of any left ventricular dysfunction. §VVIR-crossover group different from VVIR-no crossover group at p < 0.05.

DDDR = dual-chamber pacing; HF = heart failure; NYHA = New York Heart Association; VVIR = ventricular pacing.
toward more New York Heart Association (NYHA) functional class III or IV congestive heart failure and a family history of sudden death or myocardial infarction (both \( p < 0.06 \)).

On-treatment comparisons showed slightly lower hazard ratios and \( p \) values favoring DDDR versus VVIR pacing compared to intent-to-treat comparisons (Fig. 2).

Event-free survival curves for the primary end point of death or stroke and the secondary end points of HFH and AF were similar for the intent-to-treat and on-treatment comparisons (Figs. 3 to 5).

**DISCUSSION**

Since the MOST results were published (1), questions have been raised about their interpretability because of the relatively high crossover rate. This analysis finds that the impact of crossovers on the trial results is negligible.

On-treatment analyses in which pacing mode is a time-dependent covariate, taking into account changes from single-chamber pacing to DDDR or vice versa, reveal that the randomized mode comparisons may have slightly underestimated treatment effects. None of the nonsignificant tests became significant in the time-dependent covariate models, and decreases in hazard ratio estimates were small.

**Death.** The hazard ratio for death changed only from 0.95 to 0.94, which is consistent with chance when analyzing slightly regrouped patients when no treatment effect exists. Therefore the results of the on-treatment analysis support those of the intent-to-treat analysis, i.e., no relationship between pacing mode and mortality risk. This result is consistent with that of another large pacing study, the Canadian Trial Of Physiologic Pacing (CTOPP) (6), which found no difference in mortality in a comparison of ventricular versus physiologic pacing in 2,568 patients with SND or atrioventricular block. Similarly, the UK Pacing and Cardiovascular Events (UKPACE) trial (7) found no difference in mortality among 2,021 patients with high-grade atrioventricular block randomized to dual- or single-chamber pacing. Results in smaller trials have been mixed. The Pacemaker Selection in the Elderly (PASE) trial (8) found no mortality difference in a comparison of ventricular versus dual-chamber pacing in 407 elderly patients.
with SND or atrioventricular block; Andersen et al. (9) found a difference in unadjusted but not adjusted tests in a comparison of atrial and ventricular pacing in 225 patients with SND.

**Stroke.** Strokes were a relatively rare occurrence in the MOST study, only 90 occurring during the course of the trial. It is notable that the intent-to-treat hazard ratio was 0.81, well below 1 and not much different from the hazard ratio for AF, 0.77, which was significant. With fewer stroke events, however, the confidence interval for the stroke hazard ratio was wide; therefore, the treatment comparison with respect to stroke was not significant. Thus it is not possible in the MOST study to determine whether there is no underlying effect of pacing mode on the occurrence of stroke or whether an effect exists but the trial was simply underpowered to detect it. In the on-treatment analysis, stroke showed the largest decrease in hazard ratio of all the end points, from 0.81 to 0.74, with a corresponding drop in p value from 0.33 to 0.15. This raises the intriguing possibility that a DDDR benefit in reducing stroke does exist that the MOST study could not detect, and that more patients or a longer follow-up (or both) might show a statistically significant 20% to 25% relative risk reduction. This would agree with Andersen et al. (9), who found a reduction in thromboembolic events with atrial versus ventricular pacing. In another small study, Mattioli et al. (10) found a reduction in strokes with physiologic pacing in 210 patients with SND or atrioventricular block. It would, however, be in contrast to the other large trials: CTOPP, which found no stroke reduction, and UKPACE, which found no reduction in a composite of stroke, transient ischemic attack, or other thromboembolism with dual-chamber pacing.

**HFH.** There was almost no change in the hazard ratio for HFH from the intent-to-treat analysis (0.73) to the on-treatment analysis (0.72). It is perhaps surprising that we do not see a large decrease with the latter analysis, because this is an end point for which there is evidence from the MOST study that a DDDR benefit exists.

Crossovers have only a minimal impact in on-treatment comparisons of HFH when the patients who change to the other treatment have a higher baseline risk.
of the event, because their higher risk cancels out some of the benefit gained from DDDR pacing. The VVIR patients who required crossover had a higher baseline prevalence of HF risk factors (prior HF, higher NYHA functional class, antiarrhythmic therapy [11]) than the noncrossover VVIR patients and therefore a higher baseline risk of HFH.

The significant reduction in HFH risk with dual-chamber pacing is in contrast to the CTOPP and UKPACE trials, which found none. In the PASE trial, the composite end point of death, stroke, or HFH showed a nonsignificant (p/H11005 0.07) trend toward dual-chamber benefit among patients with SND. A large retrospective study found a significant association between risk of HFH and single-compared with dual-chamber pacing in 11,426 pacemaker recipients with no prior HF (12).

Atrial fibrillation. Similarly, a decrease in hazard ratio would be expected for AF because a DDDR benefit was also apparent in this intent-to-treat analysis, and it does show a modest decrease from 0.77 to 0.72. Unlike HFH, there does not seem to be any imbalance in the most important AF risk factor in the MOST study (prior supraventricular tachyarhythmias) between crossover and noncrossover VVIR patients. Therefore there would be no higher AF risk among crossover patients to cancel out the DDDR benefit, and the decrease of 0.05 in the hazard ratio is what might reasonably be expected.

In contrast to the mixed evidence for the other end points, a reduction in AF with dual-chamber pacing has been found in almost all other trials, including the CTOPP study, the PASE trial (13), and Mattioli et al. (10). Only the UKPACE trial, in which all patients had high-grade atrioventricular block, found no difference in atrial fibrillation.

Crossovers. Hardware-randomized trials, in which implanting a new device is required to change modes, typically have low rates of crossover from ventricular to dual-chamber pacing. In the CTOPP study the rate was 2.7% after 3 years, and in the UKPACE trial 3.1% after 3 years; Andersen et al. (9) reported 3.5% crossovers in long-term follow-up (mean 5.5 years). In contrast, crossovers are much more common in trials in which only programming changes are required.
The MOST study had a 31% permanent crossover rate, and the PASE trial had 26%. Various criticisms have been made of the latter type of trial design (14–17), in which the ease of mode change may artificially increase the estimate of the rate of pacemaker syndrome or other reasons for crossover and otherwise confound pacing mode comparison. In this analysis, we have determined that the rate of crossovers from ventricular to dual-chamber pacing in the MOST study did not, in fact, affect the interpretability of the results of the intent-to-treat analysis.

Study limitations. This analysis was not prespecified in the trial design. Although pacing mode was randomized, crossovers were not, so there are possible sources of bias in on-treatment analyses. Because mode is no longer truly random in the on-treatment analyses, it is not possible to imply cause and effect. However, it is important to note that the purpose of this analysis was to supplement the intent-to-treat results rather than to supersede them in any way.

Although we have considered pacing mode as the only variable that changes over time, clearly there are other variables that may have changed over the course of long-term follow-up, including comorbidities, medical therapy, and device-related factors. Had all of these changes been recorded during the trial and allowed to change over time in the analysis, the relationships of pacing mode to events might have been different.

For the end points of AF and HFH, only the first occurrence was considered, and subsequent episodes were neither systematically recorded nor adjudicated during the trial. It is possible that a change from VVIR to DDDR reduced the risk of subsequent episodes of AF or HFH. If all episodes had been considered, a more dramatic difference might have been seen between intent-to-treat and on-treatment analyses.

Conclusions. Although treatment crossovers accounted for more than 25% of follow-up time in the VVIR-randomized group, this did not affect study results. End point comparisons between randomized modes are accurate reflections of DDDR and VVIR pacing in this study population.

REFERENCES


Reprint requests and correspondence: Anne S. Hellkamp, MS, 20177 105th Avenue NE, Bothell, Washington 98011. E-mail: anne.hellkamp@duke.edu.