After decades of study, the appropriate strategy for managing asymptomatic patients with severe mitral regurgitation (MR) remains controversial. Because no randomized, controlled trials (RCTs) have been performed, current “practice guidelines” are based on inferences drawn from comparison of observational cohorts. Generally perceived consistency among these studies has led to a relatively firm consensus that asymptomatic patients should undergo mitral valve surgery when they manifest any one of several “high risk descriptors,” identified from the observational cohorts (1–3), unless symptoms develop first. Moreover, as more and more data have accumulated and the risks of surgery have progressively diminished, it has been suggested that severe MR, itself, should be an indication for surgery and that “watchful waiting” (i.e., close follow-up to detect the earliest occurrence of symptoms or high risk descriptors) should be abandoned (4). A caveat associated with the latter recommendation is that surgery is appropriate only if mitral valve repair (MVr) is highly likely to be used rather than mitral valve replacement (MVR) (2). This caveat is based on the increasingly more strongly held “article of faith” that, when surgery is needed for MR, MVr outperforms and is preferable to MVR (5) because of lower perioperative and, possibly, long-term mortality, avoidance of long-term anticoagulation, superior residual left ventricular function (due to preservation of valve and subvalvar structure and function, even in comparison with valve replacement performed with chordal preservation), and so on, despite apparently less durability of MVr than of MVR. Many published studies support the superiority of MVr to MVR, but, as for surgery versus watchful waiting, such support is based on inferences from matched observational cohorts, not from RCTs. RCTs never have been performed among patients with degenerative MR (i.e., MR due to leaflet abnormalities, as in mitral valve prolapse or flail leaflet), although very recently the first landmark RCT was performed in patients with functional MR (i.e., MR due to myocardial disease) (6).

Although their study is a RCT, Kang et al. (7) in this issue of the *Journal* have added importantly to the data informing selection of early surgery versus watchful waiting by creating a relatively large and recent series and by performing propensity matching to attempt to simulate the strengths of the RCT. Their study included “610 consecutive patients” who were asymptomatic and lacked high-risk descriptors or clinically apparent coronary artery disease and who underwent surgery between 1996 and 2009 at 2 large medical centers in South Korea. Although presented as “prospective”, this study is similar to others of this genre in that patients actually were part of ongoing registries established at the 2 centers for other purposes; patients were identified as having undergone surgery retrospectively and analyzed post hoc. This approach must be contrasted with that of Rosenhek et al. (8), who designed a prospective study specifically to assess the watchful waiting strategy and performed serial assessments in all patients for this purpose. The Rosenhek et al. (8) design has the advantage of precision and lower likelihood of unintentional selection bias, but it too was a RCT and used historical controls, potentially misleading, to conclude that watchful waiting is not inferior to early surgery.

The Kang et al. (7) design inherently involves potential unintentional selection biases that preclude unambiguous interpretation of the results: patients who undergo repair tend to be younger, less sick, less affected by comorbidities, and, perhaps most importantly in this particular study, less technically complex than those who undergo replacement. Moreover, in the study of Kang et al. (7), the second center was added after the first center already had reported results of an analysis of its registry, specifically to add power for a more definitive conclusion. The details of the construction of the second registry are not provided. Nonetheless, although suboptimal, such studies often produce useful and thought-provoking data that, although technically hypothesis generating, are used clinically as if they were hypothesis testing. The reason for this is that, unfortunately, the optimal RCT study design is very difficult to mount (e.g., large numbers needed for adequate power are difficult to recruit, biases of physicians
and of patients often preclude the presumption of equipoise and, not surprisingly, has never been performed to assess optimal strategy for asymptomatic degenerative MR.

Kang et al. (7) found significantly fewer clinical outcome events among the group undergoing early surgery than among those who underwent watchful waiting. The results clearly support the strategy of early surgery based on the presence of severe MR alone. The authors’ methodology is praiseworthy in that patients were censored when they reached a high-risk descriptor that would have triggered surgery if guidelines had been observed, thus avoiding inclusion of events that, in theory, might have been preventable. The 610-patient study is relatively large for a valve disease outcome study, but quite small when compared with the thousands commonly included in, for example, heart failure studies (9). More importantly, however, only 92 outcome events were recorded among these patients, a relatively small number that minimizes confidence in the quantitative precision of the relative risk calculations. Moreover, the authors concluded that patients assigned to 1 strategy or the other were not rigorously comparable, minimizing the value of a statistical assessment. Therefore, they undertook a meticulous propensity-matching strategy to create groups better matched for comparisons. The propensity-matched population included two-thirds of the original cohort, or 414 patients, which produced only 61 outcome events, adding greater uncertainty to interpretations of results. Unfortunately, even among the propensity-matched patients, unintentional selection bias seems likely: in the early surgery group, 94% underwent valve repair, putatively the solution providing the better outcomes. In the “conventional treatment” group, significantly fewer of those who underwent surgery had repairs (82%, \( p < 0.001 \)). In a retrospective review of data from a registry designed for a different purpose, it is not surprising that we have no information about the reason for nonselection of repair in the “conventional treatment” group, but it seems likely that factors that mitigated against performance of repair led to selection of “conventional” treatment rather than early surgery, in the first place, potentially minimizing the validity of any statistical comparison.

Although other issues raise concern about interpretation of post-hoc analysis of these registry data, 1 issue, the use of drugs, may be of particular concern. Approximately one-third of the study group received cardioactive drug therapy. The reason for this therapy is unclear. The authors have implied that all therapy was for hypertension, alone, which was said to be present in approximately one-third of the cohort. The retrospective nature of the data extraction precludes rigorous confirmation of this inference or determination of the reasons for selection of specific drugs. Certainly, it seems intuitively reasonable to treat for hypertension in patients with valve disease (although, in the absence of any evidence, clinicians often use vasodilators in nonhypertensive patients with degenerative MR), but the selection of drugs is another issue. Both experimental and (the very few available) clinical data provide no evidence of the benefit and, perhaps, even detriment, when angiotensin-converting enzyme inhibitors and angiotensin receptor blockers are used, and recent experimental data have demonstrated a higher mortality when severe MR is treated with beta-blockade than with nothing (10). Kang et al. (7) provide no information about the relationship of drug use to specific outcomes, good or bad. Here, again, a RCT design would have minimized concern about this potential confounder.

The data of Kang et al. (7) are consistent with the current growing but unproven consensus that early surgery is better than watchful waiting for severe asymptomatic MR. However, in my view, they provide a better argument for the urgent need for a well-organized, appropriately powered RCT to assess the best strategy for such patients. Although such a study never has been mounted, the recent RCT among patients with functional MR (6) indicates that this critically important step may be feasible. Indeed, the functional MR study strengthens the argument for an RCT for degenerative MR. Until the performance of the study, it was strongly believed that, for functional MR as for degenerative MR, repair is the only appropriate choice. The randomized trial revealed no apparent difference in survival or in objective functional/geometric outcomes at 12 months after surgery, although the repair group more frequently manifested recurrent MR. Like all trials, this RCT had notable deficiencies, but its outcome profoundly altered strongly held conclusions developed solely on a base of observational data. Although its strengths are several, the ambiguities of the study of Kang et al. (7) argue most heavily in favor of a serious effort toward RCTs to define appropriate management of asymptomatic degenerative MR.

Reprint requests and correspondence: Dr. Jeffrey S. Borer, State University of New York Downstate Medical Center, 450 Clarkson Avenue, Brooklyn, New York 11203. E-mail: jsborer1@gmail.com.

REFERENCES


Key Words: echocardiography • mitral valve • regurgitation • surgery.