Clinical trial design principles and endpoint definitions for transcatheter mitral valve repair and replacement: part 1: clinical trial design principles

A consensus document from the mitral valve academic research consortium

Gregg W. Stone1,2,*, Alec S. Vahanian3, David H. Adams4, William T. Abraham5, Jeffrey S. Borer6, Jeroen J. Bax7, Joachim Schofer8, Donald E. Cutlip9, Mitchell W. Krucoff10, Eugene H. Blackstone11, Philippe Généreux1,2,12, Michael J. Mack13, Robert J. Siegel14, Paul A. Grayburn15, Maurice Enriquez-Sarano15, Patrizio Lancellotti16, Gerasimos Filippatos17, and Arie Pieter Kappetein18, for the Mitral Valve Academic Research Consortium (MVARC)

1Columbia University Medical Center/New York-Presbyterian Hospital, New York, New York; 2Cardiovascular Research Foundation, New York, New York; 3Hôpital Bichat, Paris, France; 4Mount Sinai Health System, New York, New York; 5The Ohio State University, Columbus, Ohio; 6SUNY Downstate Medical Center, Brooklyn, New York; 7Leiden University Medical Center, Leiden, the Netherlands; 8Hamburg University Cardiovascular Center, Hamburg, Germany; 9Beth Israel Deaconess Medical Center, Boston, Massachusetts; 10Duke University Medical Center, Durham, North Carolina; 11Cleveland Clinic, Cleveland, Ohio; 12Hôpital du Sacré-Cœur de Montréal, Montreal, Quebec, Canada; 13Baylor University Medical Center, Dallas, Texas; 14Cedars-Sinai Medical Center, Los Angeles, California; 15Mayo Clinic, Rochester, Minnesota; 16University Hospital of Liège, Liège, Belgium; 17Athens University Hospital Attikon, Athens, Greece; and 18Erasmus University Medical Center, Rotterdam, the Netherlands. For complete information on the MVARC members and participants, please see the Online Appendix

Received 5 March 2015; revised 5 May 2015; accepted 21 May 2015

See page 1849 for the editorial comment on this article (doi:10.1093/eurheartj/ehv334)

Mitral regurgitation (MR) is one of the most prevalent valve disorders and has numerous aetiologies, including primary (organic) MR, due to underlying degenerative/structural mitral valve (MV) pathology, and secondary (functional) MR, which is principally caused by global or regional left ventricular remodelling and/or severe left atrial dilation. Diagnosis and optimal management of MR requires integration of valve disease and...
heart failure specialists, MV cardiac surgeons, interventional cardiologists with expertise in structural heart disease, and imaging experts. The introduction of transcatheter MV therapies has highlighted the need for a consensus approach to pragmatic clinical trial design and uniform endpoint definitions to evaluate outcomes in patients with MR. The Mitral Valve Academic Research Consortium is a collaboration between leading academic research organizations and physician-scientists specializing in MV disease from the United States and Europe. Three in-person meetings were held in Virginia and New York during which 44 heart failure, valve, and imaging experts, MV surgeons and interventional cardiologists, clinical trial specialists and statisticians, and representatives from the U.S. Food and Drug Administration considered all aspects of MV pathophysiology, prognosis, and therapies, culminating in a 2-part document describing consensus recommendations for clinical trial design (Part 1) and endpoint definitions (Part 2) to guide evaluation of transcatheter and surgical therapies for MR. The adoption of these recommendations will afford robustness and consistency in the comparative effectiveness evaluation of new devices and approaches to treat MR. These principles may be useful for regulatory assessment of new transcatheter MV devices, as well as for monitoring local and regional outcomes to guide quality improvement initiatives.

**Keywords**
Heart failure • Mitral regurgitation • Mitral valve • Valve intervention • Valve surgery (or cardiac surgery)

**Abbreviations and acronyms**

<table>
<thead>
<tr>
<th>FDA</th>
<th>U.S. Food and Drug Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDMT</td>
<td>guideline-directed medical therapy</td>
</tr>
<tr>
<td>LA</td>
<td>left atrial</td>
</tr>
<tr>
<td>LV</td>
<td>left ventricular</td>
</tr>
<tr>
<td>LVEF</td>
<td>left ventricular ejection fraction</td>
</tr>
<tr>
<td>MR</td>
<td>mitral regurgitation</td>
</tr>
<tr>
<td>MV</td>
<td>mitral valve</td>
</tr>
<tr>
<td>MVARC</td>
<td>Mitral Valve Academic Research Consortium</td>
</tr>
<tr>
<td>TEE</td>
<td>transoesophageal echocardiography</td>
</tr>
<tr>
<td>TTE</td>
<td>transthoracic echocardiography</td>
</tr>
</tbody>
</table>

Mitral regurgitation (MR) is the most prevalent valvular disease in the United States and Europe, and along with aortic stenosis, is one of the most frequent valve disorders referred for surgical correction.\(^1\)\(^-\)\(^4\) In contrast to aortic stenosis, which is typically characterized by severe and homogenous cusp calcification, MR is heterogeneous in aetiology, mechanisms, and pathoanatomy. MR may develop either from primary pathology involving any of the components of the mitral valve (MV) apparatus (primary MR, also known as organic MR, usually due to degenerative MV disease) or arise secondarily to left ventricular (LV) dysfunction or occasionally from left atrial (LA) dilation (secondary MR, also known as functional MR).\(^1\)\(^-\)\(^3\) \(^6\)\(^-\)\(^7\) Surgical MV repair is the recommended approach for severe primary MR, with a recently accepted role for transcatheter repair for patients who are at very high or prohibitive surgical risk.\(^1\)\(^-\)\(^3\) \(^6\)\(^-\)\(^7\) Conversely, secondary MR is typically treated with medications and (if indicated) biventricular pacing for heart failure, and coronary revascularization when appropriate, with the utility of MV surgery and transcatheter devices representing active areas of investigation.\(^8\) Few randomized trials, however, have been performed to evaluate the safety and efficacy of MV therapies. The introduction of transcatheter MV devices and the performance of a randomized trial comparing 1 such device to MV surgery\(^9\) have exposed the complexities required to properly evaluate MR therapies, specifically regarding the appropriate study population and control group, background medications and procedures, efficacy and safety endpoints, learning curve issues, and analysis cohort and statistical considerations.\(^8\)\(^,\)\(^9\) Moreover, although the outcomes of patients with MV disorders are sometimes tracked at single centres\(^10\)\(^,\)\(^11\) or in national databases,\(^12\)\(^,\)\(^13\) no standardized endpoints and definitions have been proposed to provide consistency and uniform interpretability of reported results.

The Academic Research Consortium was organized as a collective endeavour between leading academic research organizations and physician-scientists to reach consensus as to what constitutes meaningful clinical endpoints and definitions for evaluation of cardiovascular devices.\(^14\) In collaboration with the U.S. Food and Drug Administration (FDA) and supported by device manufacturers, prior Academic Research Consortium initiatives have addressed consensus endpoints for events following percutaneous coronary intervention and transcatheter aortic valve replacement (TAVR),\(^15\)\(^-\)\(^17\) as well as bleeding definitions,\(^18\) and have been adopted to improve the uniformity and interpretation of clinical studies.\(^19\) The Mitral Valve Academic Research Consortium (MVARC) working group was therefore assembled to develop endpoint definitions for clinical studies of MR therapies. In addition, given the complexity of issues that must be considered for MV trials, MVARC has also developed design principles for clinical trials and registries investigating transcatheter device therapies to treat MR, which may also be applied to surgical and other approaches. Three in-person meetings were held in 2012 to 2014 in which stakeholders and experts in MV disease and therapeutics from the United States and Europe convened to comprehensively review the principles and elements required to successfully investigate and evaluate the relative risks versus benefits of MV therapies. As listed in the Online Appendix, these multidisciplinary gatherings included specialists in general cardiology and valve disorders, heart failure, cardiac surgery, inter-ventitial cardiology, imaging, statistics and epidemiology, and clinical trials. Representatives from the FDA Center for Devices and Radiological Health participated in an advisory role. MVARC was funded by multiple industry sponsors who did not participate in either the sessions or document preparation, but were provided a copy of the report before submission. No fees or honoraria were provided to the writing group or participants.

The present document that resulted from this effort is meant to summarize the current state of knowledge and consensus expert
opinion for MR therapies and is organized in 2 parts: recommenda-
tions for clinical trial design principles (Part 1), and consensus end-
point definitions (Part 2). We acknowledge that the field of MV
therapeutics is highly dynamic and evolving, and we anticipate regu-
lar revisions to these recommendations. Finally, we have concen-
trated our current effort on therapies for primary and secondary
MR; however, many of the principles in this document may also
be applied to other MV conditions, including treatment of mitral
stenosis, degenerated mitral bioprostheses, and failed surgical
valvuloplasty.

**Overview: investigative and regulatory perspectives**

Clinical trials that are intended to support device regulatory
approval or expansion of indications must have clearly documen-
ted objectives and be performed in a highly rigorous manner. In
Europe, the CE mark process requires demonstration that the
device is safe and functions both medically and technically as the
manufacturer intends. Effectiveness is usually investigated
after CE mark approval, and post-marketing surveillance is an inte-
gral part of ongoing clinical evaluation. Either randomized trials or
well-performed registries may support CE mark approval. For ex-
ample, both the MitraClip edge-to-edge device (Abbott Vascular,
Santa Clara, California) and the Carillon coronary sinus annulo-
plasty device (Cardiac Dimensions, Kirkland, Washington)
received a CE mark to treat MR on the basis of registry data dem-
onstrating safety.

For U.S. FDA regulatory approval, high-risk class III devices must
demonstrate “reasonable assurance” of both safety and effective-
ness in a well-defined population for its intended use. Pivotal eval-
uations of breakthrough technologies such as transcatheter mitral
repair systems or percutaneous implantable valves will, in most
cases, necessitate randomized controlled trial designs wherein the
new device is compared with the currently established standard of
care therapy, unless approval for a very limited patient cohort
is desired for which randomization is not feasible. For example,
the MitraClip was approved in the U.S. to treat symptomatic pa-
tients with severe primary MR at prohibitive surgical risk on the basis
of high-quality registry data.

For U.S. approval trials, depending on the comparator group,
either a superiority or non-inferiority design for the primary end-
point may be appropriate. Although superiority in either safety
and/or effectiveness is typically preferred for FDA regulatory ap-
proval, a new device may demonstrate non-inferiority for both
and still be approvable as an alternative therapy to the existing
standard of care, depending on the benefit-risk balance. In studies
addressing an unmet clinical need for a severe disease in which the
available therapeutic alternatives are suboptimal, the benefit-risk
profile of an investigational device may also be favourable even if
effectiveness is somewhat less than that of the comparator if
treatment with the investigational device shows evidence of sub-
stantial safety benefits (and is more effective than a putative pla-
cebo). As knowledge accumulates and technology matures,
non-inferiority designs (e.g., comparing a new design to a
previously approved transcatheter device) and even non-
randomized comparisons to performance goals or objective per-
formance criteria may become reasonable to evaluate device
iterations and to expand the indications for use (label expansion)
of existing approved devices.

Primary effectiveness should be evaluated with a clinically rele-
vant endpoint, either a single event type (e.g., hospitalization for
heart failure) or a composite measure (e.g., death or hospitaliza-
tion for heart failure). Additional support for effectiveness can be
obtained through the use of validated instruments demonstrating
improved quality-of-life, improvement in symptom status (e.g.,
New York Heart Association [NYHA] functional classification),
and improved exercise performance. Although at the present
time these measures are not usually sufficient for principal FDA
regulatory device approval, increasing attention is being paid to
patient-centred benefit-risk metrics in device approval decisions.
Evidence of meaningful MR reduction by the device that is sus-
tained over time is important to demonstrate, and improvement
in ventricular volumes and function during follow-up are addition-
al supportive secondary effectiveness endpoints that should be
assessed. Safety assessments may include both short- and long-
term procedural and device-related complications, and a primary
safety endpoint (separate from the primary effectiveness end-
point) should be pre-specified (see Primary and Secondary End-
points). Finally, the duration of follow-up must be sufficient to
guarantee adequate device durability, relevant to the population
being studied and comparable to alternative therapies, if available.
Late device failures may occur after the primary endpoint of
pre-market studies, necessitating robust postmarket surveillance
to monitor long-term device performance after regulatory
approval.

Identifying the intended population for use (e.g., primary vs. sec-
nondary MR, high vs. low surgical risk, and so on) may importantly af-
flect decisions on comparator therapies (e.g., medical, surgical, or
other transcatheter devices; see Control Group Therapies). As a
general principle, because the pathophysiology, prognosis, control
groups, and response to therapies for primary and secondary MR
vary greatly, these 2 conditions should be studied in separate inves-
tigations unless randomization is stratified and each cohort is indi-
vidually powered for both safety and effectiveness. As a corollary,
Inclusion and exclusion criteria must be carefully selected to define
the population of use (see Inclusion and Exclusion Criteria). Because
transcatheter devices for MR are likely to be evaluated over a range
of disease severity and comorbidities, detailed anatomic and clinical
characterization is required, in addition to key surrogates such as
MR quantification and structural cardiac evaluation using imaging
techniques (see Assessment of Mitral Regurgitation: Role of Non-
invasive Imaging).

Determining operative risk is central to defining the population
for intended use of a new device as well as selecting the appropriate
comparator arm. Current scoring systems such as the Society of
Thoracic Surgeons (STS) and EuroSCORE II indexers may not
by themselves be sufficient to define risk or operability in all pa-
tients. Assessment of patient operability (which may define clinical
trial eligibility) should be determined by a local multidisciplinary
heart team after comprehensive patient evaluation (including risk
score assessment). For MR studies, the heart team should usually include valve and heart failure specialists, MV surgeons, interventional cardiologists experienced in transcatheter MV procedures, imaging experts, and potentially others depending on the specific population and device being studied (see also the subsection Role of the Heart Team).

Several trials may now be cited wherein the use of a sham control helped to demonstrate a lack of device efficacy, contrary to the results of prior unblinded investigations.\textsuperscript{24–26} Use of sham controls (if possible) are thus desirable and, in most cases, are ethically justifiable (see also discussion on sham controls in Control Group Therapies). When a sham control is not feasible, additional efforts should be considered to blind the patient and participants involved in data collection to the extent possible (e.g., the use of patient headphones to mask device allocation during the procedure; not recording randomization allocation in the chart; and using separate research coordinators and physicians for device implantation and follow-up). Patient-related outcomes, such as quality-of-life, are considered more robust in studies that can be blinded. For pivotal device trials, the use of independent core laboratories and event adjudication and data safety and monitoring committees are mandatory to ensure patient safety, reduce reporting bias, and enhance credibility, accuracy, and interpretability of study findings, especially when patient and physician blinding is not possible.

For both randomized trials and registry studies of MR therapies, written informed consent must be obtained from all patients unless waivers are provided with specific ethical oversight. Within the framework of a randomized trial, study-eligible patients who decline randomization should ideally be followed in a separate registry to provide additional insights into potential study selection bias and the natural history of the control population. If exploratory comparison with randomized trial arms is contemplated, the statistical methodology must be pre-specified and justified (e.g., propensity scoring analysis with appropriate covariates, and so on).

Finally, although randomized trials with primary clinical endpoints are strongly recommended, given the logistical, time and cost constraints, MVARC acknowledges that many investigations of MV therapies will collect observational or registry data only (preferably compared with either a concurrent or historical control group), or if randomized, will not be powered for clinical endpoints. Potential efficacy endpoints for these studies may include reduction in MR grade, improvement in LV pressures and chamber dimensions, improved quality of life, and enhanced functional capacity (see Primary and Secondary Endpoints). However, currently none of these endpoints have been sufficiently linked to a major clinical outcome such as death or heart failure hospitalization to be considered a true surrogate, especially as procedural risks must be taken into account when considering the benefit-risk profile of a novel therapy. As such, these studies should be considered hypothesis generating with regard to clinical utility. Nonetheless, such investigations are valuable in their own right, and they provide important supportive data when considering the utility of a new device or approach. Further studies are warranted to strengthen the association between these nonclinical endpoints and clinical outcomes such that, in the future, they might serve as primary endpoints in FDA regulatory trials.

Primary versus secondary mitral regurgitation: similarities, differences, and implications for trial design

Classification of mitral regurgitation and implications for mitral valve therapies

Accurate diagnosis of the underlying MV anatomy and pathophysiology is essential to understand the aetiology, mechanism, lesion localization, and severity of MR; to establish its prognosis; and to design appropriate trials of MR therapies. The MV complex is a dynamic structure including the annulus, the anterior and posterior leaflets and commissures, different level chordae tendineae, the papillary muscles, the underlying LV myocardium, and the LA. Pathological changes in any of the components of the MV can lead to MR, and often lesions are present in more than 1 structural component of the valve. Assessment of MR involves comprehensive evaluation of its aetiology and mechanism (the lesion or deformation resulting in valve dysfunction), including the dysfunction type (leaflet motion abnormality).\textsuperscript{22–27} Of note, annular dilation is almost universally present in patients with severe MR, regardless of other structural abnormalities, although it typically develops late. One exception is MR arising secondary to LA dilation (often in the setting of atrial fibrillation), in which annular dilation may be the principal mechanism of MR.\textsuperscript{5,6} Comprehensive characterization of the underlying aetiology and MV lesion(s) in each patient is especially critical in the new device era, as many transcatheter devices mechanistically target only a single component of the MV or a single mechanism of MR.

The mechanism of MR may be described by Carpentier’s classification of leaflet motion: type I: normal leaflet motion (e.g., annular dilatation, leaflet perforation, or clefts), type II: excessive leaflet motion (e.g., chordal elongation or rupture), and type III: restricted leaflet motion (Figure 1).\textsuperscript{10} Type III dysfunction is further subclassified according to restricted leaflet motion predominantly in diastole but also in systole (type IIIa [e.g., rheumatic disease]) versus only in systole (type IIIb [e.g., ischaemic or non-ischemic LV remodelling with leaflet tethering due to local or diffuse ventricular dilatation]). Carpentier’s segmental leaflet anatomy classification is a useful construct when describing MV disease and planning and performing an intervention.\textsuperscript{10}
Primary MR usually implies Carpentier type II dysfunction, but may be type I in endocarditis and type IIIa in cases of rheumatic origin. Primary MV disease is the most common form of MR referred for surgical correction and covers a large spectrum of lesions, ranging from abnormalities in an isolated scallop to multisegment (or generalized) prolapse, and from thin/non-redundant leaflets to thickened leaflets with excess tissue (Barlow’s disease). Prolapse location, the presence of valvular/annular calcification, and the severity of annular dilation may affect the feasibility and choice of surgical and transcatheter mitral repair techniques.

Secondary MR usually implies a Carpentier type IIIb dysfunction, although type I dysfunction with isolated annular dilation may occur secondary to LA dilation. Secondary MR most commonly develops despite a structurally normal MV due to mitral leaflet tethering secondary to ventricular deformation/remodelling, annular dilation/dysfunction, and insufficient LV-generated closing forces. Assessing global LV function and dilation (diameters, volumes, sphericity, mass) and local remodelling (displacement of papillary muscles) as well as MV deformation (coaptation depth, tenting area, and tenting volume in 3 dimensions) is of paramount importance in evaluating the potential for reparability and results of treatment. Tethering may be limited to an isolated leaflet segment on the basis of “localized” ventricular remodelling or be present along the entire MV closure line in end-stage and diffuse ventricular remodelling. The degree of secondary MR may vary greatly depending on loading conditions (more so than in primary MR).

Secondary type IIIb MR can further be sub-classified as arising from underlying ischaemic heart disease (usually prior myocardial infarction) versus non-ischaemic dilated cardiomyopathy (whether idiopathic or due to specific causes such as hypertension). The mitral jet is typically eccentric or commissural in the setting of ischaemic disease and posterior infarction, resulting in posterior leaflet tethering with medial commissural gap, and is central in most cases when the LV is globally dilated due to anterior infarction or non-ischaemic cardiomyopathy, resulting in more symmetric displacement of both papillary muscles.

It is particularly important to differentiate and separate populations of patients with primary versus secondary MR in clinical trial design, as the comorbidities, prognosis, and therapeutic approaches in these patients vary greatly. Most patients with primary MR due to degenerative MV disease achieve long-term event-free survival similar to an age-matched population after MV surgery, provided MR correction is achieved through valve repair surgery rather than valve replacement, and before significant deterioration in LV geometry or function. In contrast, patients with secondary MR have varying degrees of myocardial remodelling and dilation, and...
usually have significant LV dysfunction. Most patients with secondary MR are treated with heart failure therapies (guideline-directed medical therapy [GDMT] ± cardiac resynchronization therapy [CRT] when appropriate) as well as coronary revascularization if substantial ischaemia is present. For patients failing those initial treatments, advanced therapies including LV assist devices and heart transplantation may be considered. In patients with severe LV dysfunction, the long-term prognosis may be dictated more by the extent of ventricular dysfunction and remodelling than the severity of secondary MR. There is currently little evidence that survival or the natural history of the underlying myocardial disease are affected by mitral intervention in patients with secondary MR, although reduction or correction of MR may provide symptomatic relief.34–36

### Assessment of mitral regurgitation: role of non-invasive imaging

**Echocardiographic evaluation of mitral regurgitation**

Echocardiography is fundamental in evaluating the aetiology, mechanisms, and severity of MR, and its effect on cardiac structures and function. In addition, serial echocardiography is essential to demonstrate the effects of medical therapy, devices, and surgical MV repair and replacement over time. Routine 2-dimensional (2D) transthoracic echocardiography (TTE) differentiates whether MR is due to primary valve degeneration or is secondary to LV dysfunction or LA dilation. For primary MR, 2D TTE discriminates the specific pathological changes in the MV complex. In the presence of mixed pathologies, classification can be more difficult (e.g., secondary MR with notable annular calcification or leaflet thickening), although usually a predominant aetiology can be assigned.

Specific anatomical measurements are also useful in assessment of secondary MR (Figure 2), including leaflet length, leaflet angles (particularly the postero-lateral angle, indicating posterior leaflet tethering), coaptation distance (apical displacement of the coaptation point), coaptation length, and tenting area. Asymmetric tenting indicates posterior leaflet restriction, whereas symmetric tenting indicates bileaflet restriction. Measurements of global LV remodelling include LV diameters/volumes and the sphericity index. Measurements of local LV remodelling include apical displacement of the postero-medial papillary muscle, second order chords, and the inter-papillary muscle distance (Figure 2).29,37 Finally, echocardiographic measures of annular dimensions (anterior-posterior diameter >35 mm or the ratio of the anterior-posterior diameter to mid-diastolic anterior MV leaflet length >1.3) due to LV dysfunction, dilation, or dysynchrony have prognostic significance.37,38

### Quantification of mitral regurgitation

Three echocardiography grades of MR severity are generally recognized: mild, moderate, and severe. Whereas this 3-group classification is preferred, a 4-group quantitative scale is sometimes used as well, wherein 1 + = mild MR, 2 + = moderate MR, 3 + = moderate-to-severe MR, and 4 + = severe MR. Because each echocardiographic measurement has specific limitations and lack of precision, an integrated approach incorporating multiple variables should be used to assess MR severity, with somewhat different criteria for primary and secondary MR (Tables 2 and 3).29,39 These include qualitative findings (MV morphology, colour flow, and continuous wave signals of the MR jet), semiquantitative measures (vena contracta width, pulmonary vein flow, mitral inflow), and quantitative measures (regurgitant volume [RVol] and effective regurgitant orifice area [EROA]), as well as supportive findings (enlarged LV and/or LA, increased pulmonary artery pressure [PAP]) (Figures 3 and 4). MR severity should be evaluated by 2D TTE in the non-sedated, non-anaesthetized patient, although 2D and 3-dimensional (3D) transoesophageal echocardiography (TEE) may improve assessment, particularly in secondary MR (Figure 5).

Moreover, for consideration of patient eligibility for a trial evaluating treatment of chronic MR, the echocardiographic severity of MR must be evaluated during a period of clinical stability. If the patient presents with decompensated LV failure, the degree of MR should not be assessed until at least 30 days after the patient has stabilized on a maximal medical regimen.

Colour flow imaging is not solely used for grading MR severity. Localization, duration, timing, and direction of the regurgitant jet into the LA may be useful to evaluate MR, both at baseline and during follow-up after device or surgical intervention. When feasible, the vena contracta width and the flow convergence method (proximal isovelocity surface area [PISA], which permits assessment of RVol and EROA) are strongly recommended. Inherent limitations of the PISA method should be appreciated, however, including

---

**Table 1** Implications of the aetiology of mitral regurgitation

<table>
<thead>
<tr>
<th>Prognosis</th>
<th>Primary Mitral Regurgitation</th>
<th>Secondary Mitral Regurgitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal management strategy (standard of care)</td>
<td>Primarily dependent on the severity of mitral regurgitation and secondarily on left and right ventricular function and pulmonary pressures</td>
<td>Primarily dependent on the degree of underlying left ventricular dysfunction and secondarily on the severity of mitral regurgitation</td>
</tr>
<tr>
<td></td>
<td>Mitral valve surgery when severe (repair preferred to replacement); MitraClip may be considered in patients at prohibitive surgical risk with appropriate anatomy</td>
<td>GDMT for heart failure ± cardiac resynchronization therapy ± coronary revascularization when indicated; mitral valve surgery (repair or replacement) is not common clinical practice but may be considered in selected cases</td>
</tr>
</tbody>
</table>

GDMT, guideline-directed medical therapy.

---

3 G.W. Stone et al.

---

6
reduced accuracy with eccentric or multiple jets (especially common in secondary MR or after transcatheter MV repair with certain devices), changes in PISA radius throughout systole, and difficulty in precisely locating the regurgitant orifice. In addition, the assumption that the proximal flow convergence is hemispheric (vs. ellipsoidal or irregularly shaped, as in secondary MR, leading to underestimation of MR severity) and that it occurs over a flat surface (requiring angle correction in some cases, including post-MitraClip) are important limitations. By permitting direct planimetry of the vena contracta (as well as multiple jets), 3D-TEE may provide a more accurate assessment of MR severity, especially in secondary MR. However, both 2D and 3D colour flow Doppler may overestimate the orifice area due to aliasing and blooming artefacts. Despite these limitations, PISA is a practical method that correlates well with the severity of MR and prognosis.

Importantly, systolic regurgitant flow lasts only as long as mitral leaflet malcoaptation persists; therefore, EROA and RVol are dynamic. For example, in MV prolapse, the EROA appears or increases in mid-to-late systole, whereas in secondary MR, it decreases in mid systole. EROA is usually holosystolic in severe MR. In the current valve guidelines from both the United States and Europe, an EROA ≥ 40 mm² (RVol ≥ 60 ml) indicates severe primary MR, whereas an EROA ≥ 20 mm² (RVol ≥ 30 ml) indicates severe secondary MR. These different thresholds for severe MR due to primary and secondary MV dysfunction have been largely derived from outcome studies demonstrating the prognostic effect of varying degrees of quantitatively measured MR in the 2 conditions. In both cases, however, the regurgitant fraction is ≥ 50%. Of note, however, a regurgitant fraction ≥ 50% can be produced by different values of EROA and RVol, depending on LV volumes and ejection fraction, which can vary widely.
Table 2  Grading the severity of primary mitral regurgitation by echocardiography

<table>
<thead>
<tr>
<th>MR Severity*</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MV morphology</td>
<td>Mildly abnormal leaflets (e.g., mild rheumatic thickening, limited prolapse)</td>
<td>Moderately abnormal leaflets (e.g., moderate thickening or prolapse)</td>
<td>Severe valve lesions (e.g., flail leaflet, ruptured papillary muscle, severe retraction, large perforation)</td>
</tr>
<tr>
<td>Colour flow MR jet</td>
<td>Small LA penetration or not holosystolic</td>
<td>Moderate LA penetration or large penetration and late systolic jet</td>
<td>Deep LA penetration and holosystolic jet</td>
</tr>
<tr>
<td>Flow convergence zone†</td>
<td>Not visible, transient or small</td>
<td>Intermediate in size and duration</td>
<td>Large throughout systole</td>
</tr>
<tr>
<td>CW signal MR jet</td>
<td>Faint/partial/parabolic</td>
<td>Dense but partial or parabolic and light density</td>
<td>Holosystolic and dense or triangular</td>
</tr>
<tr>
<td>Vena contracta width, mm</td>
<td>&lt;3</td>
<td>Intermediate</td>
<td>≥7 (&gt;8 for biplane)†</td>
</tr>
<tr>
<td>Pulmonary vein flow</td>
<td>Systolic dominance</td>
<td>Systolic blunting§</td>
<td>May be normal with low LA pressure. Systolic flow reversal</td>
</tr>
<tr>
<td>Mitral inflow¥</td>
<td>A-wave dominant</td>
<td>Variable</td>
<td>E-wave dominant (&gt;1.5 cm/s)</td>
</tr>
<tr>
<td>TVI mitral/TVI aortic ratio</td>
<td>&lt;1.0</td>
<td>1.0–1.4</td>
<td>&gt;1.4</td>
</tr>
<tr>
<td>Quantitative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EROA, mm²</td>
<td>&lt;20</td>
<td>20–29; 30–39§</td>
<td>≥40</td>
</tr>
<tr>
<td>Regurgitant volume, ml</td>
<td>&lt;30</td>
<td>30–44; 45–59¶</td>
<td>≥60</td>
</tr>
<tr>
<td>LV and LA size</td>
<td>Usually normal</td>
<td>Usually normal or mild dilation</td>
<td>Usually dilated#</td>
</tr>
<tr>
<td>PA systolic pressure, mm Hg</td>
<td>Usually normal</td>
<td>Usually normal</td>
<td>May be normal; &gt;50 at rest without other cause</td>
</tr>
</tbody>
</table>

General considerations: All measurements have limitations, and an integrated approach must be used that weighs the strength of each echocardiographic measurement. All signs and measures should be interpreted in an individualized manner that accounts for body size, sex, and all other patient characteristics. Finally, there may be uncertainty in classifying mild versus moderate and moderate versus severe MR. Further differentiation may be obtained with additional testing (e.g., exercise echocardiography, cardiac magnetic resonance imaging, right and left heart catheterization) if clinically indicated or needed for clinical trial classification. Bolded qualitative and semi-quantitative signs are considered specific for their MR grade. *Mild MR = 1; moderate MR = 2; moderate-severe MR = 3; and severe MR = 4. †With Nyquist limit >50 to 60 cm/s. §For average between apical 2- and 4-chamber views. ¶Signs are non-specific and are influenced by many other factors (LV diastolic function, atrial fibrillation, LA pressure). #Signs are non-specific, are most valid in patients >50 years of age, and are influenced by other causes of elevated LA pressure. *The 2 ranges indicate mild/moderate and moderate/severe MR respectively. EROA 30 to 39 mm² or RVol 45 to 59 ml may be consistent with severe MR in individuals of small body size, particularly women. ¶LV and LA can be within the “normal” range for patients with acute severe MR or with chronic severe MR who have small body size, particularly women, or with small LV size preceding the occurrence of MR. Modified with permission from Lancellotti et al.29 and Zoghbi et al.39

CW, continuous wave; EROA, effective regurgitant orifice area; LA, left atrium; LV, left ventricular; MR, mitral regurgitation; MV, mitral valve; PA, pulmonary artery; TVI, time velocity integral.
LV dysfunction, worsening MR with increase in EROA by exercise-induced pulmonary hypertension. In asymptomatic patients with primary MR and borderline normal values of LV function with exercise is associated with a poor prognosis. Increasing LV integration of all echocardiography data (in secondary MR. Therefore, defining severe MR requires careful integration of all echocardiography data (Tables 2 and 3).

Exercise echocardiography can demonstrate the dynamic nature of MR (mild-moderate MR increasing to severe MR during exercise) and exercise-induced pulmonary hypertension. In asymptomatic patients with primary MR and borderline normal values of LV function and size, worsening of MR (with increasing systolic PAP) and lack of contractile reserve during exercise echocardiography are associated with worse outcomes. In patients with secondary MR and chronic LV dysfunction, worsening MR with increase in EROA by ≥13 mm² with exercise is associated with a poor prognosis. Increasing LV dysynchrony with increased MR can also occur during exercise and may improve after CRT. Improved regional wall motion during (low-level) exercise indicates residual viability, whereas worsening regional wall motion indicates ischaemia. Although exercise echocardiography is increasingly used, the accurate assessment of MR severity during peak exercise remains technically challenging. Pharmacological stress alone is incapable of comprehensively evaluating dynamic changes in MR. Further studies are warranted to evaluate the role of exercise echocardiography in the risk stratification of patients with MR.

Assessing the consequences of mitral regurgitation

LV diameters are derived from M-mode echocardiography or 2D imaging. LV end-systolic diameter > 40 to 45 mm and left ventricular ejection fraction (LVEF) < 60% are indicators of LV systolic dysfunction/dilation in the patient with severe MR. The 2D-based biplane Simpson’s method is recommended for estimation of LV volumes and LVEF; 3D assessment of LV function is generally more accurate than 2D imaging. The LA dilates in chronic volume and pressure overload; the biplane area-length method using apical 2- and 4-chamber views is recommended for assessing LA size. An LA volume index > 60 ml/m² predicts a poor prognosis in primary MR. However, LA dilation is more non-specific than LV dilation, as LA enlargement can also occur in the setting of

<table>
<thead>
<tr>
<th>Table 3 Grading the severity of secondary mitral regurgitation by echocardiography</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Qualitative</strong></td>
</tr>
<tr>
<td>MV morphology</td>
</tr>
<tr>
<td>Colour flow MR jet</td>
</tr>
<tr>
<td>Flow convergence zone</td>
</tr>
<tr>
<td>CW signal MR jet</td>
</tr>
<tr>
<td>Semi-quantitative</td>
</tr>
<tr>
<td>Vena contracta width, mm</td>
</tr>
<tr>
<td>Pulmonary vein flow</td>
</tr>
<tr>
<td>Mitral inflow</td>
</tr>
<tr>
<td>Quantitative</td>
</tr>
<tr>
<td>EROA, mm²</td>
</tr>
<tr>
<td>Regurgitant volume, ml</td>
</tr>
<tr>
<td>LV and LA size and systolic PAP</td>
</tr>
</tbody>
</table>

General considerations: All measurements have limitations, and an integrated approach must be used that weighs the strength of each echocardiographic measurement. All signs and measures should be interpreted in an individualized manner that accounts for body size, sex, and all other patient characteristics. These recommendations are for holosystolic MR. The values of EROA and RVol associated with severe MR (regurgitant fraction > 50%) should be consistent with LV end-diastolic volume, LVEF, and LV forward stroke volume calculated by other methods. The values presented here are rough guides. Functional MR is dynamic, and EROA changes during systole (may be limited to early and late systole) and over time (depending on loading conditions). In such circumstances, single-frame PISA or 3-dimensional measurements may overestimate MR severity. There may be uncertainty in classifying mild versus moderate and moderate versus severe MR. Further differentiation may be obtained with additional testing (e.g., exercise echocardiography, cardiac magnetic resonance imaging, right and left heart catheterization) if clinically indicated or needed for clinical trial classification. Mild MR = 1+; moderate MR 2+; moderate-severe MR = 3+; and severe MR = 4+. At a Nyquist limit of 50 to 60 cm/s. For average between apical 2- and 4-chamber views. Pulmonary venous flow and mitral inflow are indirect signs of MR and are influenced by many other factors such as LV systolic and diastolic function, LA size and pressure, atrial arrhythmias, and the presence of mitral inflow obstruction. Pulmonary venous flow reversal, which is specific to severe primary MR, is rarely observed in severe functional MR. Usually in patients > 50 years of age.

EROA and regurgitant volume by PISA may be substantially underestimated in secondary MR if the regurgitant orifice is elliptical or has multiple jets, as is often the case. Several but not all studies have shown an adverse prognosis with EROA ≥ 30 mm² and regurgitant volume ≥ 30 ml in secondary MR. It is not clear what the cut-off values for mild vs moderate MR should be, in part because of absence of a clear gold standard. 3-dimensional imaging of EROA should be considered in such patients, although it tends to overestimate actual EROA. In secondary MR, LV and LA size and PAP may be increased by the underlying LV systolic and diastolic dysfunction and, therefore, may be increased in all grades of MR. PAP, pulmonary artery pressure; PISA, proximal isovelocity surface area; other abbreviations as in Table 2.
atrial fibrillation or secondary to an increase in LV end-diastolic pressure, whether due to LV diastolic or systolic dysfunction. A systolic PAP (obtained as the sum of the transtricuspid pressure gradient and the estimated right atrial pressure) greater than 50 mm Hg at rest or 60 mm Hg with exercise is strongly associated with adverse outcomes in primary MR. Elevated PAP results in right ventricular pressure overload, which may induce right ventricular failure, and echocardiographic signs include right ventricular hypertrophy, dilation, impaired function, and increased tricuspid regurgitation (peak jet velocity >3 m/s). Tricuspid annular dilation (≥40 mm or ≥21 mm/m²) contributes to tricuspid regurgitation after MV surgery, in which case tricuspid annuloplasty may be considered concomitant with MV surgery.

Echocardiographic eligibility criteria for surgical and transcatheter mitral valve repair or replacement

Surgical and transcatheter MV repair or replacement is generally reserved for severe MR (3+ to 4+). Echocardiography eligibility criteria must also carefully consider the likelihood of procedural success for surgery and the experimental transcatheter device. For example, in primary MR, successful surgical MV repair may be compromised in the setting of multiple complex regurgitant jets, extensive leaflet or valve calcification, and/or when ≥3 scallops (particularly affecting the anterior leaflet) are involved. In secondary MR, the risk of unsuccessful surgical repair or MR recurrence is increased with the presence of severely altered geometry of the MV apparatus, severe global LV remodelling, and/or extensive basal LV scar or aneurysm (Table 4).

Given their varying mechanisms of action, the echocardiographic determinants of successful transcatheter repair of MR are likely to be device specific. For example, the MitraClip reduces MR by grasping and approximating the anterior and posterior mitral leaflets. Echocardiography is indispensable in determining the complexity of the anatomic lesion, and whether the amount of leaflet tissue and coaptation depth and length are sufficient to afford leaflet grasping and approximation by the MitraClip (Table 5). The precise echocardiographic features for procedural success or failure for transcatheter devices that reduce MR by other mechanisms, such as direct or indirect annuloplasty and MV replacement, are notably different and unique to each device.

Imaging during and after mitral valve repair and replacement procedures

Echocardiography is vital for assessing the acute and late results of both surgical and transcatheter mitral interventions. As regards
surgery, 2D TEE (complemented by 3D TEE when available) is performed acutely in the operating theatre after surgical MV repair to exclude more than mild residual valvular MR (e.g., vena contracta width > 3 mm). Adequate leaflet coaptation (length ≥ 8 mm) should be verified. Leakage due to anatomic/technical problems or ring dehiscence and MV stenosis (MV area < 1.5 cm², mean transmirtal gradient ≥ 5 mm Hg) should be excluded. Systolic anterior motion of the MV and injury to the left circumflex artery (expressed as wall motion abnormalities in the basal and mid inferolateral LV segments) due to the close proximity of sutures needed for annuloplasty ring fixation or compression by the ring itself should also be excluded.

LV function may worsen after surgical MV repair and should thus be evaluated in the immediate post-operative period. Historically, this has been attributed to the increase in LV afterload due to reduction in MR. However, after MitraClip repair, cardiac output generally increases, LV filling pressures tend to normalize, and significant LV dysfunction is uncommon, even in patients with severe baseline LV dysfunction. This suggests that the LV dysfunction observed in some patients after MV surgery may be attributable to myocardial oxidative stress, systemic inflammation and free radical injury from cardiopulmonary bypass, cardiac arrest, and cardioplegia, rather than to increased afterload due to the reduction in MR.

In addition to assessing the acute results of transcatheter device repair or replacement of MR, echocardiography is essential to guide most transcatheter MV procedures. For example, 2D and 3D TEE are used to guide each step of MitraClip implantation, complementing fluoroscopy. Immediate post-procedural echocardiographic evaluation includes assessment of residual MR, potential MV stenosis, and exclusion of complications (e.g., pericardial effusion/tamponade, thrombus formation on clips, [partial] clip detachment, and entrapment of chordae by the clip).

Depending on the device, echocardiographic assessment of MR severity after transcatheter MV procedures may pose unique challenges. For example, MR quantification with colour flow Doppler is complex in the setting of a double MV orifice after the MitraClip, and artefacts from the clip(s) hamper quantification. Pulse wave Doppler of the pulmonary veins is useful to corroborate a reduction in MR. Specifically, pulmonary vein flow reversal should be eliminated, and there is often a conversion from the pulmonary vein flow from a “D” dominant pattern (consistent with an elevated LA pressure) to an “S” dominant pattern, reflecting a drop in LA pressure secondary to MR reduction. MV stenosis should be excluded, as evidenced by mean transmirtal valve gradient < 5 mm Hg and MV area ≥ 1.5 cm². Assessment of paravalvular leak is particularly important in patients undergoing transcatheter and surgical MV replacement.

In addition to assessing the acute results of transcatheter device repair or replacement of MR, echocardiography is essential to guide most transcatheter MV procedures. For example, 2D and 3D TEE are used to guide each step of MitraClip implantation, complementing fluoroscopy. Immediate post-procedural echocardiographic evaluation includes assessment of residual MR, potential MV stenosis, and exclusion of complications (e.g., pericardial effusion/tamponade, thrombus formation on clips, [partial] clip detachment, and entrapment of chordae by the clip).

In addition to assessing the acute results of transcatheter device repair or replacement of MR, echocardiography is essential to guide most transcatheter MV procedures. For example, 2D and 3D TEE are used to guide each step of MitraClip implantation, complementing fluoroscopy. Immediate post-procedural echocardiographic evaluation includes assessment of residual MR, potential MV stenosis, and exclusion of complications (e.g., pericardial effusion/tamponade, thrombus formation on clips, [partial] clip detachment, and entrapment of chordae by the clip).

Figure 4 Quantification of mitral regurgitation using the proximal isovelocity surface area method. To calculate the effective regurgitant orifice area (EROA), the radius (r) of the hemispheric convergence flow is measured on a zoomed apical 4-chamber view. To better visualize the largest flow convergence, the colour scale baseline (Nyquist) is reduced to velocities around 35 cm/s. The Nyquist limit is considered the velocity of aliasing (Va) and is introduced in the formula. From the continuous wave Doppler of the regurgitant jet, the peak velocity (Vmax) is measured and the velocity time integral (VTI) is calculated. Regurgitant volume (RVol) is then calculated from the EROA and the VTI.

Figure 5 Evaluation of secondary mitral regurgitation with 3-dimensional transoesophageal echocardiography. (A) Full volume of a dilated left ventricle and tethered mitral leaflets. The left atrial (LA) “en face” view of the mitral valve shows normal mitral leaflets with lack of coaptation between the central scallops (B, arrows). With 3-dimensional transoesophageal echocardiography colour Doppler data, the regurgitant flow is observed from the LA “en face” view of the mitral valve (C, arrow). Post-processing software permits reconstruction of a 3-dimensional model (D) of the mitral valve showing flattening of the mitral annulus and tethering of the mitral leaflets (blue). A = anterior; AL = anterolateral; Ao = aorta; P = posterior; PM = posteromedial.
For both surgical and transcatheter MV procedures, evaluation of the immediate post-repair/replacement results should be performed when the patient’s blood pressure is at least equal to the basal state (and after the effects of anaesthesia have worn off). Comprehensive follow-up TTE is typically recommended at 1 month, at 6 months, and then annually to serially assess MR severity, chamber volumes and pressures, and structural and functional device performance (including the detection of specific device-related technical failure issues and complications as discussed in part 2 of this document).

For clinical trials using serial echocardiographic imaging to assess device performance, study-specific site training and certification in imaging quality before enrolment are recommended, and should be conducted in collaboration with an independent echocardiographic core laboratory.

### Table 4  Unfavourable transthoracic echocardiographic characteristics for surgical mitral valve repair in secondary mitral regurgitation

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ideal Valve Morphology</th>
<th>Unsuitable Valve Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitral valve remodelling</td>
<td>Coaptation distance ( \geq 10 \text{ mm} )</td>
<td>Perforated mitral leaflets or clefts, lack of primary and secondary chordal support</td>
</tr>
<tr>
<td></td>
<td>Tenting area ( \geq 2.5–3.0 \text{ cm}^2 )</td>
<td>Severe calcification in the grasping area</td>
</tr>
<tr>
<td></td>
<td>Complex regurgitant jets</td>
<td>Haemodynamically relevant mitral stenosis</td>
</tr>
<tr>
<td></td>
<td>Posterolateral angle ( \geq 45^\circ )</td>
<td>Length of posterior leaflet ( &lt; 7 \text{ mm} )</td>
</tr>
<tr>
<td>2. Local left ventricular remodelling</td>
<td>End-diastolic diameter ( &gt; 65 \text{ mm} )</td>
<td>Rheumatic valve disease (restriction in systole and diastole) or endocarditic valve disease</td>
</tr>
<tr>
<td></td>
<td>End-systolic diameter ( \geq 51 \text{ mm (end-systolic volume } &gt; 140 \text{ ml}} )</td>
<td>3D TEE gap between leaflets ( &gt; 2 \text{ mm} )</td>
</tr>
<tr>
<td></td>
<td>Systolic sphericity index ( &gt;0.7 )</td>
<td></td>
</tr>
</tbody>
</table>

Adapted with permission from Lancellotti et al.29

### Table 5  Relationship between the morphological characteristics of the mitral valve and suitability for the mitraclip procedure

<table>
<thead>
<tr>
<th>Ideal Valve Morphology</th>
<th>Unsuitable Valve Morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitral regurgitation originating from the mid-portion of the valve (degenerative or functional aetiology)</td>
<td>Perforated mitral leaflets or clefts, lack of primary and secondary chordal support</td>
</tr>
<tr>
<td>Lack of calcification in the grasping area</td>
<td>Severe calcification in the grasping area</td>
</tr>
<tr>
<td>Mitral valve area ( &gt; 4 \text{ cm}^2 )</td>
<td>Haemodynamically relevant mitral stenosis</td>
</tr>
<tr>
<td>Length of posterior leaflet ( \geq 10 \text{ mm} )</td>
<td>Length of posterior leaflet ( &lt; 7 \text{ mm} )</td>
</tr>
<tr>
<td>Non-rheumatic or endocarditic valve disease</td>
<td>Rheumatic valve disease (restriction in systole and diastole) or endocarditic valve disease</td>
</tr>
<tr>
<td>Flail width ( &lt; 15 \text{ mm} ), flail gap ( &lt; 10 \text{ mm} )</td>
<td>3D TEE gap between leaflets ( &gt; 2 \text{ mm} )</td>
</tr>
<tr>
<td>Sufficient leaflet tissue for mechanical coaptation: coaptation depth ( &lt; 11 \text{ mm} ), coaptation length ( &gt; 2 \text{ mm} )</td>
<td></td>
</tr>
</tbody>
</table>

Adapted with permission from Wunderlich et al. 39

3D, 3-dimensional; TEE, tranoesophageal echocardiography.
in the setting of atrial fibrillation. In the future it is likely that CMR and MDCT will be increasingly used for pre-procedural assessment and planning of both surgical and transcatheter MR repair and replacement procedures, and post-intervention surveillance.

**Control group therapies**

Selection of the appropriate control group is essential to interpreting the benefit-risk profile of a new device. For randomized MR clinical device trials, 3 control groups may be considered:
GDMT alone (with or without a sham control) when GDMT is standard of care; (2) GDMT plus surgical therapy when surgical therapy is standard of care; and (3) GDMT plus an active comparator device if an alternative device is available and is considered a standard of care.

Ensuring the use of appropriate GDMT is a requirement for all patients enrolled in randomized controlled trials and registries. It is the basis upon which the safety and incremental efficacy of procedural therapies may be judged. GDMT in symptomatic patients with severe MR includes treatments for heart failure (for all patients with secondary MR due to LV dysfunction, and for those with primary MR with symptoms of heart failure or volume overload (class D), especially those in whom surgery is not performed or will be delayed). GDMT includes not only the use of specific recommended therapies, but also titration of those therapies to recommended target doses, as tolerated. Optimal GDMT use before study enrolment minimizes the likelihood of major changes in medication dosing during the course of a trial, defined for each drug class as an increase in dose by $\geq 100\%$ or decrease in dose by $\geq 50\%$ from baseline. Thus, patients should meet pre-defined GDMT dosing stability criteria before randomization, as the initiation, discontinuation, or titration of therapies after randomization (in either the treatment of control groups) may otherwise seriously confound interpretation of the study results. Although it may not be possible to always prevent major changes in drug dosing (e.g. after improvement in hemodynamics with effective MR therapy), in general such changes should be minimized to isolate the effect of the randomized treatment, unless they are pre-specified and considered as part of the treatment arm strategy (including, for instance, a prospective approach to reduction of heart failure medications).

Achieving and maintaining maximally tolerated guideline recommended doses of beta-blockers, angiotensin-converting enzyme inhibitors or angiotensin receptor blockers, and mineralocorticoid receptor antagonists is especially important before enrolment in secondary MR trials, as reduction in LV dimensions and LV remodelling with effective medical therapy in heart failure may substantially reduce MR in individual patients, obviating the need for advanced or experimental therapies. MR severity and appropriateness for study eligibility should be reassessed at least 30 days (and preferably 90 days) after any major change in GDMT.

Compliance with optimal GDMT in individual patients is often challenging and should be documented at baseline and throughout the course of the study. Before enrolment, the adequacy of GDMT in individual patients (including drug class, dose, and patient compliance) should be verified by a central eligibility committee to reduce bias associated with subjects changing their behaviour under observation post-enrolment (Hawthorne type effect) (see also Role of the Central Eligibility Committee). Intolerance to a drug or drug class or limitation in drug dosing should be on the basis of objective clinical criteria, according to the known adverse effects of specific agents, and must be well-documented in the medical chart and study case report form. Examples include symptomatic hypotension with angiotensin-converting enzyme inhibitors, hyperkalemia with mineralocorticoid receptor antagonists, and symptomatic bradycardia with beta-blockers.

Figure 8 Transoesophageal echocardiogram evaluation of mitraclip implantation in a patient with severe secondary mitral regurgitation. From the midesophageal 4-chamber (A) and bicommissural (B) views, the vena contracta width of the central regurgitant jet can be measured. The 3D LA “en face” view shows lack of coaptation between the anterior and posterior mitral leaflets at the central level (C, arrows). With 3D colour Doppler data, the convergence flow can be observed along the coaptation line from the LV view (D, arrows). Three MitraClip devices were successfully implanted with significant reduction of MR as observed from the colour Doppler biplane views of the MV (E). On 3D transoesophageal echocardiogram full volume of the mitral valve, the LA “en face” view shows a double orifice mitral valve after MitraClip implantation (F). The clips were positioned at the central and anterolateral levels (arrow) leading to a large orifice at the posteromedial level and a small anterolateral orifice (F, asterisks). (G) The colour Doppler 3D “en face” view of the mitral valve with 2 residual mild regurgitant jets. Abbreviations as in Figures 5 and 7.
In addition to GDMT for heart failure, appropriate patients should also be treated with biventricular pacing (CRT) and coronary revascularization when substantial ischaemia is present, according to contemporary clinical practice guidelines, such as those from the American College of Cardiology Foundation/American Heart Association and the European Society of Cardiology/European Association for Cardio-Thoracic Surgery. CRT is indicated (Class I) in patients with NYHA functional class II to IV symptoms on GDMT with LVEF ≤ 35%, sinus rhythm, a left bundle branch block pattern, and QRS duration ≥ 150 ms. In such patients, CRT may substantially decrease LV dimensions and reduce MR in as many as 50% of patients. CRT may also be considered (Class IIa) for selected patients with a left bundle branch block pattern and QRS duration < 150 ms, and for those with a non-left bundle branch block and QRS duration ≥ 150 ms (Class IIa). Surgical or percutaneous coronary revascularization in patients with substantial ischaemia may also, on occasion, reduce secondary MR and should be performed in appropriate patients before study enrolment. After CRT or coronary revascularization, at least 30 days (and preferably 90 days) should pass, after which TTE or other relevant imaging tests are repeated to assess MR severity and appropriateness for study eligibility. Similar to optimal GDMT use, whether CRT and/or coronary revascularization are indicated and utilized should be verified by the central eligibility committee before study enrolment.

**Appropriate scenarios for guideline-directed medical therapy alone (with or without a sham) as the control group**

GDMT should be used alone as the comparator (control) group when a surgical comparator is either not indicated (i.e., is not standard of care) or is contraindicated due to high surgical risk, and no other active comparator exists. Examples of this scenario are seen in recent studies of TAVR for critical aortic stenosis in extreme surgical risk patients and from a single arm registry of the MitraClip for primary MR in prohibitive surgical risk patients. Another example comes from the ongoing COAPT (Cardiovascular Outcomes Assessment of the MitraClip Percutaneous Therapy for Heart Failure Patients with Functional Mitral Regurgitation) trial of the MitraClip for severe secondary MR in chronic heart failure patients (NCT01626079). Although some practice variability exists in this setting, GDMT (rather than MV surgery) is considered the default

---

**Figure 9** Multidetector row computed tomography for assessment of mitral valve geometry in secondary mitral regurgitation. From the reconstructed short-axis view of the mitral valve, orthogonal planes can be placed across the anterolateral, central, and posteromedial levels of the MV leading to the left ventricular outflow tract view at each level. The angles (Aα and Pα) and tenting (MVTh) of the mitral leaflets can be measured at the anterolateral (A1-P1), central (A2-P2), and posteromedial (A3-P3). Reproduced with permission from Delgado et al. AC = anterior commissure; Ao = aorta; PC = posterior commissure; RA = right atrium; RVOT = right ventricular outflow tract.
therapy for most patients with secondary MR, a conclusion sup-
ported by both the current United States and European guide-
lines. Thus, the control group in COAPT is GDMT alone for
patients in whom MV surgery is not considered appropriate after
comprehensive individualized evaluation by the local heart team
(see also the subsection Role of the Heart Team).

For patients randomized to the control group, a sham control
procedure, in which an invasive procedure is performed but the de-
vice is not implanted, should be strongly considered when feasible.
Although the implanting physician cannot be blinded, use of a sham
control minimizes bias by facilitating blinding of study patients as well
as the clinicians and investigators responsible for follow-up study
assessments. There are now several notable examples in which
favourable results from unblinded studies were not supported by
sham-controlled randomized trials, including studies of percutan-
eous myocardial laser revascularization for refractory angina,24 clos-
ure of patent foramen ovale for migraines,25 and renal denervation
for hypertension.26 The major limitations to the use of sham con-
trols are: (1) ethical concerns (e.g., risk of harm with no chance of
benefit); and (2) difficulties in maintaining the blind. The nature of
the sham control will vary according to the control procedure,
and should be selected to maximize the goal of maintaining the blind
while minimizing patient risk. For example, for procedures in which
the experimental procedure requires a transseptal puncture, the
sham control procedure may include femoral venous access and
right heart catheterization. Conversely, a sham control may not
be possible for an invasive procedure such as apical insertion of a
transcatheter mitral valve. Use of a sham control may be less critical
if the primary endpoint of the trial is mortality, although even in this
case bias in an open-label study may differentially affect medical
compliance and crossover to other therapies.

**Appropriate scenarios for surgical therapy as the control group**

Surgical therapy (on a background of GDMT) should be considered
as the control group when surgical therapy is the standard of care
and patients are acceptable surgical candidates. This is the situation
for most patients with primary MR who are not considered to be at
very high operative risk.

**Appropriate scenarios for an active comparator device as the control group**

An active comparator device (on a background of GDMT) may be
considered as the control group when another approved device is
indicated for use in the population being studied. For example, in
the United States, for symptomatic patients with severe primary
MR at prohibitive surgical risk (defined by the FDA as an STS score
for 30-day mortality of ≥8 [replacement calculator] or ≥6 [repair
calculator] or the presence of 1 or more high-risk features that, in
the opinion of an experienced MV surgeon, otherwise precludes
surgery), the MitraClip might currently serve as an active compara-
tor for either a randomized trial or single-arm registry, assuming ap-
propriate MV anatomy. Specific recommendations for control
groups in MR trials appear in Table 6.

**Inclusion and exclusion criteria**

Table 7 details numerous considerations for inclusion and exclusion
criteria for investigational MR trials that may be used as a frame-
work. Although each trial will need to tailor these criteria to the
specific device and patient population being studied, general princi-
plcs may be applied when selecting patients with primary and sec-
ondary MR for enrolment in MV trials.

**Risk scores and surgical candidacy**

A major decision point that must be reached early is whether the
patient is an acceptable surgical candidate. Several risk scores are
in widespread use to determine short-term morbidity and mortality
after cardiac surgery that account for patient comorbidities and ven-
tricular function. General recommendations for the use of risk
scores and assessments of comorbidities for patients undergoing
TAVR have been recently reviewed in the Valve Academic Research
Consortium-2 consensus document.17 MVARC recommends adop-
tion of a similar approach to integration of risk scores and co-
morbidities for studies of devices treating MR, in particular with
regard to the classification of surgical risk related both to prognosis
and selection of the appropriate control group.

The STS score and EuroSCORE II are currently most commonly
recommended for this purpose.22,23 The STS score provides separ-
ate scores for surgical MV repair and MV replacement, and is
recommended for use in clinical trials. Conventionally, very high or "prohibitive" surgical risk is defined by an estimated surgical 30-day mortality of ≥8% using the STS replacement calculator or ≥6% using the STS repair calculator; however, such scores, although having good discrimination, have relatively poor calibration and therefore limited accuracy in identifying extreme risk patients. Moreover, few patients with prohibitive risk for surgical treatment of MR were included in the cohorts used to develop and validate the STS and EuroSCORE II.

Of note, the STS and EuroSCORE II were developed from outcomes in patients who actually underwent surgery, whereas transcatheter devices for MR may warrant evaluation in patients too sick for surgery, who are not represented by these scoring systems. Thus, similar to the Valve Academic Research Consortium-2 recommendations, other variables that are not captured in these scores should also be considered when deciding whether a patient is at excessive risk for surgery, including frailty, major organ system compromise (e.g., cirrhosis), and procedure-specific impediments (Table 8). Additional factors that may preclude surgery include severe mitral annular calcification, the presence of a hostile chest (e.g., prior mediastinal radiation or chest malformation), patent left internal mammary artery bypass graft crossing the midline, prior tracheotomy, and severe pulmonary hypertension with or without right ventricular dysfunction.

A specific issue unique to MV therapeutics refers to use of the STS repair versus replacement calculator to determine surgical risk. This is an important consideration, as the replacement calculator yields greater predicted perioperative mortality. For studies of primary MR in non-prohibitive risk patients in whom surgical MV repair is generally the standard of care, it is appropriate to use the STS mitral repair calculator to determine surgical risk. Conversely, for studies of secondary MR, MV surgery is not generally considered the standard of care, and GDMT (±CRT as appropriate) is the mainstay therapy. In secondary MR patients who are operated on, MV repair has not been proven superior to MV replacement; most patients will be of at least moderate surgical risk given their underlying cardiomyopathy; and most high-risk patients with secondary MR who are operated on currently receive MV replacement rather than repair (at least in the United States). It is, therefore, reasonable to use the STS mitral replacement calculator to determine surgical risk for studies of devices for secondary MR. Additional considerations regarding the choice of the appropriate surgical risk calculator should take into account the device characteristics, access, mode of action, and the procedure the device is intended to replace.

**Role of the heart team**

The standard of care for any individual patient, including assessment of surgical candidacy, appropriate use of GDMT, and potential clinical trial eligibility, should be determined by a multidisciplinary heart team consisting of local experts experienced in the care of patients with MV disease. At a minimum, the heart team should include a heart failure/valve cardiologist, an interventional cardiologist skilled in the relevant access and device implantation procedures, an MV
The heart team tailors adjustment of the decision-making process, and functional status as well as survival must be considered. Importantly, expected improvement in symptoms, quality-of-life, with other known prognostic variables, including assessment of frailty. During the consideration of surgical eligibility, anticipation of individual expected improvement in symptoms, quality-of-life, and functional status as well as survival must be considered. Importantly, the heart team tailors adjustment of the decision-making process according to local expertise and standards of care. Thus, a patient who is considered to be very high risk for MV surgery at one institution may appropriately be considered to be at low or intermediate surgical risk at a different centre. Clinical trials can accommodate such systematic site-based variability by stratification at the time of randomization on the basis of risk assessment by objective scores or the central eligibility committee. Finally, in clinical practice, patient preferences (shared decision-making) play an important role, and arbitrary age- and risk-score-based cut-offs are no longer the dominant basis for treatment selection. For trials leading to regulatory approval or indication expansion, the local heart team determination of surgical risk and eligibility supersedes other considerations, thus ensuring enrolment of a clinically appropriate control group and minimizing crossovers during trial conduct.

### Table 6 Recommended control groups for transcatheter device trials in patients with mitral regurgitation

<table>
<thead>
<tr>
<th>Control Group</th>
<th>Mitral valve surgery (repair preferable to replacement):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary mitral regurgitation</td>
<td>±GDMT* (if heart failure or left ventricular dysfunction present)</td>
</tr>
<tr>
<td>Acceptable surgical risk</td>
<td>GDMT* or MiraClip</td>
</tr>
<tr>
<td>High surgical risk</td>
<td>GDMT*+</td>
</tr>
<tr>
<td>Secondary mitral regurgitation</td>
<td>GDMT*+</td>
</tr>
<tr>
<td>Acceptable surgical risk</td>
<td>GDMT*+</td>
</tr>
<tr>
<td>High surgical risk</td>
<td>GDMT*+</td>
</tr>
</tbody>
</table>

*Maximally tolerated doses of recommended medications for heart failure. Appropriate patients also should have been treated with CRT and/or coronary revascularization before study enrolment. The definition of high surgical risk may vary according to national standards of care. In the United States, per current U.S. Food and Drug Administration guidelines, patients with primary mitral regurgitation should be determined to be at "prohibitive surgical risk" for GDMT or approved transcatheter devices to be considered as an acceptable control group in regulatory trials. In patients for whom the local standard of care for secondary mitral regurgitation is not surgical mitral valve repair or replacement, mitral valve repair or replacement might also be a suitable control group for selected patients in whom the local standard of care for secondary mitral regurgitation is mitral valve surgery, depending on the experimental device characteristics (e.g., for studies of transcatheter mitral valve replacement). GDMT, guideline-directed medical therapy.

### Frailty

Assessment of patient frailty deserves special emphasis. The existence of frailty in an elderly population is an important parameter for risk stratification before major cardiovascular interventions and has demonstrated substantial prognostic capability. Frailty is a geriatric syndrome that reflects a state of decreased physiological reserve and vulnerability to stressors, and is characterized by a progressive decline in muscle mass and strength. Multiple frailty criteria and scales have been proposed, although the single best assessment tool remains uncertain. Most experts agree that the combination of 5-m gait speed, grip strength, unintentional weight loss, inactivity, and exhaustion represent the most validated frailty measurements. Disability, defined as the inability or dependency to carry out activities of daily living and/or managing one's medications or finances (instrumental activities of daily living), is also an essential part of the initial geriatric evaluation. Assessment of baseline and post-procedure cognitive impairment with clinically established scales such as the MiniMental State Examination, the modified Telephone Interview of Cognitive Status, and/or the Clinical Dementia Rating Scale should also be considered when evaluating the utility and cost-effectiveness of invasive procedures among an elderly population. Involvement of experts in neurocognitive dysfunction (e.g., geriatricians, neurologists) with serial evaluations pre- and post-procedure is essential for meaningful appraisal; however, it should be acknowledged that evaluation and collection of frailty parameters can be time consuming and resource intensive. Further research is required to determine the extent to which frailty assessments should be a routine part of MV clinical trials and registries.

### Primary and secondary endpoints

#### General considerations

Selection of the primary and major secondary clinical endpoints should afford an assessment of new technologies that is meaningful both for regulatory considerations and to guide clinical use. Such endpoints should give a robust determination of the benefit-risk balance afforded to patients receiving the therapy relative to other therapeutic options. Separate pre-specified primary powered safety and effectiveness endpoints are recommended for most trials of new transcatheter MV therapies, and the use of a single composite clinical safety and effectiveness endpoint, especially when the individual components of safety and efficacy may move in opposite directions, is not recommended. Although a single primary endpoint (or set of endpoints) cannot be pre-determined in this document for all possible clinical trials of MR devices and therapies, general principles may be proposed:

- Major endpoints should address both the safety and effectiveness of the proposed new device. In general, separate safety and effectiveness measures are desirable as coprimary endpoints, and both safety and effectiveness hypotheses should typically be...
Clinical trial design principles

19

Table 7  Recommended major inclusion and exclusion criteria for transcatheter device trials in patients with mitral regurgitation

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ≥18 years</td>
<td>Patient (or legal guardian) unable or unwilling to provide written, informed consent before study enrolment</td>
</tr>
<tr>
<td>Degree of MR: Severe (or 3+ and 4+)†</td>
<td>Need for emergent or urgent surgery for any reason or any planned cardiac surgery within the next 12 months</td>
</tr>
<tr>
<td>LVEF &gt; 20% (primary MR) or ≥20% to #60% (secondary MR)‡</td>
<td>Absence of CRT with Class I indication criteria for biventricular pacing</td>
</tr>
<tr>
<td>Symptom status: NYHA functional class II to IV‡</td>
<td>Implant or revision of any rhythm management device (CRT or CRT-D) or implantable cardioverter-defibrillator within 1 month before randomization</td>
</tr>
<tr>
<td>Treatment and compliance with optimal guideline-directed medical therapy for heart failure for at least 30 days (preferably 90 days)</td>
<td>Untreated clinically significant coronary artery disease requiring revascularization</td>
</tr>
<tr>
<td>MR mechanism/anatomy: Appropriate to the design specifications of each device</td>
<td>Any percutaneous cardiovascular intervention, cardiovascular surgery, or carotid surgery within 30 days</td>
</tr>
<tr>
<td>Surgical risk: Specific STS risk score criteria and/or the presence of high-risk features or comorbidities, depending on the specific trial aims</td>
<td>Tricuspid valve disease requiring surgery or severe tricuspid regurgitation</td>
</tr>
<tr>
<td>Completion of required functional tests (e.g., 6-min walk) and/or quality-of-life assessments</td>
<td>Aortic valve disease requiring surgery</td>
</tr>
</tbody>
</table>

Note:

†Ideally as assessed by an independent echocardiographic core laboratory. Different quantitative criteria may apply for primary and secondary MR. See Assessment of Mitral regurgitation: Role of Non-invasive Imaging. †As a starting point for consideration. The upper limit of LVEF should be selected to ensure inclusion of patients with true secondary MR due to LV dysfunction. By unloading the LV, severe MR increases the LVEF, and LVEF #60% is consistent with LV dysfunction; however, lowering the upper limit of the LVEF range (e.g., to #50%) may be considered to increase specificity. Similarly, the lower level of LVEF should be selected to ensure exclusion of patients who might not be capable of benefitting from MR reduction. An acute increase in afterload by reducing or eliminating MR may also (rarely) result in hemodynamic compromise in the early post-intervention period. In general, a lower limit LVEF of 20% is recommended. Lower and upper limits for LV dimensions should also be considered on the basis of the specific device being tested. †As a starting point for consideration. Patients should be symptomatic, and most patients should be ambulatory (able to complete a 6-min walk test). Dedicated trials, however, may be designed for asymptomatic or end-stage patients. †Assessed by echocardiography or right heart catheterization, unless active vasodilator therapy in the catheterization laboratory is able to reduce the pulmonary vascular resistance to <3 Wood Units or between 3 and 4.5 Wood Units, with v-wave less than twice the mean of the pulmonary capillary wedge pressure.

ACC, American College of Cardiology; AHA, American Heart Association; CRT, cardiac resynchronization therapy; CRT-D, cardiac resynchronization therapy defibrillator; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association; STS, Society of Thoracic Surgeons; TAVR, transcatheter aortic valve; UNOS, United Network for Organ Sharing.
Table 8  Risk assessment in valvular heart disease, combining society of thoracic surgery risk estimates, frailty, major organ system dysfunction, and procedure-specific impediments for intervention

<table>
<thead>
<tr>
<th></th>
<th>Low risk (all criteria in this column must be present)</th>
<th>Intermediate risk (at least 1 criterion in this column must be present)</th>
<th>High risk (at least 1 criterion in this column must be present)</th>
<th>Prohibitive risk (any 1 criterion in this column must be present)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STS PROM</strong>†‡</td>
<td>&lt;4%</td>
<td>4%-8%</td>
<td>&gt;8%</td>
<td>Predicted risk with surgery of death or major morbidity (all-cause)</td>
</tr>
<tr>
<td><strong>Frailty</strong></td>
<td>None</td>
<td>1 index (mild)</td>
<td>≥2 indexes (moderate to severe)</td>
<td>&gt;50% at 1 year</td>
</tr>
<tr>
<td><strong>Major organ system compromise not to be improved post-operatively</strong>§</td>
<td>None</td>
<td>1 organ system</td>
<td>No more than 2 organ systems</td>
<td>≥3 organ systems</td>
</tr>
<tr>
<td><strong>Procedure-specific impediment</strong>§</td>
<td>None</td>
<td>Possible procedure-specific impediment</td>
<td>Possible procedure-specific impediment</td>
<td>Severe procedure-specific impediment</td>
</tr>
</tbody>
</table>

*Use of the STS predicted risk of mortality (PROM) to predict risk in a given institution with reasonable reliability is appropriate only if institutional outcomes are within 1 SD of STS average observed/expected ratio for the procedure in question.
†Seven frailty indexes: Katz Activities of Daily Living (independence in feeding, bathing, dressing, transferring, toileting and urinary continence) and independence in ambulation (no walking aid or assist required for 5 m walk in <6 s). Other scoring systems can be applied to calculate no, mild, or moderate-to-severe frailty. Examples of major organ system compromise: Cardiac: severe LV systolic or diastolic dysfunction or RV dysfunction, or fixed pulmonary hypertension; CKD stage 3 or worse; pulmonary dysfunction with FEV1 <50% or DLCO <50% of predicted; CNS dysfunction: dementia, Alzheimer’s disease, Parkinson’s disease, or CVA with persistent physical limitation; GI dysfunction: Crohn’s disease, ulcerative colitis, nutritional impairment, or serum albumin <3.0; cancer: active malignancy; and liver: any history of cirrhosis, variceal bleeding, or elevated INR in the absence of VKA therapy. §Examples: tracheostomy present, heavily calcified ascending aorta, chest malformation, arterial coronary graft adherent to posterior chest wall, or radiation damage. Adapted with permission from Nishimura et al. CKD, chronic kidney disease; CNS, central nervous system; CVA, cerebrovascular accident (stroke); DLCO, diffusion capacity for carbon dioxide; FEV1, forced expiratory volume in 1 s; GI, gastrointestinal; INR, international normalized ratio; LV, left ventricular; PROM, predicted risk of mortality; RV, right ventricular; STS, society of thoracic surgeons; VKA, vitamin K antagonist.

- Endpoints should be selected for their ability to measure both clinical events and meaningful improvements in clinical outcomes, and be measured at relevant intervals that are appropriate for the disease state and the intervention. The primary and secondary endpoints should be selected such that meeting these endpoints will demonstrate reasonable assurance of safety and effectiveness and a favourable benefit-risk profile.
- All endpoints should be well defined and linked to meaningful clinical outcomes (e.g., reduction in all-cause mortality, reduction in hospitalization, improvement in quality of life, and improvement in functional status).
- Additional secondary endpoints may include functional measures or surrogate measures of the primary endpoints studied, rather than simply a technical assessment of device success.
- The observed benefit was due to the device intervention, for life-threatening and life-limiting conditions.
- Additional secondary endpoints may include functional measures or surrogate measures of the primary endpoints studied, rather than simply a technical assessment of device success.
- The primary effectiveness endpoint should be a relevant clinical outcome for the population studied, rather than a technical assessment of device success.
- The primary and secondary safety endpoints should be selected such that meeting these endpoints will demonstrate reasonable assurance of safety and effectiveness and a favourable benefit-risk profile.
- The primary and secondary safety endpoints should be selected such that meeting these endpoints will demonstrate reasonable assurance of safety and effectiveness and a favourable benefit-risk profile.
- The primary and secondary safety endpoints should be selected such that meeting these endpoints will demonstrate reasonable assurance of safety and effectiveness and a favourable benefit-risk profile.
- The primary and secondary safety endpoints should be selected such that meeting these endpoints will demonstrate reasonable assurance of safety and effectiveness and a favourable benefit-risk profile.
related to effectiveness and/or device safety that may be observed during the relevant period of observation. The individual components of the composite endpoint should share a common pathophysiology or represent specific major complications of device therapy, and should be expected to trend in the same direction. Major and minor events should be clearly distinguished to avoid grouping outcomes of variable clinical significance.

- The statistical analysis of these endpoints should conform to commonly accepted principles, such as accounting for competing risks and multiplicity (see also Statistical Considerations).

Selection of appropriate primary and secondary endpoints to assess device therapies for MR is especially challenging, because of a complicated matrix that includes the underlying risk and comorbidities of the target population, the specific pathogenic mechanisms of MV dysfunction (e.g., primary vs. secondary and ischaemic vs. non-ischaemic aetiologies), and whether options for therapy include MV repair or replacement surgery, coronary revascularization, and ventricular resynchronization therapy, in addition to GDMT. For example, whether mortality is the optimal stand-alone primary endpoint may depend on the expected survival rate of the target population with currently available treatment options. For patient cohorts in whom short- or intermediate-term mortality is low, other measures, such as outcomes related to heart failure and functional capacity, may be more clinically meaningful during the course of a clinical trial observation if associated with continued successful device performance.

For studies in which MV surgery is the control group (e.g., for primary MR in acceptable surgical candidates), major endpoints must assess the safety of the new device relative to the complications of surgery, and demonstrating superiority in safety with non-inferiority in efficacy (within a reasonable margin reflecting therapeutic interchangeability) is a reasonable goal. Conversely, if the control group is GDMT with or without coronary revascularization and/or CRT as appropriate (e.g., for secondary MR in heart failure), device safety endpoints must be weighed against superiority measures of clinical efficacy that would justify the associated procedural and device-related risks. Whether or not pre-specified endpoints are met, FDA approval is based upon the totality of the data when considered as part of a thorough benefit-risk determination.20

**Primary endpoints**

Clinically meaningful effectiveness measures to be considered for MR device therapies are presented in Table 9. All-cause mortality should be incorporated into the primary efficacy endpoint (as either a standalone measure or as part of a composite) if there is a reasonable expectation that MR reduction might improve survival (e.g., for primary MR). For clinical trials in which the mortality rate during the time-course of observation is expected to be high, all-cause mortality as a pre-specified and adequately powered standalone primary efficacy endpoint should strongly be considered, as a significant improvement in survival is likely to support a favourable benefit-risk determination even if other device-related safety issues emerge. It is also the endpoint least affected by lack of blinding. Otherwise, all-cause or cardiac mortality may be part of a carefully constructed primary composite clinical effectiveness endpoint in which the pathophysiology between the components is shared and all events are considered clinically important. In this regard, after all-cause mortality, heart failure-related outcomes may be the best clinical measure of effectiveness to use in trials of MR therapies. As outlined in part 2 of this document, specific criteria defining hospitalization or hospitalization equivalents for heart failure may be crafted to allow independent adjudication of this event. Hospitalization for heart failure may serve as a primary standalone endpoint for conditions in which MR reduction is anticipated to improve quality-of-life, but may not necessarily improve survival (e.g., secondary MR in patients with severe LV dysfunction).

Whether functional measures (e.g., 6-min walk distance) or quality-of-life indexes are sufficiently clinically important and robust to warrant incorporation into a primary effectiveness endpoint is controversial. This consideration is especially relevant in unblinded trials, in which placebo and Hawthorne effects as well as assessment bias may make interpretation of these measures difficult. Proposed measures of functional performance are presented in Table 9. Such measures might be regarded as having intermediate value as reviewed by Temple.88 because they are meaningful to patients and may eventually result in a favourable effect on survival. If functional measures or quality-of-life outcomes are considered as a primary endpoint (either standalone or as part of a composite endpoint), all possible efforts for blinding should be used (including assessment of the success of the blinding procedures), and the data should demonstrate continued evidence of device performance at the

<table>
<thead>
<tr>
<th>Table 9</th>
<th>Clinical and functional outcome measures that may be considered for primary or secondary effectiveness endpoints in mitral regurgitation trials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endpoint</strong></td>
<td><strong>Primary or Secondary</strong></td>
</tr>
<tr>
<td>Clinical measures</td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td></td>
</tr>
<tr>
<td>All-cause</td>
<td>Primary or secondary&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Procedure-related</td>
<td></td>
</tr>
<tr>
<td>Cardiac&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Primary or secondary&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Heart failure rehospitalization</td>
<td>Primary or secondary&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mitral valve reintervention</td>
<td>Secondary&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Need for LVAD or heart transplant</td>
<td>Secondary&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Functional measures</td>
<td></td>
</tr>
<tr>
<td>6-min walk distance</td>
<td>Usually secondary</td>
</tr>
<tr>
<td>Cardiopulmonary exercise testing</td>
<td>Usually secondary</td>
</tr>
<tr>
<td>Validated quality-of-life scales</td>
<td>Usually secondary</td>
</tr>
<tr>
<td>Change in New York Heart Association functional class</td>
<td>Secondary</td>
</tr>
</tbody>
</table>

<sup>a</sup>Or part of a primary composite endpoint. <sup>b</sup>In general, all-cause mortality is preferred to cardiac mortality as a primary endpoint. In studies enrolling patients with numerous noncardiac comorbidities that may result in a high rate of noncardiac mortality, accurately adjudicating the cause of death may be difficult. If cardiac mortality is used as a primary endpoint, a neutral effect on noncardiac mortality with the intervention should be present (accounting for competing risks). LVAD, left ventricular assist device.
time of assessment and reasonable confidence of lack of harm, including mortality.

The primary safety endpoint is usually a composite endpoint specific to the device and underlying cardiac condition, and should incorporate the need for unplanned MV surgery (or reoperation) due to progressive or recurrent MR or device-related complications.

As an example, these principles have been incorporated into the design of the COAPT trial, an ongoing prospective, multicenter randomized trial performed under an FDA Investigational Device Exemption in which the MitraClip is being compared to GDMT in patients with symptomatic severe secondary MR in whom surgery is not considered appropriate after local heart team evaluation. The primary effectiveness endpoint is hospitalization for heart failure during follow-up (measured by the Andersen-Gill test to take into account the number of heart failure hospitalizations), powered to demonstrate superiority of the MitraClip. The primary safety endpoint is the composite of single leaflet device attachments, device embolization, endocarditis requiring surgery, core laboratory-confirmed mitral stenosis requiring surgery, LV assist device implant, heart transplant, or any device-related complications requiring non-elective cardiovascular surgery at 12 months, to a pre-specified performance goal.

**Secondary endpoints**

Secondary endpoints should include the individual components of the primary endpoint if a composite was used, as well as other measures of effectiveness that were not a part of the primary endpoint, including functional, symptom, and quality-of-life assessments; imaging-related measures of MR reduction and cardiac structure and performance; and major and minor safety outcomes (Table 10). Secondary endpoints of interest that may be affected by the intervention should be pre-specified, but the study may or may not be adequately powered to demonstrate statistical significance for such endpoints even if differences truly exist. Powered secondary endpoints with pre-specified statistical hypotheses are necessary to make labelling claims for approved medical devices in the United States. If not powered, secondary endpoints are considered hypothesis generating, even if pre-specified. Nonetheless, secondary mechanistic endpoints may provide valuable guidance for new device designs or iterations, especially in this early era of MR device development.

Endpoints should be classified according to device- and procedure-relatedness and timing of occurrence as previously proposed: acute, within 24 h; early, after 24 h but within 30 days; intermediate, after 30 days but within 1 year; late, between 1 and 5 years; and very late, after 5 years. For secondary outcome measures (as for the primary endpoints), safety should be assessed separately from effectiveness, except possibly for all-cause mortality and stroke, endpoints that reflect both safety and effectiveness of an intervention. Secondary safety endpoints should further evaluate procedural and/or device-related complications and assess any adverse effects of the intervention on the disease state, future treatments, and prognosis (e.g., whether device failure impedes the likelihood to perform successful surgical MV repair). If not already identified as a standalone primary endpoint, all-cause mortality should always be pre-specified as a secondary endpoint (even if not adequately powered) to consider whether a new therapy might result in increased or decreased survival.

Various measures of mechanistic and technical success are important to include as pre-specified secondary outcomes in MR device trials. Although quantitative reduction in MR is the *sine qua non* for treatment effectiveness, and the associated physiological

---

**Table 10** Major safety, technical, and mechanistic endpoints in mitral regurgitation trials

<table>
<thead>
<tr>
<th>Major safety endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device or procedure-related adverse events (specific to each device and procedure)</td>
</tr>
<tr>
<td>Major bleeding complications (transfusion reported separately)</td>
</tr>
<tr>
<td>Major vascular complications</td>
</tr>
<tr>
<td>Pulmonary complications (device or procedure-related)</td>
</tr>
<tr>
<td>Stroke and other cerebrovascular events (assessed by a stroke neurologist and CT/CMR imaging; disabling and nondisabling; change in modified Rankin score)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
</tr>
<tr>
<td>Acute kidney injury or progression of chronic kidney disease (dialysis reported separately)</td>
</tr>
<tr>
<td>New onset atrial fibrillation</td>
</tr>
<tr>
<td>Unplanned mitral valve surgery due to device/procedure failure or malfunction</td>
</tr>
<tr>
<td>Requirement for valve replacement versus repair</td>
</tr>
<tr>
<td>Unplanned cardiac surgery for any cause</td>
</tr>
<tr>
<td>Requirement/insertion of an implantable cardiac defibrillator</td>
</tr>
<tr>
<td>Requirement/insertion of biventricular pacemaker for cardiac resynchronization therapy</td>
</tr>
<tr>
<td>Device failure resulting in the inability to perform successful surgical mitral valve repair</td>
</tr>
</tbody>
</table>

**Technical success**

- Device success (specific definition)
- Implant rate
- Device time and procedure duration
- Contrast utilization
- Ionizing radiation exposure
- Procedural success (specific definition)
- Mechanistic endpoints
  - Imaging measures
  - Mitral regurgitation severity (integrated assessment; see text and Tables 2 and 3)
  - Mitral valve area and mean gradient
  - Left atrial and pulmonary artery pressures
  - End-systolic dimension and volume
  - End-diastolic dimension and volume
  - Left ventricular sphericity
  - Left ventricular ejection fraction
  - Left atrial dimension and volume
  - Right ventricular pressures, dimension, volume, and ejection fraction
  - BNP and/or NT-pro BNP levels

**Mechanistic endpoints**

- Absolute levels and incremental change from baseline.
- BNP, B-type natriuretic peptide; CMR, cardiac magnetic resonance; CT, computed tomography; NT-proBNP, N-terminal pro-B-type natriuretic peptide.
measures (e.g., reduced LA and LV volumes and PAP, improved LVEF) are consistent with successful device performance over time, these surrogate endpoints are insufficient to serve as primary effectiveness endpoints (either standalone or as a component of a composite measure) because they may not be associated with clinically meaningful improvements. Furthermore, small statistically significant differences in continuous measures may not result in clinically measurable benefits. Nevertheless, a high rate of early technical success, coupled with continued device functional performance and a beneficial physiological response over time, should be present to support the biological plausibility of the primary clinical effectiveness endpoint. These measures should, therefore, be evaluated in all MR trials and reported as secondary efficacy outcomes. MR severity over time is also an important measure of durability of the treatment effect and should be assessed at regular intervals throughout the study duration. All echocardiographic measures should be evaluated by a central core laboratory to standardize reporting and limit potential bias.

Analysis of primary and secondary endpoints

It is recommended that primary and secondary endpoints be assessed at the intervals noted in Table 11. In most MR trials, the primary effectiveness endpoint should be assessed no sooner than 1 year after randomization, whereas the primary safety endpoint may be assessed as soon as 30 days after randomization to account for procedural complications (each taking into account between group differences in time from randomization to treatment initiation). Depending on the device, however, follow-up longer than 30 days may be appropriate for the primary safety endpoint assessment.

The analysis plan should incorporate achievement of device- and patient-oriented outcomes to assess overall safety and effectiveness as well as benefit-risk. As detailed in part 2 of this document, device-oriented outcomes include technical success with associated mechanistic outcomes and device- and procedure-related safety endpoints. Patient-oriented outcomes include the components of the primary safety and effectiveness endpoints. A hierarchical analysis plan should be pre-specified beginning with assessment of the primary safety and effectiveness endpoints followed by analysis of powered major secondary endpoints assessing functional and mechanistic outcomes, with attention to preserving type I error (alpha) at the 0.05 level. Thereafter, non-powered secondary endpoints are assessed, as well as subgroup analyses to examine consistency, although these exploratory analyses offer lower levels of evidence and are considered hypothesis-generating in most circumstances.

Finally, it should be noted that over time the level of evidence required for serial iterations of MR devices or even novel devices will likely change as experience grows with therapies for treating MR patients.

Role of the central eligibility committee

Even with the use of local heart teams and detailed protocol inclusion/exclusion criteria, site-based variability in patient selection for studies of devices treating MR remains a concern. To enhance interpretability, particularly for regulatory trials of novel MR devices, it is strongly recommended that each patient be presented to a central eligibility committee for evaluation of patient appropriateness for enrolment. The central eligibility committee serves multiple important functions (Table 12). The members of the central eligibility committee for MR trials should include at a minimum a moderator, a heart failure specialist, and an experienced MV surgeon. Each patient should be presented to the committee by the local principal investigator and, depending on the specific requirements for the trial, a cardiologist with expertise in valvular heart disease, a heart failure specialist, and/or an MV cardiac surgeon. The MV surgeon plays a central role in all MR randomized trials, whether the control arm is surgery (as for most trials of primary MR) or GDMT (as for most trials of secondary MR). Before the committee meeting, a central echocardiography core laboratory should have reviewed the qualifying TTE (+ TEE) to ensure that the MR meets severity criteria and, depending on the trial, to confirm that other eligibility criteria are met (e.g., LV volumes, LVEF, MR aetiology and anatomy, absence of mitral stenosis, and so on). Depending on the specific device, the central echocardiographic core laboratory may also be asked to determine whether appropriate anatomy is present for device eligibility. Although implementation of a central eligibility committee and pre-review by a central echocardiographic core laboratory entail extra time and cost, ensuring that only appropriate patients are enrolled in the clinical trial will substantially increase the power of the study and the likelihood of success.

An important distinction must be made between the roles of the local heart team and the central eligibility committee. Both multidisciplinary groups possess expertise to evaluate surgical risk and clinical trial eligibility. However, the thresholds at which MV surgery might be considered reasonable vary from centre to centre and surgeon to surgeon, depending on local experiences and volume considerations. The subtleties of the patient’s clinical condition (and surgical risk) are also best assessed by those able to speak to and examine the patient. Thus, the local heart team determines the relative surgical risk and operability of a patient, rather than the central eligibility committee. The MV surgeon (and others) on the central eligibility committee may, however, query the local surgeon as to his/her criteria for operability to ensure, for example, that crossover to surgery will not be considered should the patient be randomized to a nonsurgical therapy (unless permitted by the protocol).

Table 11 Timing of endpoint assessment (follow-up intervals)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute</td>
<td>(during procedure or within 24 h)</td>
</tr>
<tr>
<td>Procedural</td>
<td>(30 days post-procedure or until discharge from hospital or acute care facility)</td>
</tr>
<tr>
<td>90 days</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td></td>
</tr>
<tr>
<td>Annual</td>
<td>(for a minimum of 5 years)</td>
</tr>
</tbody>
</table>
Trial endpoints and analysis

Endpoints fall into 3 categories: (1) early post-procedure events and measures (e.g., death, stroke, valvular regurgitation); (2) time-related events (e.g., death, stroke); and (3) periodically sampled longitudinal data (e.g., valvular regurgitation, chamber dimensions, NYHA functional class). Early events are compared as odds ratios, early measurements are compared as differences, time-related events are compared as actuarial curves and hazard ratios, and longitudinal data are compared as time-related differences in ensemble averages. Although each trial must carefully evaluate and adopt the statistical methodology most appropriate for its goals, the following general principles may be useful to consider.

Nonfatal time-related events
Nonfatal events can repeat (e.g., stroke, rehospitalization). All occurrences should be analysed, not just time to first occurrence, using the Nelson, Andersen-Gill, or other estimators. These methods make different assumptions with respect to the independence of events, hazard function after each occurrence, and informativeness of death and other competing risks.

Weighted events
Although generally not done, consideration may be given to weighing nonfatal events (e.g., by applying the National Institutes of Health stroke scale and considering the duration and cost of rehospitalization), which may be further analysed as cumulative functions, a common industrial method when considering costs.

Longitudinal data
Longitudinal data reflect an endpoint’s state at time of assessment; they are not time-to-event data. Examples are drug use (binary), functional status (ordinal), and EROA (continuous). Such endpoints should be analysed by longitudinal repeated measures methods.
Composite endpoints
The use of composite endpoints to reduce sample size is a practical convention, but if not carefully constructed, may lead to difficulties in interpretation. Typically, each component is equally weighted, although the hazard function for each may be different (e.g., the Andersen-Gill method assumes proportional hazards [91]). However, clinical hierarchy or patient preference for each component may differ. If the components can be hierarchically arrayed, tree-structured gatekeeping tests or pair-wise winner-loser strategies may be used. Family-wise tests of individual components emphasize consistent direction of effect. The most controversial composite endpoints combine disparate component categories. Several groups, such as Finkelstein and Schoenfeld, have developed methods that combine time-to-event components with periodic longitudinal assessments. Others have extended this to continuous longitudinal data combined with weighted time-to-event data. On the horizon are joint models that account simultaneously for different intensity functions of each event and longitudinal components and their interrelations.

Conclusions
In contrast to calcific aortic stenosis, a relatively simple disease with limited aetiologies and a straightforward pathophysiology, MR is a more complicated entity, due to the greater complexity of the MV structure and the numerous lesions and mechanisms that may lead to its failure. Central to the understanding of MR is recognizing that MR is indeed 2 disorders: 1 of the valve apparatus itself and 1 of the LV (or LA), which secondarily disrupts normal MV function. Continuing the analogy, developing effective therapies (and surgical approaches) for MR and demonstrating their safety and effectiveness in clinical trials is much more challenging than for aortic stenosis, and requires the intimate collaboration between physician-scientists across numerous disciplines, clinical trialists, statisticians, and industry and regulatory authorities. Although each device trial will entail its own nuanced considerations, adopting the principles espoused in this document as a template for clinical investigation of mitral therapeutics should allow sponsors and investigators to avoid the most common errors that can render interpretation of their findings problematic.

Appendix
For complete information on the MVARC members and participants, please see the online version of this article.

References
84. McHorney CA, Ware JE Jr., Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247–263.
103. Castaneda J, Gerritsen B. Appraisal of several methods to model time to multiple events per subject: modeling time to hospitalizations and death. Revista Colombiana de Estadistica 2010;43:61.