Use of Cardiac Magnetic Resonance and Echocardiography in Population-Based Studies
Why, Where, and When?

Thomas H. Marwick, MBBS, PhD, MPH; Stefan Neubauer, MD, FRCP, FMedSci; Steffen E. Petersen, MD, DPhil

Population research studies are often directed toward eliciting the association of physiological measurements (eg, left ventricular [LV] function) and clinical variables (eg, hypertension) with outcomes. Although cardiovascular magnetic resonance (CMR) is known to be accurate and versatile, until recently, the diffusion of CMR technology was too limited by technical and logistic challenges to consider its use in population studies on a large scale (>10,000). However, with recent technical developments, CMR has reached a level of maturity and ease of use, which makes its use in large-scale population studies a practical reality for the first time. The goal of this review is to facilitate the process of selecting imaging methods for population research studies based on design requirements and existing experience with the techniques.

Design Considerations
There are 4 fundamental aspects of imaging that are pertinent in population studies: validity, feasibility, accuracy, and reproducibility.

Validity
Selection bias is a critical issue in population studies. Inappropriate patient selection may lead to problems in extrapolating sample information even to the population from which the sample was derived, which is a core task of a population-based study. External validity may also be limited by selection. In contrast to this, greater variance in imaging measurements may require larger numbers, as discussed in the later section on accuracy and validity, but may have relatively less importance in population studies than in clinical trials. Patient selection is, therefore, critical to external validity. This has been considered carefully in the echocardiographic literature. For example, before the incorporation of 2-dimensional (2D) measurements of LV mass, in addition to M mode, the feasibility of LV mass measurement was somewhat limited, for example, in the Framingham Heart Study and Cardiovascular Health Study. In contrast, more recent studies, such as the Strong Heart Study and Losartan Intervention For Endpoint Reduction in Hypertension (LIFE) study, have had a feasibility of ≈95% in most age groups. Age, sex, body mass index, and pulmonary disease are weakly associated with feasibility of measuring LV mass, but these patients did not have a different outcome compared with the majority in whom LV mass was measurable. The impact of CMR access and absence of contraindications (implanted devices, claustrophobia, renal impairment if contrast used) on the ability to gather an unselected population should be considered (and if possible, measured) in population-based CMR studies. Similarly, the prevalence of obesity and, especially, lung disease on acoustic windows should be considered for studies performed with echocardiography. The relative roles of these features are likely to differ in different subpopulations.

In a population-based study, a representative sample of a defined population is selected for longitudinal assessment of exposure–outcome associations. For example, the UK Biobank gathered data from 500,000 of the 23.5 million people aged 40 to 69 years on the English National Health Service Registry between 2006 and 2010. Many such population-based studies are part of a comprehensive evaluation that includes storage of biological samples, surveys and questionnaires, many physical measures, as well as outcomes. The combination of these features provides a resource for the detection of generalizable associations between characteristics recorded at baseline and health outcomes during follow-up. The value and cost-effectiveness of such population-based studies increase over time as outcomes accrue and more enhancement measures are done.

Feasibility
In these studies, the balance between data breadth and data depth is important. In some epidemiological studies, data acquisition is not limited to cardiovascular disease. If this is the case, cardiovascular imaging is but one component of an extensive evaluation and may be restricted to a limited time (eg, 20 minutes) because other imaging and other testing are being performed. This is sufficient for imaging to address most cardiovascular questions (eg, LV structure and function, right ventricular [RV] function and pulmonary pressure, valvular disease, aortic size, and vascular function), but not all of them. In the situation where imaging time is limited, the study design will, therefore, need to target specific questions.
Data acquisition and analysis should provide accurate measurements of the parameters of interest and should allow quality control, standardization, and reproducibility. Because of the size and complexity of large-scale population-based studies, participant safety and comfort, and feasibility are central considerations. It should be kept in mind that external validity is greatest when testing is achievable as close as possible to all subjects. Notwithstanding the safety of all modern methods for cardiovascular imaging, the safety of radiation exposure, medication, or contrast agents by healthy subjects is such that they are more likely to participate when the methods are noninvasive and do not use any radiation. For these reasons, chest radiographs, scintigraphy, positron emission tomography, and invasive coronary angiograms are not ideal because of their invasiveness and associated radiation exposure. Cardiac computed tomography with new state-of-the-art equipment often can be performed with low-dose radiation, and calcium score, which can be determined without contrast, has been used in many previous studies. Although other noncontrast targets include heart size, chest, visceral and subcutaneous fat, aortic size, and liver density, the most commonly desired measurements (LV size and function) cannot be acquired at low radiation dose and without contrast. Therefore, for the purposes of identifying multiple cardiac diseases, 2 noninvasive techniques are best suited for large-scale population studies: echocardiography and CMR.

CMR is an expensive and sophisticated methodology. Although many populations are potentially of interest, some are more amenable to CMR than others. Populations of patients attending outpatient clinics are highly suitable. Populations derived from the community, particularly in socioeconomically depressed and rural areas, pose significant access challenges for CMR, which may be partially but not completely addressed by mobile scanners. On a worldwide basis, much population research is currently being performed in the process of epidemiological transition, as developing countries develop disease burdens in degenerative and man-made diseases, rather than infectious disease, malnutrition, and more simple chronic disease, such as hypertension. Although mobile MR scanners have been used for on-site imaging in Europe, most of these environments in the developing world are unsuitable for CMR, both regarding the availability of the scanners and the infrastructure to support them.

### Accuracy and Validity

CMR has accuracy benefits relative to other tests, although not for all measurements. The important differences in the accuracy of CMR compared with 2D echocardiographic measurements of LV mass, volume, and ejection fraction (EF) are based on the fact that CMR is a volumetric technique with high contrast and spatial resolution and have been recognized for more than a decade. In the original landmark work by Bellenger et al., 20 patients with heart failure and 20 controls underwent CMR, and a comparison was made with the echocardiographic literature. A direct comparison between 2D echocardiography with the same patients was made in 2001 by Strohm et al., who showed an interstudy difference of EF of 24±18%, compared with only 17±19% with CMR. These variations seem drastic in the current era and may reflect the use of a former-generation echo machine. As one might expect, the differences were mostly marked between 2D echocardiography and CMR in the postinfarct population, where the role of CMR as a fundamentally 3-dimensional (3D) approach led to estimation of a lower EF (44±12% versus 51±8%), although this may have been accentuated by somewhat earlier performance of CMR (2.8±1.6 versus 3.4±1.7 days after infarction). In a recent meta-analysis, end-diastolic and end-systolic volumes were underestimated by 2D echocardiography by 48±25.5 and 27±45.7 mL, respectively, although the bias for EF was small (0.1±1.39%). Indeed, these differences in LV volumes have been markedly reduced in comparison with 3D echocardiography, an equivalent volumetric technique. A recent meta-analysis using this method showed that volume differences were reduced to 19.1±34.2 and 10.1±29.7 mL, with a small difference in EF (0.6±11.8%). To what extent the use of 3D imaging has improved the accuracy of echocardiographic examination of LV size and function is less clear. Ongoing concerns regarding 30% to 40% underestimation of RV volumes by 3D echocardiography are difficult to reconcile with a recent meta-analysis that reported a small (but still significant) underestimation of CMR-derived RV volumes by 3D echocardiographic methods in multiple recent human studies.

Although CMR can be used for assessing valvular and diastolic function, these nonetheless possess challenges related, in large part, to temporal resolution, and the test is not the first choice. The importance of these limitations will vary by context. If the study is being performed with an interest in accurate LV measurements, CMR may be the best option. If the questions relate to valvular regurgitation or diastolic dysfunction, the inclusion of echocardiography may be more attractive.

The ability to recognize different tissue types is a major attraction of CMR. Most of these applications currently require imaging after infusion of intravenous gadolinium-based contrast agents. Although this has been used in population studies, there are some disadvantages related to patient acceptability, cost, additional time (≥15 minutes with cost implication), uncertain impact on other MRI measures done after contrast (eg, brain MRI), additional incidental findings, and a small risk of serious adverse reaction. Thus, the use of contrast in a population study presents several considerations that may pertain to external validity of the data set and needs to be tailored carefully to the goals of the study. However, this may be changed in the future, because T1 mapping permits detection and quantification of the extent of several other tissue pathologies related to chronic myocardial injury or infiltra tion. Measurements of T2 characteristics can be performed without contrast agents, allowing the identification of myocardial edema and iron deposition.

### Reproducibility

Not only is CMR more reproducible than echocardiography, but CMR reproducibility data have also focused on interstudy reproducibility that assesses a combination of acquisition and analysis reproducibility, whereas the frequently reported inter- and intraobserver variability reported with echocardiography assesses the reproducibility of analysis. The availability of more accurate and reproducible measurements from CMR...
has an important impact on power calculations in clinical trials, such as randomized controlled trials testing new antihypertensive treatments. Study power is dependent on effect size, the arbitrary definition of significance level, and the square root of the number of patients and is inversely proportional to the variance of the measurement. Tests with a high level of variance for repeat samples have a low power to detect change, and this needs to be compensated by an increment of numbers. Consequently, in the setting of a clinical trial before and after measurements, where the randomization and selection process hold other variables to be equal between the populations, measurement error becomes the only source of variability other than the treatment effect (Table 1).

The assumptions in population studies are a little more nuanced, and the study design is critical. If the primary interest is to perform sequential imaging follow-up, then the superior test–retest reproducibility of CMR is desirable. Similarly, if the plan is to use a baseline measure to predict later events, CMR allows smaller sample size for the same power (or for given sample size, greater power) because the SD that determines sample size is determined by both interindividual variability and reproducibility. However, understanding the role of imaging relative to other influences on outcome is more difficult. In this setting, most factors that are affecting outcome are uncontrolled or may be not even known, there may be a lot of variance between individuals in several variables, and these other variables may have an important impact compared with the treatment or exposure effect. The effect size of many of the parameters is relatively small (most risk factors carry a relative risk of 1.5–2), and often their prevalence is low (eg, 10%). Thus, if the goal of a population study is to define the risk of an imaging finding relative to these clinical risk factors, a high reproducibility of the imaging test may have a small effect relative to the power requirement of defining associations of an uncommon factor carrying a limited risk burden (Table 2).

Thus, the selection of CMR imaging needs to take into account the question being asked. The study design that is most amenable to population use of CMR relates to when a physiological measurement is being studied (eg, EF or end-systolic volume) and when the interest pertains to how this changes over time. Such a study might include the evaluation of cardiotoxicity or remodeling. In contrast, the high reproducibility of CMR is less relevant to studies where the nonimaging determinants of an event are associated with between-subject variability, as the latter may be the main driver of sample size requirements. Table 3 illustrates 3 situations where the assessment of associations in cross-sectional studies was not influenced by differences in reproducibility between imaging modalities, such that studies with relatively minor differences in numbers between echo and CMR studies demonstrated essentially the same findings. These are examples of population health studies that require not only imaging measurements, but also an understanding of the interaction of risk factors, which have a relative risk in the range of 1.5 to 2.0. Thus, in a large study, where the prognostic role of a physiological signal from imaging (eg, EF) is sought relative to other variables, between-modality differences may have a minor role.

<table>
<thead>
<tr>
<th>Table 1. Differences Between Clinical Trials and Population Studies That May Be Important in Selecting Imaging Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Trial</td>
</tr>
<tr>
<td>Close to idealized experiment</td>
</tr>
<tr>
<td>Experimental units comparable (blinding)</td>
</tr>
<tr>
<td>Often months to years</td>
</tr>
<tr>
<td>Measurement error is the only source of variability other than Rx effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Power (%) to Detect Relative Risks for Associations of Dichotomous Risk Factors and Clinical Cardiovascular Disease Outcomes by Risk Factor Prevalence in the Multi-ethnic Study of Atherosclerosis Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence 5%</td>
</tr>
<tr>
<td>Prevalence 10%</td>
</tr>
<tr>
<td>Prevalence 20%</td>
</tr>
</tbody>
</table>

Even in a large study of 6500 patients, power to associate a risk factor of low prevalence with a small effect size (Relative risk<2) is inadequate.
between modalities—a CMR scanner is 4 to 10 times the cost of a standard high-quality 3D echocardiographic system, and operating costs are higher. In the presence of large numbers of patients being studied at a limited number of sites, this cost difference becomes less important, but if the study requires evaluation of a dispersed population at a large number of sites, the use of a larger number of less-expensive imaging equipment may make the difference between success and failure. With either modality, extensive training of a large number of technologists is a critical component. Observer expertise is important with both echocardiography and CMR. In a classic article, the limits of change of echocardiographic measurement using a 10% classification error were 20 mL for End-systolic volume and 8.5% for EF.45 These inter-reader differences with echocardiography relate to difficulties in tracing endocardial contours. The high contrast resolution of CMR minimizes these difficulties, with the consequence that variability is less with CMR—representative 95% confidence intervals for systolic volumes are 18 mL, with differences in EF of 9%. The intervals for novice readers were 26 mL and 15%.46 Variation with both methods may relate to inclusion/exclusion of papillary muscles and trabeculations, although this may be automated with CMR.47 The exact nature of the imaging requirement is critical. From the earliest days of CMR, this test has been shown to

### Table 3. Similarities Between Echocardiographic and CMR-Based Studies That Seek to Link Imaging With Clinical Findings in Population Studies

<table>
<thead>
<tr>
<th>Topic</th>
<th>Echocardiography</th>
<th>CMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remodeling and LV dysfunction</td>
<td>Lollipop study(^a) (n=441)</td>
<td>MES(^a) (n=1074)</td>
</tr>
<tr>
<td></td>
<td>Concentric left ventricular hypertrophy was independently associated with significantly worse systolic (P=0.02) and diastolic function (P&lt;0.001) and higher LV filling pressure (P=0.003) compared with subjects with normal LV geometry. Similar results were found for nonhypertrophic concentric remodeling</td>
<td>In men, a gradual decline in peak global circumferential strain was seen with increasing M/V ratio (P&lt;0.001). In women, strain was lower only in the fifth quintile of M/V ratio (P=0.1)</td>
</tr>
<tr>
<td>IGT and LV mass</td>
<td>Strong Heart(^a) (n=1343)</td>
<td>MES(^a) (n=588)</td>
</tr>
<tr>
<td></td>
<td>IGT was associated with increased LV mass in men (P=0.05) and women (P=0.002) and increased posterior wall thickness in men (P=0.002) and women (P=0.001)</td>
<td>IGT was associated with increased LV mass in women (P=0.001) and increased posterior wall thickness in men and women (both P=0.001)</td>
</tr>
<tr>
<td>Systolic BP and LV mass</td>
<td>Cardia study(^a) (n=5115)</td>
<td>MES study(^b) (n=4869)</td>
</tr>
<tr>
<td></td>
<td>(r=0.37–0.65) depending on race and sex</td>
<td>(r=0.46)</td>
</tr>
</tbody>
</table>

BP indicates blood pressure; CMR, cardiovascular magnetic resonance; IGT, impaired glucose tolerance; LV, left ventricular; MESA, The Multi-ethnic Study of Atherosclerosis; and M/V, mass/volume ratio.

### Table 4. Large-Scale Population-Based Studies (≥1000 Subjects) That Have Used or Are Planning to Use CMR

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Study Patients</th>
<th>CMR Patients</th>
<th>CMR Protocol</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK Biobank(^c)</td>
<td>General population: 40–69 y</td>
<td>500 000</td>
<td>100 000</td>
<td>M</td>
<td>Planning stage</td>
</tr>
<tr>
<td>Iceland MI (AGES substudy)(^d)</td>
<td>Age &gt;67 y</td>
<td>12 000</td>
<td>1000</td>
<td>S</td>
<td>Completed</td>
</tr>
<tr>
<td>Jackson Heart Study(^e)</td>
<td>35–84-year-old blacks in Jackson, Mississippi (one of the highest rates of CVD in the United States)</td>
<td>5 301</td>
<td>≈2 000</td>
<td>S</td>
<td>Ongoing</td>
</tr>
<tr>
<td>SHIP(^f)</td>
<td>General population: 20–79 y</td>
<td>9 000</td>
<td>4 000</td>
<td>S</td>
<td>Ongoing</td>
</tr>
<tr>
<td>Dallas Heart Study(^g,h)</td>
<td>Multietnic (54% black), age 18–65 y (imaging substudy 30–65 y)</td>
<td>6 101</td>
<td>2 971</td>
<td>S</td>
<td>Completed</td>
</tr>
<tr>
<td>FHS offspring study(^i)</td>
<td>&lt;70 in 1971, offspring of original FHS cohort</td>
<td>5 124</td>
<td>1 800</td>
<td>S</td>
<td>Completed</td>
</tr>
<tr>
<td>MESA(^j)</td>
<td>Asymptomatic participants of 4 ethnicities; age: 45–84 y</td>
<td>6 814</td>
<td>5 000</td>
<td>M</td>
<td>Ongoing</td>
</tr>
<tr>
<td>German National Cohort(^k)</td>
<td>General population: 20–79 y</td>
<td>2 000 000</td>
<td>40 000</td>
<td>M</td>
<td>Planning stage</td>
</tr>
</tbody>
</table>

The table is unlikely to be complete as some studies have not published data yet, but it demonstrates the increasing popularity of using cardiovascular imaging in large-scale studies and that CMR has been used successfully for this purpose. CMR indicates cardiovascular magnetic resonance; CVD, cardiovascular disease; FHS, Framingham Health Study; LGE, late gadolinium enhancement; LV, left ventricular; MESA, The Multi-ethnic Study of Atherosclerosis; RV, right ventricular; SSFP, steady state free precession; T, Tesla; and T2w, T2 weighted.
have a high accuracy for cardiac chamber measurements based on excellent spatial and contrast resolution that allows reproducible delineation of endo- and epicardial borders and the free, but standardized choice of imaging planes not limited by ultrasound windows. In contrast, the original echocardiographic technique (M mode) used in the Framingham Heart study had high temporal resolution but limited reproducibility. The development of 2D echocardiography and Doppler allowed more effective assessment of valvular disease and diastolic dysfunction, but problems with reproducibility persisted, largely because of variations in cut planes when imaging 3D structures in 2D. The traditional superiority of CMR in permitting complete coverage of the heart to enable avoidance of geometric assumptions has been blunted by the transition to 3D echocardiography, because it too avoids geometric assumptions. However, although meta-analyses attest to the fact that this has reduced the variability and improved accuracy of echocardiography, the experience of using this method in population-based studies is relatively new. There are still problems of (albeit smaller) underestimation of LV volumes, and LV mass calculations remain problematic because of the challenges of determining epicardial borders. Furthermore, because image quality problems may prevent acquisition of accurate cardiac data in patient groups with chronic obstructive airway disease or obesity, the acquisition of incomplete data may not be random. Nonetheless, in combination with the strength of echocardiography in valvular and diastolic dysfunction evaluation, the availability of 3D has enhanced the competitiveness of echocardiography relative to CMR.

Several limitations are common to both CMR and echocardiography. First, feasibility can be a problem with either method. Obesity and chronic pulmonary disease remain challenging for echocardiography. Although LV opacification can be used to ameliorate this problem, its use contravenes the common desire in population studies to avoid intravenous access or injections. However, CMR feasibility falls short of 100% as a result of claustrophobia and metallic implants. Second, the evolution of technology can pose important challenges to follow-up studies, as both different CMR sequences and different echocardiographic methods (M mode, 2D and 3D) may provide differences in temporal and spatial resolution. In the Multi-ethnic Study of Atherosclerosis study, sequential comparisons involved use of cine-segmented k-space gradient echo methods at baseline, with follow-up studies being performed with steady state free precession methods. Third, both methods are susceptible to variations between measurements at different sites, based on differences in equipment and different operators.

Conclusions

The low variance between multiple CMR measurements has made this technique the test of choice in the evaluation of patients in some clinical trials, in preference to alternative strategies for LV evaluation, including echocardiography. However, large-scale cardiovascular imaging in population-based studies requires different considerations to trials and clinical work. In population studies, there is often interest in the interactions between physiological measurements and environmental factors that have low prevalence and low relative risk, in which circumstance it is these factors that drive the required size of a population trial.

Thus, the use of cardiac MRI in population studies needs to take account of the exact question being asked, the impact on bias, the need for appropriate reading skill, and the setting of the patient. The best cardiovascular imaging modality will depend on the design, aims, and circumstances of the study. CMR is the reference method for LV and RV anatomy and function, and tissue characterization may be a major attraction of CMR. Echocardiography remains superior for valvular and hemodynamic evaluation. Sustainable high quality is probably more challenging with echocardiography compared with CMR. Finally, the issue of feasibility, based on access to equipment and, to a lesser extent, contraindications to testing, may be an important consideration in population studies. Barriers to scanning the entire population generate a source of potential bias, which may limit the external validity of study findings.

Acknowledgments

We thank the UK Biobank Imaging Working Group and the UK Biobank Cardiovascular MRI Advisory Group for invaluable contributions to this topic in numerous discussions.

Sources of Funding

Dr Petersen was directly funded by the National Institutes for Health Research Cardiovascular Biomedical Research Unit at Barts.

Disclosures

None.

References


**Key Words:** cardiac magnetic resonance • echocardiography • epidemiology • magnetic resonance imaging • population
Use of Cardiac Magnetic Resonance and Echocardiography in Population-Based Studies: Why, Where, and When?
Thomas H. Marwick, Stefan Neubauer and Steffen E. Petersen

Circ Cardiovasc Imaging. 2013;6:590-596
doi: 10.1161/CIRCIMAGING.113.000498
Circulation: Cardiovascular Imaging is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-9651. Online ISSN: 1942-0080

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circimaging.ahajournals.org/content/6/4/590

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Imaging can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Imaging is online at:
http://circimaging.ahajournals.org//subscriptions/