




Variation in left ventricular cardiac magnetic resonance normal reference ranges: systematic review and meta-analysis

Zahra Raisi-Estabragh ^{1,2*}, Asmaa A.M. Kenawy^{1,2}, Nay Aung^{1,2}, Jackie Cooper¹, Patricia B. Munroe¹, Nicholas C. Harvey^{3,4}, Steffen E. Petersen^{1,2}, and Mohammed Y. Khanji^{1,2}

¹William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK; ²Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, West Smithfield, London EC1A 7BE, UK; ³MRC Lifecourse Epidemiology Unit (MRCLEU), Tremona Rd, Southampton SO16 6YD, UK; and ⁴NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Tremona Road, Southampton, SO16 6YD, UK

Received 22 January 2020; editorial decision 8 April 2020; accepted 9 April 2020; online publish-ahead-of-print 27 May 2020

Aims

To determine population-related and technical sources of variation in cardiac magnetic resonance (CMR) reference ranges for left ventricular (LV) quantification through a formal systematic review and meta-analysis.

Methods and results

This study is registered with the International Prospective Register of Systematic Reviews (CRD42019147161). Relevant studies were identified through electronic searches and assessed by two independent reviewers based on predefined criteria. Fifteen studies comprising 2132 women and 1890 men aged 20–91 years are included in the analysis. Pooled LV reference ranges calculated using random effects meta-analysis with inverse variance weighting revealed significant differences by age, sex, and ethnicity. Men had larger LV volumes and higher LV mass than women [LV end-diastolic volume (mean difference = 6.1 mL/m², *P*-value = 0.014), LV end-systolic volume (MD = 4 mL/m², *P*-value = 0.033), LV mass (mean difference = 12 g/m², *P*-value = 7.8 × 10⁻³)]. Younger individuals had larger LV end-diastolic volumes than older ages (20–40 years vs. ≥65 years: women MD = 14.0 mL/m², men MD = 14.7 mL/m²). East Asians (Chinese, Korean, Singaporean-Chinese, *n* = 514) had lower LV mass than Caucasians (women: MD = 6.4 g/m², *P*-value = 0.016; men: MD = 9.8 g/m², *P*-value = 6.7 × 10⁻⁵). Between-study heterogeneity was high for all LV parameters despite stratification by population-related factors. Sensitivity analyses identified differences in contouring methodology, magnet strength, and post-processing software as potential sources of heterogeneity.

Conclusion

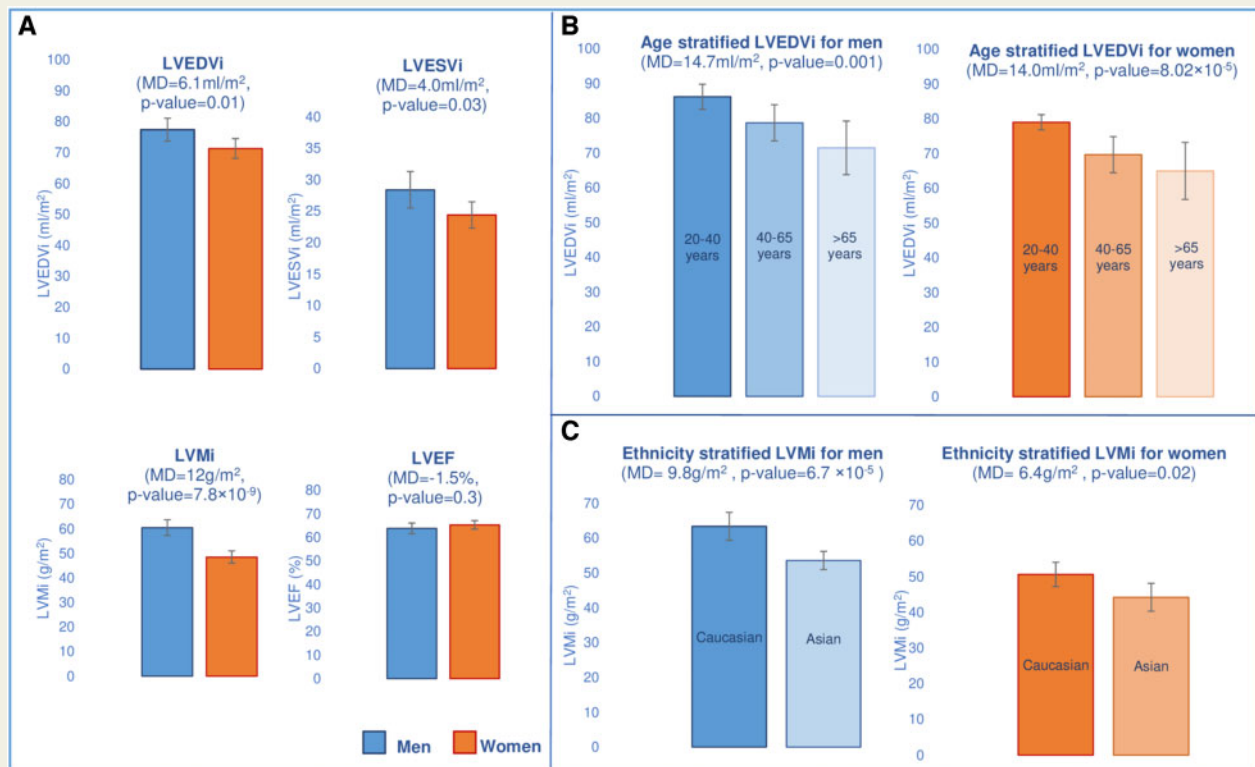
There is significant variation between CMR normal reference ranges due to multiple population-related and technical factors. Whilst there is need for population-stratified reference ranges, limited sample sizes and technical heterogeneity precludes derivation of meaningful unified ranges from existing reports. Wider representation of different populations and standardization of image analysis is urgently needed to establish such reference distributions.

* Corresponding author. Tel: +44 20 37658766. E-mail: zahraraisi@doctors.org.uk

© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical Abstract



Keywords

cardiac magnetic resonance • reference range • normal range • left ventricle

Introduction

Accurate quantification of left ventricular (LV) structure and function is key to clinical decision making in cardiology. LV cavity volumes in end-systole (LVESV) and end-diastole (LVEDV) reflect adverse myocardial remodelling.¹ LV mass (LVM), is an independent prognostic marker in individuals with and without cardiovascular disease.²⁻⁴ LV ejection fraction (LVEF) provides an estimate of LV systolic function and is the determinant of many important clinical decisions such as cardiac-resynchronization therapy, valve interventions, and management of heart failure syndromes.⁵⁻⁸

Cardiac magnetic resonance (CMR) imaging is the reference test for cardiac chamber quantification and is increasingly used to guide difficult clinical decisions. However, there is lack of consensus on normal reference ranges with variation in published reports.⁹ Whilst there are known sex, age, and ethnic differences in cardiac morphology,^{10,11} these differences have not been adequately studied with CMR and commonly quoted reference ranges are based on small cohorts that do not always represent the populations to which they are applied.

Previous attempts to pool results from different CMR reference ranges were limited by the datasets available at the time, with small sample sizes, inability to provide age and ethnicity stratification, or perform a formal meta-analysis.¹² In the last 5 years, there has been a

surge of publications reporting normal CMR reference ranges from around the world. The objective of this study is to determine population-related (sex, age, and ethnicity) and technical sources of heterogeneity through a formal systematic review and meta-analysis of published CMR reference ranges.

Methods

This study is registered online with the International Prospective Register of Systematic Reviews (PROSPERO, <https://www.crd.york.ac.uk/PROSPERO/>; registration number: CRD42019147161, 27 April 2020, date last accessed). Methods are in accordance with the PRISMA statement (Transparent Reporting of Systematic Reviews and Meta-Analyses, <http://prisma-statement.org/>, 27 April 2020, date last accessed). The PRISMA checklist is provided in the [Supplementary data](#) online.

Selection criteria

We selected studies that defined a normal reference range in healthy adults (>18 years-old) with sample sizes of ≥ 50 , reported in the English language. We required confirmation of healthy status of participants. We accepted studies with 1.5-T or 3-T scanners from all vendors. We restricted to studies using steady state free precession (SSFP) sequences, as this reflects current clinical standards for volume quantification. We

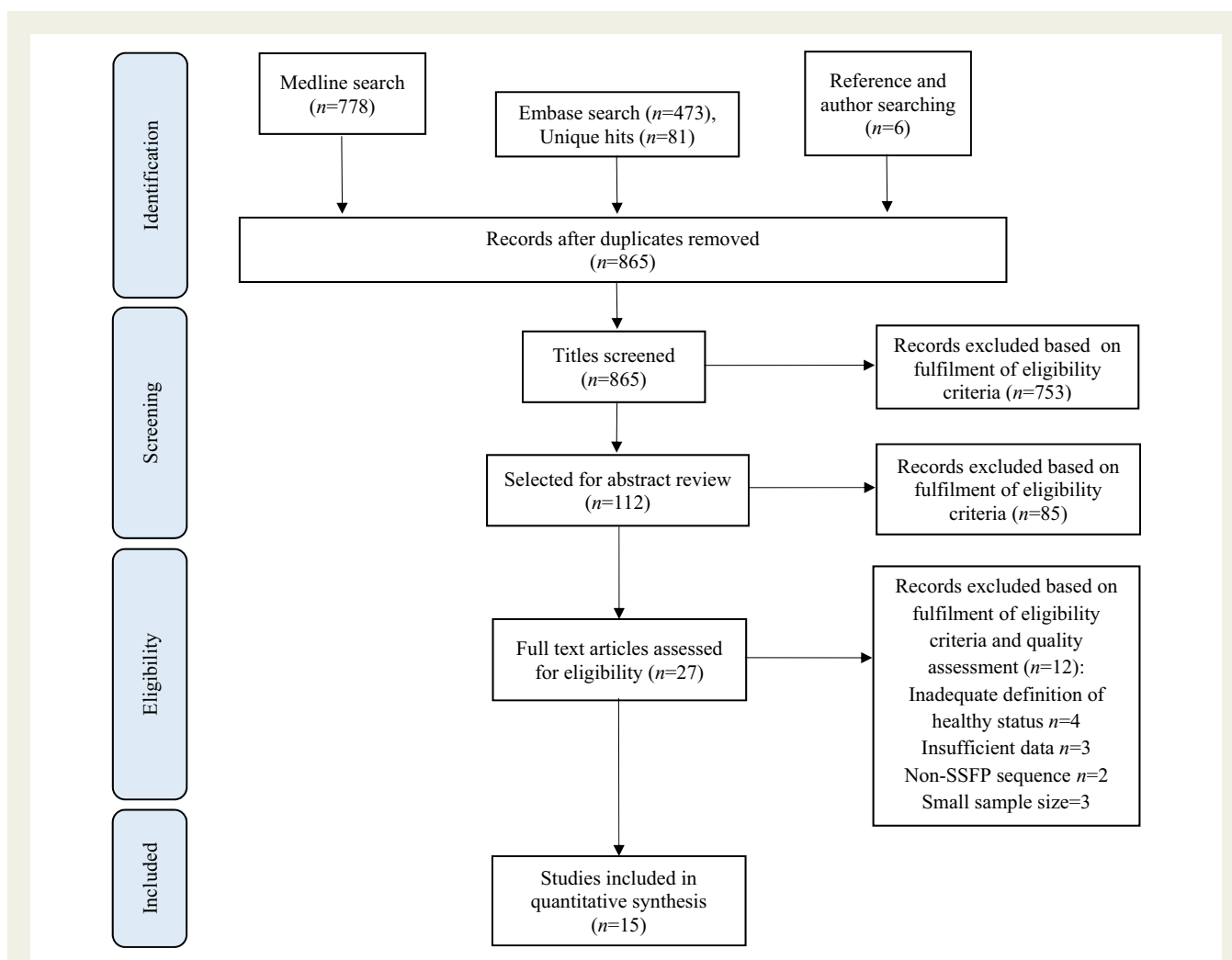


Figure 1 Flow diagram summarizing flow of information through different phases of the systematic review.

required LV quantification to be made using short-axis cine images using a predefined standard operating procedure for image acquisition and analysis. Studies selected for quantitative analysis were required to report sex-stratified LVM, LVEDV, LVESV, and LVEF in a manner where mean and standard deviation values in indexed formats [indexed to body surface area (BSA), denoted by $\bar{}$] could be extracted.

Search strategy

Z.R.E. and A.A.M.K. independently searched Ovid Medline (1946–April 2019) and Embase electronic databases. Relevant subject headings were used to conduct the search using MeSH terms (Medical Subject Headings) for Medline and the equivalent, Emtree, for Embase. Subject headings and their ‘trees’ were examined, and relevant subheadings were selected, related terms were included using the explode command (Supplementary data online, Table S1). Search terms were combined using Boolean operators. Selected terms were included in the search as keywords. We performed separate keyword search of titles and abstracts to ensure capture of newer publications not yet incorporated into MeSH/Emtree classifications. The final output was limited to studies in adults (>18 years-old) and in the English language.

Study selection

Study selection was through a process of title screening, abstract review, and full-text review carried out independently by A.A.M.K. and Z.R.E. At each iteration, results were merged and duplicates removed. Further studies were identified through reference and author searching. Decision for study eligibility was based on predefined selection criteria. In case of disagreement, decisions were taken through discussion after review of full text and mediation by M.Y.K.

Quality assessment

As this review was not based on intervention-outcome studies, existing quality assessment tools were not entirely applicable. We therefore designed a quality assessment protocol tailored to our purpose based on revised elements from the ROBINS-I (The Risk Of Bias In Non-randomized Studies-of Interventions) and QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) assessment tools.^{13,14}

Data extraction

Mean, standard deviation, and sample size for sex-stratified LVM indexed (LVMi), LVESV indexed (LVESVi), LVEDV indexed (LVEDVi), and LVEF

Table 1 Summary of studies selected for inclusion in the meta-analysis

Author, year of publication	Single/multicentre	Sample size, male:female	Age ^a	Ethnicity	Scanner vendor	Field strength (T)	Analysis software	Contouring technique	Papillary muscle included/excluded from LVM
Bulow et al., ¹⁶ 2018	Single	n = 617	Men (43) Women (45)	Caucasian	Siemens Magnetom	1.5	QMass MR, Medis	Manual	Included
Lei et al., ¹⁷ 2017	Single	n = 120	20–83	Chinese	Siemens Magnetom	3	QMass MR, Medis	Manual	Excluded
Petersen et al., ¹⁸ 2017	Single	n = 800	45–74	Caucasian	Siemens Magnetom	1.5	CMR42, Circle Cardiovascular Imaging	Manual	Excluded
Aquaro et al., ¹⁹ 2017	Multi	n = 255	15–80	Caucasian	Multi-vendor	1.5	Multiple	Manual	Included
Le et al., ²⁰ 2016	Single	n = 180	20–69	Singaporean-Chinese	Philips Ingenia	3	CMR42, Circle Cardiovascular Imaging	Not stated	Included
Li et al., ²¹ 2016	Single	n = 90	40–65	Chinese	Philips Achieva	3	Philips Medical Systems, Philips	Manual	Included
Le Ven et al., ²² 2016	Single	n = 434	18–35	Caucasian	Philips Achieva	1.5	CMR42, Circle Cardiovascular Imaging	Semi-automated	Included
Yeon et al., ²³ 2015	Single	n = 852	Men (61) Women (62)	Caucasian	Philips Gyroscan	1.5	EasyVision 5.1, Philips	Manual	Excluded
Macedo et al., ²⁴ 2013	Multi	n = 107	20–80	Brazilian	Philips Achieva	1.5	Multiple	Semi-automated	Included
Chang et al., ²⁵ 2012	Single	n = 124	20–70	Korean	Siemens Magnetom	1.5	Argus, Siemens	Manual	Excluded
Teo et al., ²⁶ 2008	Single	n = 60	51	Non-aboriginal Australian (Caucasian)	Siemens Sonata	1.5	Argus, Siemens	Manual	Included
Maceira et al., ²⁷ 2006	Single	n = 120	20–80	Caucasian	Siemens Sonata	1.5	CMRtools, Cardiovascular Imaging Solutions	Semi-automated	Included
Nikitin et al., ²⁸ 2006	Single	n = 95	22–91	Caucasian	General Electric SignaCV/i	1.5	MRI-MASS, Medis	Semi-automated	Included
Hudsmith et al., ²⁹ 2005	Single	n = 108	21–68	Caucasian	Siemens Sonata	1.5	Argus, Siemens	Manual	Included
Alfakih et al., ³⁰ 2003	Single	n = 60	20–65	Caucasian	Philips Intera	1.5	MRI-MASS, Medis	Manual	Included

LVM, left ventricular mass; T, Tesla; n denotes total sample size available for analysis.

^aAge: range, or mean (years).

Table 2 Pooled mean left ventricular parameters with sex stratification and expression of between-study and subgroup heterogeneity

		Mean (95% CI)	Between study heterogeneity				Subgroup differences (men vs. women)		
			Q statistic	τ^2	I^2	P-value	Mean difference	Q statistic	P-value
LVEDVi (mL/m ²)	Men	77.4 (73.7–81.1)	655.4	51.1	97.9%	8.5×10^{-131}	6.1	6.0	0.01
	Women	71.3 (68.1–74.5)							
LVESVi (mL/m ²)	Men	28.4 (25.4–31.3)	1196.4	29.1	99.0%	1.0×10^{-248}	4.0	4.5	0.03
	Women	24.4 (22.3–26.5)							
LVMI (g/m ²)	Men	60.5 (57.3–63.7)	773.9	18.2	98.2%	4.3×10^{-156}	12	33.3	7.8×10^{-9}
	Women	48.5 (45.9–51.0)							
LVEF (%)	Men	63.8 (61.6–66.1)	1272.2	19.2	98.9%	5.1×10^{-263}	-1.5	1.0	0.3
	Women	65.3 (63.5–67.1)							

Significance level is set at *P*-value <0.05. Random effects model is used for assessment of subgroups and between-study heterogeneity. CI, confidence interval; LVEDVi, left ventricular end-diastolic volume indexed to body surface area; LVEF, left ventricular ejection fraction; LVESVi, left ventricular end-systolic volume indexed to body surface area; LVMI, left ventricular mass indexed to body surface area.

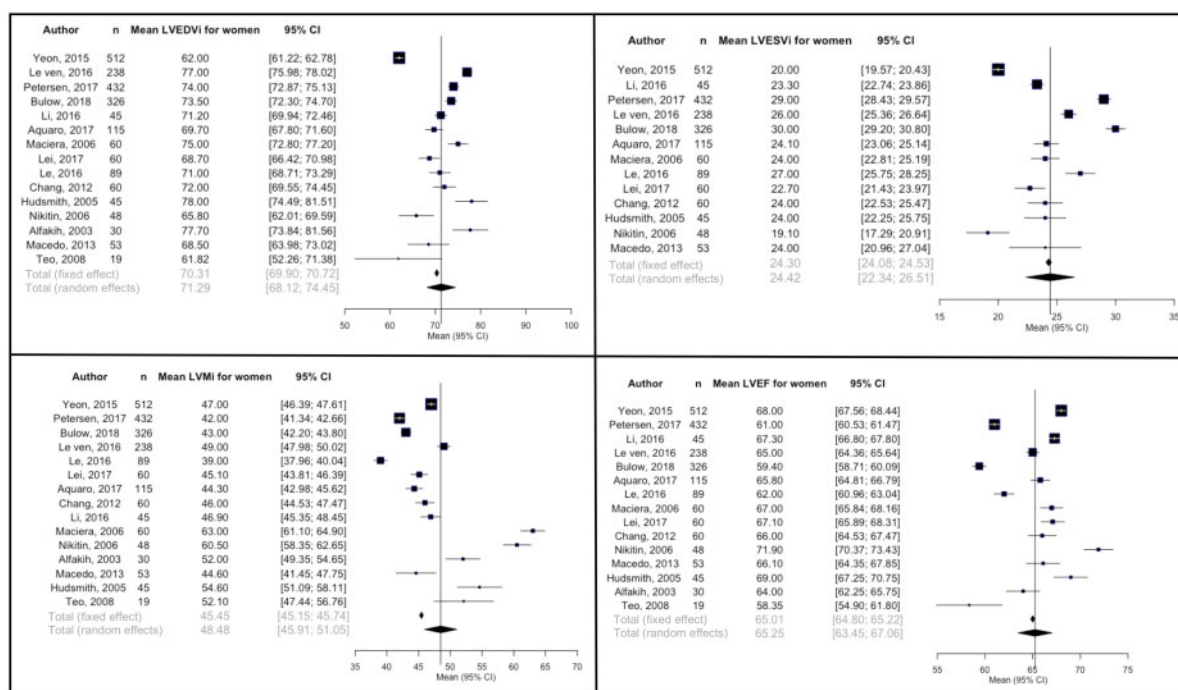


Figure 2 Forest plots of left ventricular parameters indexed to body surface area for women.^a Both fixed effect and random effects estimates are presented. The vertical reference line corresponds to random effects pooled mean estimate. CI, confidence interval; LVEDVi, left ventricular end-diastolic volume indexed to body surface area (mL/m²); LVEF, left ventricular ejection fraction (%); LVESVi, left ventricular end-systolic volume indexed to body surface area (mL/m²); LVMI, left ventricular mass indexed to body surface area (g/m²).

were extracted from individual studies. Data extraction was carried out independently by Z.R.E. and A.A.M.K. and cross-checked by Z.R.E.

Statistical analysis

Statistical analysis was with the 'meta: General package for meta-analysis' package on the R studio platform [R Core Team (2018). R: A language

and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 27 April 2020, date last accessed].¹⁵ We calculated pooled age, sex, and ethnicity-specific ranges for LV parameters indexed to BSA. We used random effects meta-analysis of single means with inverse variance weighting to calculate pooled values. Between-study heterogeneity was assessed with τ^2 , I^2 , Q statistic, and the related *P*-value. For subgroup analysis, mean

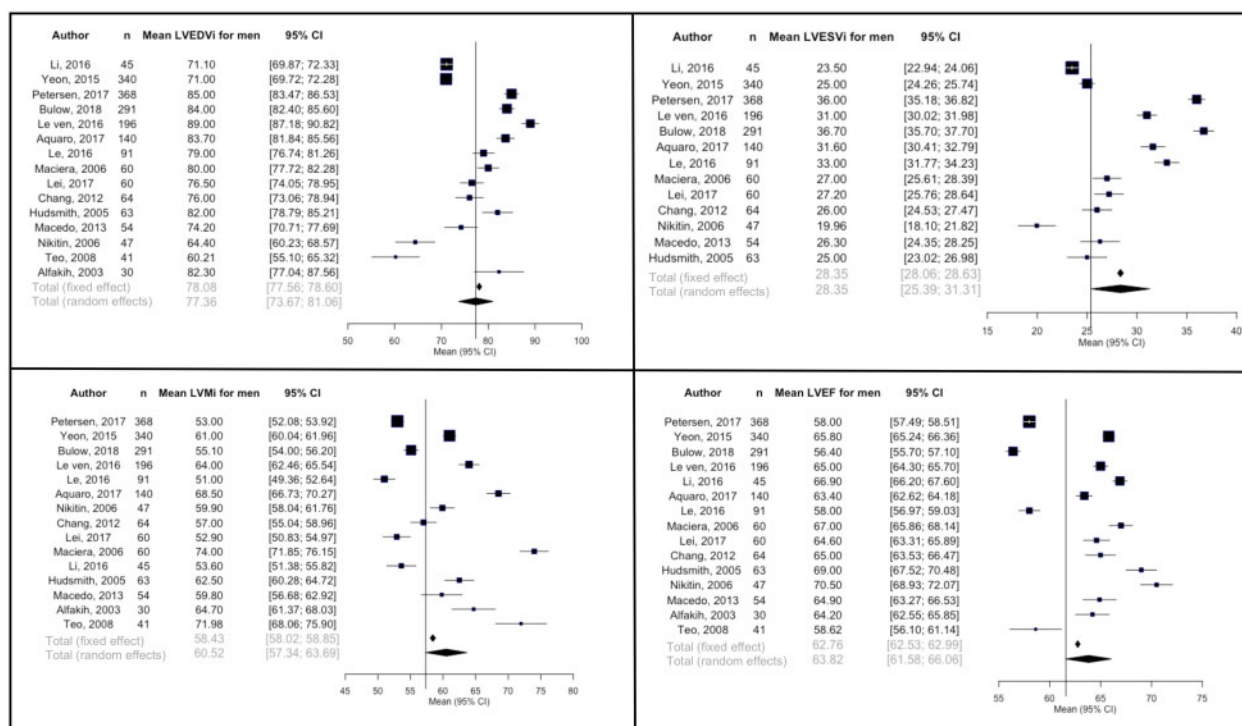


Figure 3 Forest plots of left ventricular parameters indexed to body surface area for men.^a Both fixed effect and random effects estimates are presented. The vertical reference line corresponds to random effects pooled mean estimate. CI, confidence interval; LVEDVi, left ventricular end-diastolic volume indexed to body surface area (mL/m^2); LVEF: left ventricular ejection fraction (%); LVESVi, left ventricular end-systolic volume indexed to body surface area (mL/m^2); LVMI, left ventricular mass indexed to body surface area (g/m^2).

difference (MD), Q statistic, and *P*-values are presented. We performed sensitivity analysis with the following variables: scanner vendor, magnet strength, post-processing analysis software, papillary muscles contouring (inclusion/exclusion in LVM). To assess the impact of the larger studies in the meta-analysis on the overall results, we display results for both fixed and random effects models in the figures. A large study with extreme results would lose influence under the random effects model. Our analyses demonstrate similar estimates from fixed and random effects models; therefore we conclude that variations in study sample size are not having a disproportionate impact on the results. For further illustration, we performed sensitivity analysis with exclusion of the two largest studies from the pooled estimates, which also did not significantly alter the pooled estimates.

Results

Systematic review

Our approach is summarized in the PRISMA flow diagram (Figure 1). Combined Ovid Medline and Embase searches yielded 859 unique hits; 6 additional citations were obtained from cross-referencing and author searches. After title screening, 112 citations were deemed potentially relevant and selected for abstract review. From these, 27 papers were selected for full-text review based on fulfilment of the inclusion criteria. A further 12 studies were excluded after examination of the full text based on quality assessment and consideration

of inclusion criteria. Fifteen studies were selected for inclusion in the meta-analysis. Of these, two did not report LVESV, therefore, 13 studies are included in analysis for this parameter.

Quality assessment

Pertinent quality indicators were systematically assessed for studies selected for full-text review. There were differences in the definition of 'healthy status' with variable use of clinical assessment, blood tests, and non-invasive tests (echocardiography and electrocardiogram) to exclude disease. There were also variations in the number of readers and reports of inter-/intra-observer variability. Overall, the studies included in the meta-analysis are of high quality with clearly defined study objectives and imaging protocols (Supplementary data online, Table S2).

Summary of selected studies

Overall 2132 women and 1890 men from 15 studies published between 2003 and 2018 are included in the analysis (Table 1). The age range is between 20 and 91 years. There are five studies from non-Caucasian cohorts: two from Chinese populations,^{17,21} and one study each from Singaporean-Chinese,²⁰ Korean,²⁵ and Brazilian²⁴ cohorts. There are 10 studies from Caucasian populations.^{16,18,19,22,23,26–30} Both the Chinese studies and the study from Singapore use a 3-T scanner, all others use 1.5-T scanners. Scanners included several Siemens and Philips models; one of the earlier studies used a General

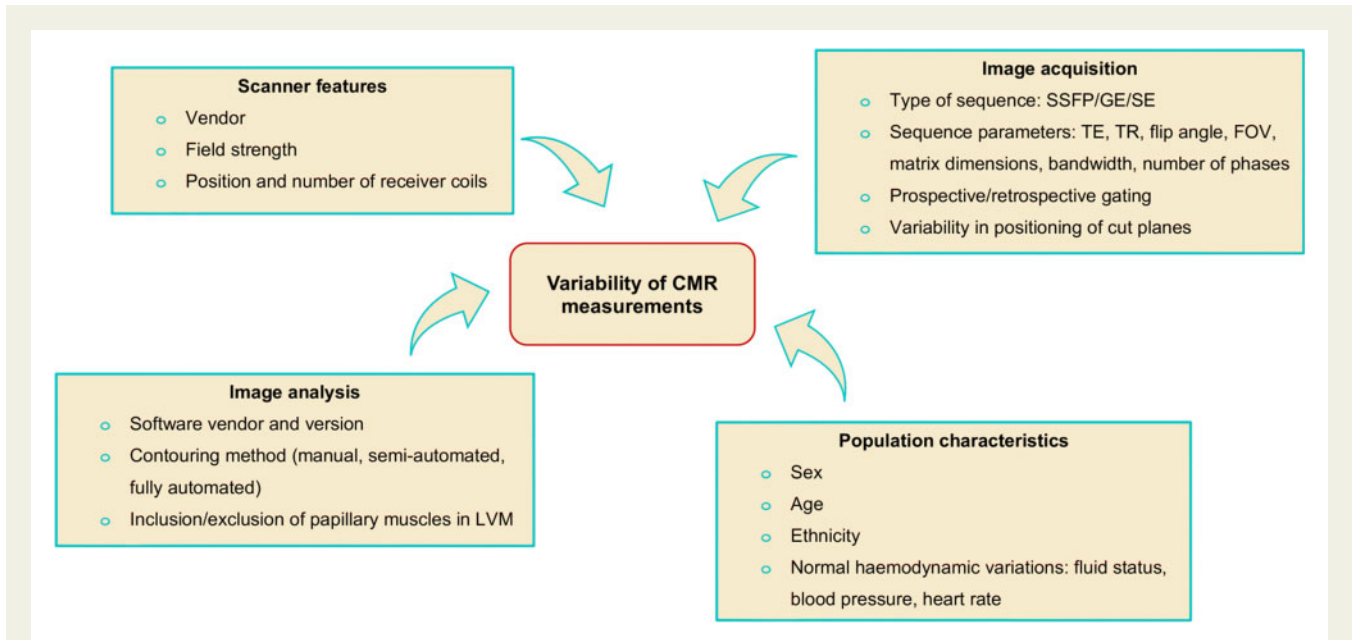
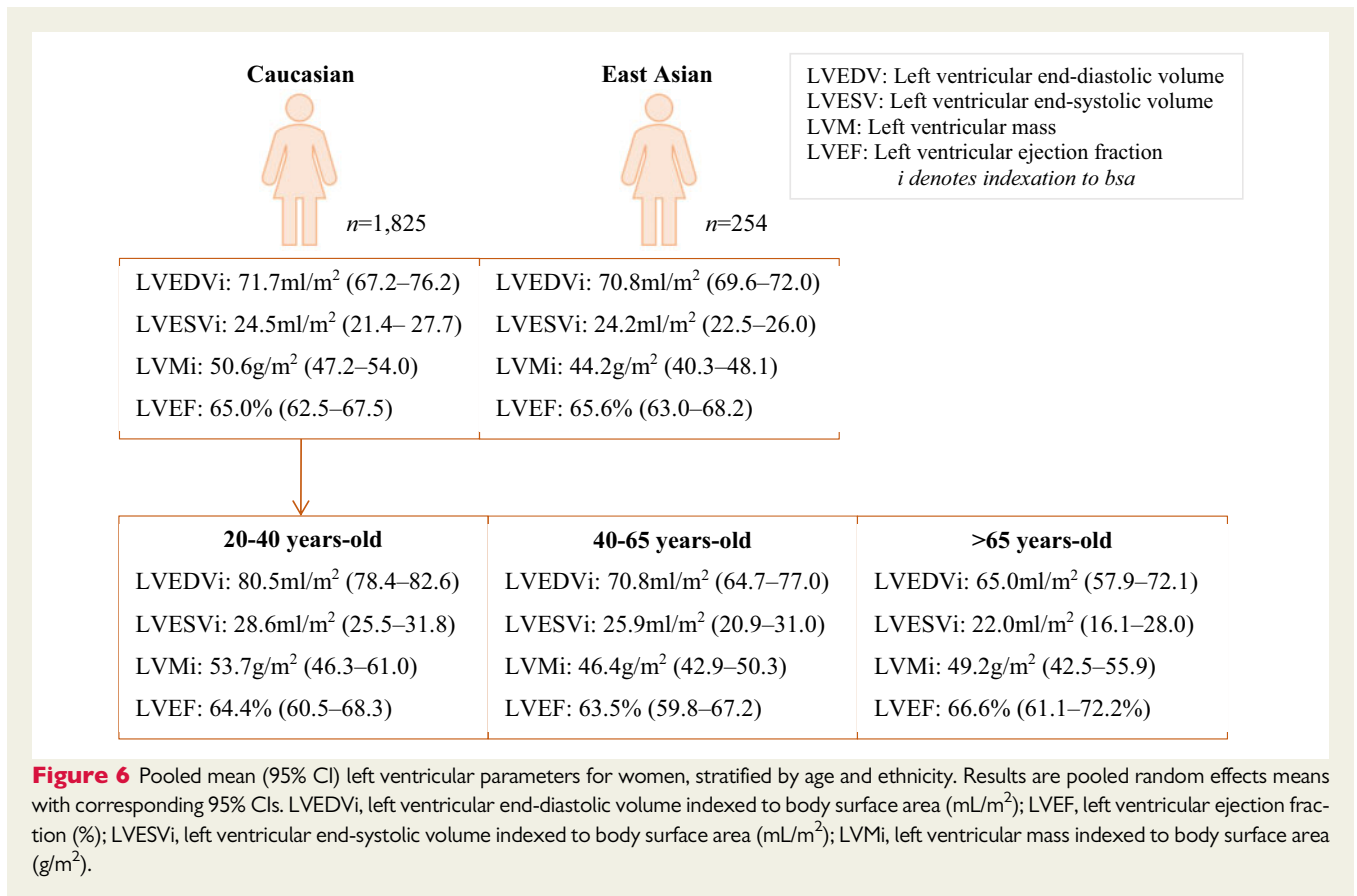


Figure 4 Potential sources of variability in cardiac magnetic resonance measurements. CMR, cardiac magnetic resonance; FOV, field of view; GE, gradient echo; LVM, left ventricular mass; SE, spin echo; SSFP, steady state free precession; TE, echo time; TR, repetition time.

Caucasian <i>n</i> =1,576		East Asian <i>n</i> =260			
LVEDVi: 78.4ml/m ² (73.5–83.2)	LVEDVi: 75.6ml/m ² (71.5–79.7)	LVEDV: Left ventricular end-diastolic volume LVESV: Left ventricular end-systolic volume LVM: Left ventricular mass LVEF: Left ventricular ejection fraction <i>i</i> denotes indexation to <i>bsa</i>			
LVESVi: 29.1ml/m ² (25.2–33.0)	LVESVi: 27.4ml/m ² (23.0–31.9)				
LVMi: 63.4g/m ² (59.3–67.4)	LVMi: 53.6g/m ² (51.0–56.2)				
LVEF: 63.8% (60.9–66.7)	LVEF: 63.6% (59.3–67.9)				
20–40 years-old		40–65 years-old		>65 years-old	
LVEDVi: 88.4ml/m ² (85.8–90.9)	LVEDVi: 79.8ml/m ² (73.5–86.2)	LVEDVi: 71.4ml/m ² (63.3–79.5)			
LVESVi: 33.7ml/m ² (29.5–38.0)	LVESVi: 30.6ml/m ² (24.7–36.5)	LVESVi: 26.0ml/m ² (19.4–32.5)			
LVMi: 67.3g/m ² (60.1–74.5)	LVMi: 63.0g/m ² (57.1–69.0)	LVMi: 59.8g/m ² (53.0–66.7)			
LVEF: 62.7% (57.8–67.5)	LVEF: 62.5% (58.6–66.5)	LVEF: 64.5% (59.0–70.0)			

Figure 5 Pooled mean (95% CI) left ventricular parameters for men, stratified by age and ethnicity. Results are pooled random effects means with corresponding 95% CIs. LVEDVi, left ventricular end-diastolic volume indexed to body surface area (mL/m²); LVEF, left ventricular ejection fraction (%); LVESVi, left ventricular end-systolic volume indexed to body surface area (mL/m²); LVMi, left ventricular mass indexed to body surface area (g/m²).



Electric (GE) scanner. Various versions of a wide range of post-processing software packages were used for endocardial contouring. Contouring technique was either manual or semi-automated with manual edits. Eleven studies included papillary muscles in the LVM, the remainder as part of the blood pool.

Meta-analysis

Stratification by sex

Results for sex-stratified analyses are summarized in *Table 2*, *Figures 2* and *3*, and *Panel A* of the graphical abstract. Compared to women, men had significantly larger LVEDVi (MD = 6.1 mL/m², *P*-value = 0.014), LVESVi (MD = 4.0 mL/m², *P*-value = 0.033), and LVMi (MD = 12.0 g/m², *P*-value = 7.8 × 10⁻⁹). LVEF was not significantly different between men and women (MD = -1.5%, *P*-value 0.33). In both men and women, there was significant between-study heterogeneity for all LV parameters (*I*² > 97% for all).

Stratification by age and sex

Three age categories were created to represent young (20–40 years), middle-aged (40–65 years), and older (≥65 years) adults. These age cut-offs allowed inclusion of the largest possible pooled sample from all studies. Age- and sex-stratified results are presented in *Supplementary data* online, *Table S3*. Both men and women had significantly larger LVEDVi in younger age (*Supplementary data* online, *Figure S1*) with similar magnitude of difference (20–40 years vs. ≥65 years: women MD = 14.0 mL/m², men MD = 14.7 mL/m²). A

trend for larger LVESVi in younger age is observed for both men and women but is not statistically significant in either. There were non-significant trends towards greater LVMi in younger and higher LVEF in older individuals. The data available did not permit analysis with age as a continuous measure or with more granular age bands. Between-study heterogeneity remained high after sex and age stratification.

Stratification by sex and ethnicity

Pooled values were calculated for two ethnicity categories: East Asian (Chinese, Singaporean-Chinese, Korean) and Caucasian (including non-Aboriginal Australian). East Asian men and women had significantly lower LVMi compared to Caucasians (women: MD = 6.4 g/m², *P*-value = 0.016; men: MD = 9.8 g/m², *P*-value = 6.7 × 10⁻⁵), this difference was more consistent and of greater magnitude in men (*Supplementary data* online, *Figure S2*). Further comparison was made between pooled values for Caucasians, East Asians, and the one Brazilian cohort. Again, significant subgroup differences were observed in LVMi for both men and women. Brazilian men and women had greater LVMi than East Asians, but lower values than Caucasians. There were no significant ethnic differences in any of the other LV parameters (*Supplementary data* online, *Table S4*). We present pooled sex-stratified results for Caucasians and East Asians with addition of age stratification for Caucasians (*Figures 5* and *6*) We are unable to provide pooled age- and sex-stratified results for East Asians due to variation in age bands and reporting of

stratified results in the original studies. There was high statistical heterogeneity between studies despite these population stratifications.

Sensitivity analyses

To explore potential technical sources of between-study heterogeneity, sensitivity analyses were performed with the following variables: scanner vendor, field strength, post-processing software, and papillary muscle contouring (included vs. excluded from LVM) (Supplementary data online, Table S5). Studies including papillary muscles as part of the LVM reported significantly higher LVM for both men (MD = 7.1 g/m², *P*-value = 0.017) and women (MD = 6.0 g/m², *P*-value = 0.029). Despite stratification for sex and contouring methodology, heterogeneity between studies remained high, with greater heterogeneity for studies contouring papillary muscles as part of LVM (Supplementary data online, Figure S3). The post-processing software used for contouring also impacted results with Argus software from Siemens Medical yielding significantly smaller LVESVi and higher LVMI in comparison to other post-processing tools. We also note a significant relationship between lower LVMI and 3-T field strength scanners. Limited samples and significant methodological heterogeneity at all levels meant that pooling of results with stratification for multiple technical and population-related factors was not possible.

Discussion

Summary of findings

We present the first formal systematic review and meta-analysis of CMR normal reference ranges incorporating results from 1890 men and 2132 women from 15 studies. Pooled results demonstrate significant differences in LV parameters by sex, age, and ethnicity. Compared to women, men had larger cavity volumes and greater LVMI. Younger individuals had larger LV volumes, higher LVMI and lower LVEF in comparison to older ages. Individuals with East Asian ancestry had lower LVMI in comparison to Caucasians. Between-study heterogeneity was high for all parameters despite stratification for population-related factors. Sensitivity analyses identified differences in contouring methodology, post-processing software, and magnet field strength as potential significant contributors to the observed between-study heterogeneity. Limited sample sizes from existing results and methodological variation at all levels precludes recommendation of robust unified reference ranges from this analysis.

Comparison with previous literature

The observed sex, age, and ethnic differences in LV measures are consistent with previous reports using cardiac computed tomography, echocardiography, and gradient echo CMR.^{31–35} Echocardiography studies report important differences in cardiac morphology of healthy individuals of South Asian and Afro-Caribbean ethnicity in comparison to Caucasians.^{32,36,37} Furthermore, there are reports of differential significance of alterations in LV parameters in different ethnic populations. For instance, Akintoye et al. report greater prognostic utility of LVMI for predicting cardiovascular events for Chinese and Hispanic populations in comparison to non-Hispanic Whites.³⁸ Similarly, there are reports of

significant ethnic differences in ventricular remodelling in response to important cardiovascular risk factors such as hypertension.³⁶ As ethnic variation exist for LV parameters, it is possible that there are also ethnic differences in the morphology of other cardiac chambers (right ventricle, atria). Whilst in recent years, there have been reports of CMR reference ranges from several non-Caucasian cohorts, data from a wide range of ethnicities remains absent, as such, our understanding of ethnic differences in CMR derived measures of cardiac morphology remains incomplete.

In addition to the expected variations by population-related factors, we identified important technical sources of heterogeneity. We identified magnet strength (3 T vs. 1.5 T) as a significant source of variation, in particular lower LVMI reported by the studies using 3 T scanners. Certainly, it is conceivable that higher spatial resolution produced by expert programming of pulse sequences with 3 T scanners provides superior endocardial border definition and thus more accurate contouring of the LV endocardium with exclusion of an intracavity trabecular layer that may be included within LVM at lower spatial resolutions. However, there are other factors that need consideration. For instance, the 3 T studies are all more recent publications (2016 onwards), image analysis for these studies has been conducted using modern post-processing software perhaps allowing for more accurate endocardial border contouring in comparison to older studies. There are also important population factors—all the studies with 3 T scanners are from East Asian cohorts, whereas all studies on Caucasians are with 1.5 T scanners. With the presence of multiple overlapping variables, it is impossible to isolate definitively the effect of 3 T vs. 1.5 T in this study. Previous studies dedicated to comparison of LV measures at 3 T vs. 1.5 T have not shown significant differences between the two.³⁹ On balance, our judgement is that the observed differences are more likely related to ethnic differences with perhaps a smaller contribution from the various technical sources of variation.

Consistent with previous reports, we identified differences in endocardial contouring as a significant source of variation.^{40,41} There was greater heterogeneity between studies that included papillary muscles within LVM compared to those that did not, perhaps reflecting difficulties in reproducibly tracing the irregular geometry of papillary muscles. Previous studies report similar variations with the potential for clinically important differences in the assessment of relevant pathologies such as hypertrophic cardiomyopathy and Fabry's disease.^{42,43} Whilst other sources of technical variation do exist and perhaps have a cumulative effect, it does seem that contouring technique is the most important. Interestingly, a small study of variation of CMR derived LV measures from the use of different software packages demonstrated no significant variation from the software programmes with the application of a standardized contouring protocol and a single scanner vendor.⁴⁴ This observation suggests that the variability in LV quantification measures may be eliminated, or certainly reduced, by development of uniform contouring practices.

Our analysis suggests that the high between-study heterogeneity is a result of cumulative effects from multiple population-related and technical sources of variation. We were unable to significantly reduce between-study heterogeneity through stratification by one or two factors and the sample size does not permit meaningful sub-analysis by greater number of variables.

Relevance for clinical practice

Our results show that for both men and women, healthy young adults have on average 21% larger LVEDVi compared to healthy older adults (age < 40 years vs. age > 65 years: women MD = 14.0 mL/m², max difference = 24.3 mL/m²; men MD = 14.7 mL/m², max difference = 26.0 mL/m²). Whilst specific recommendations for age-correction cannot be made, reporting cardiologists should consider this level of variation when applying reference ranges to individuals outside represented age groups. Similar considerations should be made regarding ethnicity. Our findings show lower LVMI in East Asians compared to Caucasians with mean percentage difference of 18% and 15% in men and women, respectively (women: MD = 6.4 g/m², max difference = 13.7 g/m²; men: MD = 9.8 g/m², max difference = 16.4 g/m²). These differences can be clinically important. For example, consider an East Asian man with LVMI of 63 g/m²—whilst this is average for a Caucasian population, it is well above the upper limit of normal for Asian cohorts (56.2 g/m²). Where possible, ethnicity-specific reference ranges should be used. Differences produced by technical factors, in particular, contouring methodology should also be considered. For instance, our findings suggest approximately 13% greater LVMI for both men and women when contouring includes papillary muscles within LVM.

Whilst CMR remains the reference standard for LV quantification, the results must be interpreted with consideration of age, sex, and ethnic differences. In addition, there are multiple technical sources of variation that may result in clinically important differences in reported values. Considering the high statistical heterogeneity between studies and the importance of technical sources of variation, we would recommend use of reference ranges that most resembles one's own clinical practice in terms of image acquisition, analysis, and population. In cases of variation in practice from the reference range of choice, it is possible to making approximate corrections using the calculations provided here.

Directions for future work

This work highlights the need for richer reference datasets with attention to incorporation of data from different ethnic groups and wider spectrum of ages. The lack of published data from ethnicities with known important differences in cardiac morphology is a significant limitation of existing literature. We should aim for development of reference ranges that are fully stratified by age, sex, and ethnicity. It is also important that we reduce the level of heterogeneity introduced by technical factors, with development of a unified approach to contouring methodology being a key step. However, it is difficult to make consensus recommendations at present, as it is not clear from existing literature, which contouring method best predicts clinical outcomes and/or discriminates disease. Therefore, prior to embarking on development of standardized approaches, research is needed into the prognostic and diagnostic value of different contouring methodologies. Finally, consideration of variability in cardiac morphometrics beyond traditional CMR indices is important for better understanding of differential disease patterns and risk profiles in different populations and would allow for deeper phenotyping of individuals and their disease susceptibilities.

Limitations

Our search strategy was thorough for published reports of CMR normal reference ranges; however, we did not seek results from

unpublished cohorts. Whilst this may have resulted in a larger sample size, quality control of data that has not been through a formal peer-review process is challenging and inclusion of such data may have compromised the quality of the study. There are important gaps in the literature with paucity of data for individuals in the youngest and oldest age categories and limited representation of non-Caucasian ethnicities. Our analysis reflects these gaps in published data. Whilst age, sex, and ethnicity explain part of the between-study heterogeneity, there are technical sources of variation that cannot be fully explored within the scope of this study (Figure 4).

Conclusions

There is significant heterogeneity in published CMR LV reference ranges. Age, sex, and ethnicity represent significant sources of variation and we should endeavour to develop reference ranges stratified to these parameters. Different endocardial contouring methodology, scanner magnet strength, and post-processing software all contribute to the observed variability. Due to multiple sources of heterogeneity, it is not possible to produce reliable normal ranges across a wide age range, by sex or ethnicity from existing reports. Wider representation of different populations and standardization of image analysis is urgently needed to establish such reference distributions, and thus ensure global comparability of CMR measures.

Data availability statement: This review uses summary data from cited publications, which are retrievable by referring to the original manuscripts.

Supplementary data

Supplementary data are available at *European Heart Journal - Cardiovascular Imaging* online.

Funding

This work was supported by a British Heart Foundation Clinical Research Training Fellowship (FS/17/81/33318 to Z.R.E.) and was supported by a Wellcome Trust Research Training Fellowship (203553/Z/16/Z to N.A.). S.E.P. acknowledges support from the National Institute for Health Research (NIHR) Cardiovascular Biomedical Research Centre at Barts NHS Trust and has received funding from the European Union's Horizon 2020 research and innovation programme (825903). S.E.P. acts as a paid consultant to Circle Cardiovascular Imaging Inc., Calgary, Canada and Servier. N.C.H. has received consultancy, lecture fees, and honoraria from Alliance for Better Bone Health, Amgen, MSD, Eli Lilly, Servier, Shire, Radius Health, UCB, Consilient Healthcare, Kyowa Kirin, and Internis Pharma.

Conflict of interest: none declared.

References

- Gjesdal O, Bluemke DA, Lima JA. Cardiac remodeling at the population level—risk factors, screening, and outcomes. *Nat Rev Cardiol* 2011;**8**:673–85.
- Leonardi S, Raineri C, De Ferrari GM, Ghio S, Scelsi L, Pasotti M et al. Usefulness of cardiac magnetic resonance in assessing the risk of ventricular arrhythmias and sudden death in patients with hypertrophic cardiomyopathy. *Eur Heart J* 2009;**30**: 2003–10.
- Bluemke DA, Kronmal RA, Lima JAC, Liu K, Olson J, Burke GL et al. The relationship of left ventricular mass and geometry to incident cardiovascular events. *J Am Coll Cardiol* 2008;**52**:2148–55.

4. Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *N Engl J Med* 1990;**322**:1561–6.
5. Bristow MR, Saxon LA, Boehmer J, Krueger S, Kass DA, De MT et al. Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N Engl J Med* 2004;**350**:2140–50.
6. Taniguchi T, Morimoto T, Shiomi H, Ando K, Kanamori N, Murata K et al. Prognostic impact of left ventricular ejection fraction in patients with severe aortic stenosis. *JACC Cardiovasc Interv* 2018;**11**:145–57.
7. McMurray J. Systolic heart failure. *N Engl J Med* 2010;**362**:228–38.
8. Aurigemma GP, Gaasch WH. Diastolic heart failure. *N Engl J Med* 2004;**351**:1097–105.
9. Petersen SE, Khanji MY, Plein S, Lancellotti P, Bucciarelli-Ducci C. European Association of Cardiovascular Imaging expert consensus paper: a comprehensive review of cardiovascular magnetic resonance normal values of cardiac chamber size and aortic root in adults and recommendations for grading severity. *Eur Heart J Cardiovasc Imaging* 2019;**44**:1–11.
10. Kishi S, Reis JP, Venkatesh BA, Gidding SS, Armstrong AC, Jacobs DR et al. Race-ethnic and sex differences in left ventricular structure and function: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *J Am Heart Assoc* 2015;**4**:e001264.
11. Strait JB, Lakatta EG. Aging-associated cardiovascular changes and their relationship to heart failure. *Heart Fail Clin* 2012;**8**:143–64.
12. Kawel-Boehm N, Maceira A, Valsangiacomo-Buechel ER, Vogel-Claussen J, Turkbey EB, Williams R et al. Normal values for cardiovascular magnetic resonance in adults and children. *J Cardiovasc Magn Reson* 2015;**17**:29.
13. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919.
14. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529.
15. Schwarzer G. Meta: an R package for meta-analysis. *R News* 2007;**7**:40–6.
16. Bülow R, Ittermann T, Dörr M, Poesch A, Langner S, Völzke H et al. Reference ranges of left ventricular structure and function assessed by contrast-enhanced cardiac MR and changes related to ageing and hypertension in a population-based study. *Eur Radiol* 2018;**28**:3996–4005.
17. Lei X, Liu H, Han Y, Cheng W, Sun J, Luo Y et al. Reference values of cardiac ventricular structure and function by steady-state free-precession MRI at 3.0T in healthy adult Chinese volunteers. *J Magn Reson Imaging* 2017;**45**:1684–92.
18. Petersen SE, Aung N, Sanghvi MM, Zemrak F, Fung K, Paiva JM et al. Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. *J Cardiovasc Magn Reson* 2017;**19**:18.
19. Aquaro GD, Camastra G, Monti L, Lombardi M, Pepe A, Castelletti S et al.; On behalf of the working group “Applicazioni della Risonanza Magnetica” of the Italian Society of Cardiology. Reference values of cardiac volumes, dimensions, and new functional parameters by MR: a multicenter, multivendor study. *J Magn Reson Imaging* 2017;**45**:1055–67.
20. Le TT, Tan RS, De Deyn M, Goh EPC, Han Y, Leong BR et al. Cardiovascular magnetic resonance reference ranges for the heart and aorta in Chinese at 3T. *J Cardiovasc Magn Reson* 2016;**18**:21.
21. Li CY, Gao BL, Guo FQ, Zhang XJ, Fan QY, Wu BL et al. Quantitative evaluation of left ventricular volume and function in middle-aged healthy Chinese people with 3 Tesla MRI. *J Magn Reson Imaging* 2016;**44**:1143–50.
22. Le Ven F, Bibeau K, De Larocheliere E, Tizon-Marcos H, Deneault-Bissonnette S, Pibarot P et al. Cardiac morphology and function reference values derived from a large subset of healthy young Caucasian adults by magnetic resonance imaging. *Eur Heart J Cardiovasc Imaging* 2016;**17**:981–90.
23. Yeon SB, Salton CJ, Gona P, Chuang ML, Blease SJ, Han Y et al. Impact of age, sex, and indexation method on MR left ventricular reference values in the Framingham Heart Study offspring cohort. *J Magn Reson Imaging* 2015;**41**:1038–45.
24. Macedo R, Fernandes JL, Andrade SS, Rochitte CE, Lima KC, Maciel AC et al. Morphological and functional measurements of the heart obtained by magnetic resonance imaging in Brazilians. *Arq Bras Cardiol* 2013;**101**:68–77.
25. Chang SA, Choe YH, Jang SY, Kim SM, Lee SC, Oh JK. Assessment of left and right ventricular parameters in healthy Korean volunteers using cardiac magnetic resonance imaging: change in ventricular volume and function based on age, gender and body surface area. *Int J Cardiovasc Imaging* 2012;**28**:141–7.
26. Teo KSL, Carbone A, Piantadosi C, Chew DP, Hammett CJK, Brown MA et al. Cardiac MRI assessment of left and right ventricular parameters in healthy Australian normal volunteers. *Heart Lung Circ* 2008;**17**:313–7.
27. Maceira AM, Prasad SK, Khan M, Pennell DJ. Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance. *J Cardiovasc Magn Reson* 2006;**8**:417–26.
28. Nikitin NP, Huan Loh P, De Silva R, Witte KKA, Lukaschuk EI, Parker A et al. Left ventricular morphology, global and longitudinal function in normal older individuals: a cardiac magnetic resonance study. *Int J Cardiol* 2006;**108**:76–83.
29. Hudsmith LE, Petersen SES, Francis JJM, Robson MDM, Neubauer S. Normal human left and right ventricular and left atrial dimensions using steady state free precession magnetic resonance imaging. *J Cardiovasc Magn Reson* 2005;**7**:775–82.
30. Alfakih K, Plein S, Thiele H, Jones T, Ridgway JP, Sivanathan MU. Normal human left and right ventricular dimensions for MRI as assessed by turbo gradient echo and steady-state free precession imaging sequences. *J Magn Reson Imaging* 2003;**17**:323–9.
31. Lin FY, Devereux RB, Roman MJ, Meng J, Jow VM, Jacobs A et al. Cardiac chamber volumes, function, and mass as determined by 64-multidetector row computed tomography: mean values among healthy adults free of hypertension and obesity. *JACC Cardiovasc Imaging* 2008;**1**:782–6.
32. Chahal NS, Lim TK, Jain P, Chambers JC, Koener JS, Senior R. Ethnicity-related differences in left ventricular function, structure and geometry: a population study of UK Indian Asian and European white subjects. *Heart BMJ* 2010;**96**:466–71.
33. Poppe KK; The EchoNoRMAL Collaboration. What is normal left ventricular ejection fraction? A global individual person data meta-analysis of the distribution across ethnicity, gender and age. *Eur Heart J* 2013;**34**:1938.
34. Buccheri S, Costanzo L, Tamburino C, Monte I. Reference values for real time three-dimensional echocardiography-derived left ventricular volumes and ejection fraction: review and meta-analysis of currently available studies. *Echocardiography* 2015;**32**:1841–50.
35. Natori S, Lai S, Finn JP, Gomes AS, Hundley WG, Jerosch-Herold M et al. Cardiovascular function in multi-ethnic study of atherosclerosis: normal values by age, sex, and ethnicity. *Am J Roentgenol* 2006;**186**:S357–65.
36. Chaturvedi N, Athanassopoulos G, McKeigue PM, Marmot MG, Nihoyannopoulos P. Echocardiographic measures of left ventricular structure and their relation with rest and ambulatory blood pressure in blacks and whites in the United Kingdom. *J Am Coll Cardiol* 1994;**24**:1499–505.
37. Poppe KK, Doughty RN, Gardin JM, Hobbs FDR, McMurray JVV, Nagueh SF et al. Ethnic-specific normative reference values for echocardiographic LA and LV size, LV mass, and systolic function: the EchoNoRMAL study. *JACC Cardiovasc Imaging* 2015;**8**:656–65.
38. Akintoye E, Mahmoud K, Shokr M, Sandio A, Mallikethi-Reddy S, Sheikh M et al. Racial/ethnic differences in the prognostic utility of left ventricular mass index for incident cardiovascular disease. *Clin Cardiol* 2018;**41**:502–9.
39. Hudsmith LE, Petersen SE, Tyler DJ, Francis JM, Cheng ASH, Clarke K et al. Determination of cardiac volumes and mass with FLASH and SSFP cine sequences at 1.5 vs. 3 Tesla: a validation study. *J Magn Reson Imaging* 2006;**24**:312–18.
40. Miller CA, Jordan P, Borg A, Argyle R, Clark D, Pearce K et al. Quantification of left ventricular indices from SSFP cine imaging: impact of real-world variability in analysis methodology and utility of geometric modeling. *J Magn Reson Imaging* 2013;**37**:1213–22.
41. Suinesiaputra A, Bluemke DA, Cowan BR, Friedrich MG, Kramer CM, Kwong R et al. Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours. *J Cardiovasc Magn Reson* 2015;**17**:63.
42. Kozor R, Callaghan F, Tchan M, Hamilton-Craig C, Figtree GA, Grieve SM. A disproportionate contribution of papillary muscles and trabeculations to total left ventricular mass makes choice of cardiovascular magnetic resonance analysis technique critical in Fabry disease. *J Cardiovasc Magn Reson* 2015;**17**:22.
43. Gommans DHF, Bakker J, Cramer GE, Verheugt FWA, Brouwer MA, Kofflard M. Impact of the papillary muscles on cardiac magnetic resonance image analysis of important left ventricular parameters in hypertrophic cardiomyopathy. *Neth Heart J* 2016;**24**:326–31.
44. Zange L, Muehlberg F, Blaszczyk E, Schwenke S, Traber J, Funk S et al. Quantification in cardiovascular magnetic resonance: agreement of software from three different vendors on assessment of left ventricular function, 2D flow and parametric mapping. *J Cardiovasc Magn Reson* 2019;**21**:12.